AL Esraa

University College Journal for Engineering Sciences

A Periodical Comprehensive Refereed Scientific Journal - Issue by: AL-Esraa University, Baghdad - Iraq

ISSN: 2709 - 7145.
E-ISSN: 2790 - 7732
The number of deposit at books and documents house,(2445), Baghdad,Iraq (2020).

جامعة الإسراء

(سُبْحَانَ الَّذِي أَسْرَىٰ بِعَبْدِهِ لَيْلًا مِّنَ الْمَسْجِدِ الْحَرَامِ إِلَى الْمَسْجِدِ الْأَقْصَى)

E U

AL- Esraa University

1432 هـ – 2011 م

Vol.(8), No.(13) June - 2026

Editor in Chief

- **Prof. Dr. Abdul- Razaq J. Al- Majidi**, Chancellor of Al-Esraa University / Iraq.

Editorial Manager

- **Assist. Prof. Dr. Ihsan A. S. Al-Shaarbaf** College of Eng. / Al-Esraa University / Iraq.

Editorial Broad

- **Prof. Dr. Mousa A. Al-Mousawy** Advisor/ Ministry of Higher Education and Scientific Research/ Iraq.
- **Prof. Dr. Abbas M. Al-Bakry** Chancellor of IT University / Iraq.
- **Prof. Dr. Thamir K. Mahmoud** College of Eng. Technology, Al-Esraa University / Iraq.
- **Prof. Dr. Hussain Al-Rizzo** Arkansas University, Elec. Eng./ U.S.A.
- **Prof. Dr. Riadh S. Al-Mahaidi** Swinburne University of Technology / Australia.
- **Prof. Dr. MuthannaH. Al-Dahhan** Missouri University, Mech. Eng. U.S.A.
- **Prof. Dr. Ramzi M. Mahmoud** Bensalvania University ,Civil. Eng./ U.S.A.
- **Prof. Dr. Abdulrazzak T. Ziboon** College of Eng. Technology, Al-Esraa University/ Iraq
- **Assist. Prof. Dr. Kadhum Aboud Al-Majidi** Al- Mustansiriya Univ. / Iraq.
- **Assist. Prof. Dr. Reiadh A. Al-Mosawy** College of Eng. Al-Esraa Univ. / Iraq.

- **Assist. Prof. Dr. Sabah N. Hassan** College of Eng. Al-Esraa Univ. / Iraq.
- **Assist. Prof. Dr. Mohammed G. Sabri** Sci. Affair Dept., Al-Esraa Univ./ Iraq
- **Assist. Prof. Dr. Abdul-Nasser A. Hafidh** Ministry of Higher Educ. and Sci. Res./ Iraq
- **Lecturer Dr. Ayad Ahmed Al-Taweel** Al-Esraa Univ./ Iraq
- **Lecturer Dr. Haider Sadiq Al-Aasam** College of Eng. Technology, Al-Esraa Univ./ Iraq
- **Associate Prof. Mahnoosh Biglari** Razi Univ., Civil Eng./ Iran
- **Associate Prof. Hassan Moradi** Razi Univ., Elec. Eng./ Iran
- **Associate Prof. Iman Ashayeri** Razi Univ., Civil Eng./ Iran

Language Consultant

- **Prof. Dr. Ghaleb F. Al-Matlabi** Al-Esraa Univ. / Iraq
- **Assist. Prof. Dr. Saad F. Al-Hassani** Al-Esraa Univ. / Iraq

Intellectual Integrity

- **Assist. Prof. Dr. Akram A. Anbar** Vice Chancellor for Admin. Affair of Al-Esraa Univ. / Iraq
- **Lect. Dr. Mohammed J. Al-Shammari** Al-Esraa Univ. / Iraq.
- **Assit. Lect. Ann S. Ibrahim** Al-Esraa Univ./ Iraq.

Financial Manager

- Assist. Lect. Bashar Q. Tayeb Al-Esraa Univ. / Iraq.

Guidelines of Publication in the Al-Esraa University College Journal for Engineering Sciences.

The Al-Esraa University College Journal for Engineering Sciences is published annually by the Al-Esraa University in term of two issues per year.

- The journal is concerned with publishing scientific papers as follows in the Engineering Sciences:
  - Construction engineering.
  - Civil engineering.
  - Chemical engineering.
  - Computer engineering.
  - Electrical engineering.
  - Material engineering.
  - Mechanical engineering..
  - Oil engineering.
  - …… etc.

Terms of publication

1. Each manuscript must be typed using a computer in a single spaced text on one face of the A4 paper (size A4) using 12 font size type (Times New Roman and Simplified Arabic), while the titles in Arabic and English should be written using 14 font size. A 2-cm margin must be left from top

and bottom, and 3 cm from right and left. Articles should not exceed more than 15 pages including tables, figures, and resources taking in consideration that the whole work is written on one face of A4 papers.

2. It is not advisable to publish an article by neither the editor-in-chief nor the members of the editorial board of the journal, whether it is a solo or joint work.
3. After being approved for publication, the article is to be presented in three hard copies and an electronic one. The article is submitted in the final form by being printed on a regular basis for all pages excluding the first one which has the title of the article and the names of the authors and their addresses in both Arabic and English language in addition to the e-mail of the first author English language, the CD copy of the article should be made using Microsoft word 2010.
4. Papers may be accepted in both Arabic and English language. However, English is highly preferred.

Author Guidelines

Below are the terms and requirements need to be considered by the researcher wishing to publish in this journal:

1- The research must not published in any other engineering journal and has not been completed for more than four years prior to publication.

2- The title of the research should be brief and expressive

3- Authors names: the names of authors and their work place addresses should be clearly written along with the first authors e-mail address.

4- An abstract should be clear and about 250- 300 words, followed by a keyword (4-6) in Arabic if the article is in Arabic language followed by

abstract and keywords in English language and virus visa.

5- Introduction: includes a review of information relevant to the subject of research in the scientific sources, ending with the aim of the study and its rationale.

6- Materials and Methods : Should be fully detailed if they are new. In case of being already published, they should be mentioned in brief with reference to the sources and the use of System International Units (S.I.U.s) for measuring weight and volume.

7- Results and Discussion: should be shown in a concise, meaningful and sequential manner. The results are presented in the best form. After being referred in the results, tables and figures should be place in their designated positions.

8- The Arabic numerical system should be used in the researches submitted for publication. The discussion of the results represents a brief expression of the results and their interpretations.

9- Writing the references in the list shall include the name (s) of the authors, the publication year, the title of the research, the name of the journal, volume number, issue number and the number of pages.

e.g. Al-Khafaji, J.M., Hameed, M.H. and Kareem, H.H., (2018) " Experimental Investigation on Concrete with Paetially Replacement of Coarse Aggregate by Junk Rubber."Al-Esraa Univer. College J., 1(1), 217-243.

e.g. Garrick, G.M., (2005) " Analysis and Testing of Waste Tire Fier Modified Concerete ", M.Sc. Thesis, University of Luisiana State, U.S.A., Louisiana, pp. 9-15.

10- The abstract in English must be clear and describe the research and the results in a precise manner and not necessarily be a literal translation of

the Arabic abstract and followed by 4-6 keywords.

Reviewer Guidelines

Below are the terms and requirements to be taken in consideration by the reviewer of the research sent for publication in this journal:

- Filling the evaluation form sent with the research to be evaluated accurately and not leave any paragraph without an answer.
- The reviewer must make sure that the titles, both Arabic and English, are linguistically identical. If not, an alternative title is to be suggested.
- The reviewer should state whether tables and figures seen in the research are thorough and expressive.
- The reviewer should state whether or not the authors uses statistical methods correctly.
- The reviewer should state whether the discussion of the results logically sufficient.
- The reviewer should determine the extent to which the authors uses modern scientific evidences.

- The reviewer should clearly indicate one of the three options as follows:
  - The research is suitable for publication without modifications.
  - The research is suitable for publication after changes are made.
  - The research is not suitable for publication

- The reviewer should clarify in a separate sheet the basic modifications suggested before accepting the article for publication.
- The reviewer has the right to get the manuscript back to him after making

the necessary modifications to make of sure of the authors commitment.

- The reviewer must register his / her name, scientific degree work , address and the evaluation date, with the signature of the evaluation form sent, accompanied by the article submitted for evaluation.

References

1. References in the text of the manuscript are indicated as follows:

   The title or last name of the author and the year if the work is done by one scholar. if there were two authors they should be mentioned along with the year. In case of being three and more, the first one is mentioned then et al., and the year.

2. Reference should be listed according to (APA) and as the examples mentioned:

A .Scientific research in a Journal.

Authors name, year, research title, journal name, volume, issue number and page , numbers.

B. Books.

Authors name, year, title of the book, edition, publishing house and number of pages.

C. Theses and dissertations.

Authors name, year, title of thesis, address of the college and university, and number of pages.

D. Scientific research in the proceedings of a scientific conference or symposium.

Authors name, year, the paper title, the name of the conference or the scientific symposium, venue, the starting and ending pages of the paper.

The journal is highly committed to preserving the intellectual property rights of authors.

Articles are sent to the Al-Esraa University College Journal for Engineering Sciences at the following address:

Al-Esraa University – Documentation and Scientific Publishing Department

Baghdad – Iraq

E_mail : al-esraajournal@esraa.edu.iq

(A Written Undertaking (Pledge) of Intellectual Property)

I /We hereby certify that ..... I am/(We are) the author(s) who have achieved and written the article entitled .....

I /We confirm that this article has never been published in any other journal whether locally or internationally . I /We submit this article for consideration for publication in (**Al-Esraa University College Journal for Engineering sciences**) issued by the Al-Esraa University.

Signature (s) :

Date:

(A Written Undertaking (Pledge) of Copyrights Transfer)

I / We hereby certify that I / We .....,am/ are the authors of the article entitled .....

I /We agree to transfer the copyright to (**Al-Esraa University College Journal for Engineering Sciences**) issued by the Al-Esraa University.

Signature(s) :

Date:

Contents

Guidelines of Publication in the Al-Esraa University College Journal for Engineering Sciences. 5

An AI and IoT Framework for Dynamic Optimization and Sustainability in Smart Cities 17
Zeinab E. Ahmed, Rashid A. Saeed, Salah Hagahmoodi,
Mamoon M. Saeed, Khalid Hamid, Sally D. Abugasim, Eyman F. A. Elsmany

Adaptive Modulation Techniques to Enhance OFDM System Performance 47
Researcher Alaa Aqeel Mohammed

Evaluation of MIMO Antenna Configurations for High Data-Rate Systems 77
Researcher Ammar Abdulimam Abbood

CFD Optimization Using ANSYS Fluent for Advanced Thermal Energy Conversion, A Case Study on Heat Recovery Systems in Iraqi Power Plants 105
Researcher Basim Mohammed Abid Ali

Comparative Adsorptive Performance of Clinoptilolite, Purolite, and Mesoporous Carbon (CMK-3) for Heavy Metals and Hydrocarbons Removal from Real Petrochemical effluent 145
Assist. Lect. Eynas Muhamad Majeed, Assist. Lect. Nuha Muhsen Ali

Real-Time Adaptive Control of Machining Parameters in CNC Milling using an Ensemble Digital Twin and In-Process Vibration Monitoring for Enhanced Surface Integrity 181
Researcher Ali Qasim Abdulwahid

Multiaxial Stress and Deformation Analysis in Mechanical Structures under Variable Dynamic Loads 221
Researcher Aqeel Naeem Fareed Al-Khafaji

Guided Mind Mapping for Enhancing Spatial Thinking in First-Year Architectural Design Studios 253
Assist. Lect.Noor Ayad Rashid and Lect. Dr. Mustafa Kamil Al-Khafaji

The Impact of Architectural Form on Seismic Performance of Buildings 291
Researcjer Durgham Munir Abdul Hadi, Prof. Dr. Abdullah Saadoon Salman,
Assist Prof. Dr. Lahib Ali Abdul Hussein

Video Compression Optimization Techniques Using Artificial Intelligence 325
Assist. Prof. Amal Abbas Kadhim, Assist. Prof. Wedad Abdul Khuder Naser,
Lecturer Nada Abdulkareem hameed

The Impact of Materials and Design Procedures Used in Architectural Facades on Improving the Thermal Performance of the Building's Inner Environment for Hot, Dry Climate Cities, Baghdad 339
Assist Lect. Abeer Qasim Jbur

Design and Development of Solar-Powered Water Desalination Systems 367
Researcher Salam Ali Hassoon Alabooda

Static Analysis-based Android Malware Detection (2021–2026): A Systematic Survey and Taxonomy 397
Assist. Lect. Ali Hussein Mohammed Ali and Assist. Lect. Musaab Riyadh Abdulrazzaq

Current Condition of the Holy Kadhimiya Shrine Surveying Using a GPR 465
Assist. Prof. Dr. Israa Hatem

Spatial Estimation for Gas Flaring and Natural Gas Waste from Oil Production using Remote Sensing and GIS Techniques 479
Lect. Dr. Hisham M. Jawad Al Sharaa

Performance Evaluation of Multi-Stage Compressors and Their Impact on Energy Conversion System Efficiency 'A Field Study at Thi-Qar Oil Company' 499
Researcher Taha Chyad Thajeel

Experimental Analysis of Diesel Engine Combustion Chamber Performance Using Field Measurements at the State Company for Heavy Engineering Equipment 549
Researcher Mohammed Kanaan Sedeeq

Experimental Analysis of Thermal - Hydraulic Enhancement in a Shell-and-U-Tube Heat Exchanger Using Copper Foam Baffles 583
Researcher Rawaa Hussein Al-meammar

Enhancement of Gas Turbine Efficiency by Using Recuperator and Intercooler Systems 601
Mohammed Saad Kamil

Enhancement of Tribological Behavior of Rotating Bearings Using Advanced Ceramic Compounds via Thermal Spray Techniques 623
Researcher Amjid Shihab Abdul Jaleel

Effect of Nano-Additives on Fracture Toughness of Engineering Materials: An Experimental Study 659
Researcher Humam Ahmed Awad

Experimental Study of the Effect of Nano-Doping on the Electrical and Thermal Properties of Organic Semiconductors 701
Chief Programmer Zainab Abdulateef Rashad

Interaction Design in Museum Architecture: A Contemporary Developmental Vision for the Iraqi Museum 739
Researcher Afaf Abdel Sahib Mohammed
Prof. Dr. Enas Salim Abdulahaad

Solar-Assisted Thermal Integration for Improving Fuel Pumping Efficiency in the Middle Oil Distribution Company – Karbala Depot 773
Researcher Ali Shakir Hial

An Intelligent System for Managing High-Priority Vehicle Routing in Smart Cities Using Artificial Intelligence and Spatial Data Analytics 829
Researcher Bahjat Abdulelah Madhloom

AI-Enhanced Heat Transfer Optimization in Magnetic Bio-Nanofluids 861
Researcher Hasan Attyah Shaboot

Strategic Optimal Design of Energy Storage Systems for Enhancing Stability and Sustainability in Electrical Distribution Networks 895
Researcher Hayder Abdulhussein Allawi

Optimization of Bearing Design in Centrifugal Pumps Using Finite Element and Experimental Methods: Application at South Refineries Company, Basra.....933
Researcher Mohammed Hasab Naser Al-Mtasher

Quality of Life Indicators in Historic Urban Spaces:
A Case Study of Al-Mustansiriya Madrasa, Baghdad / Iraq 973
Assist. Lect. Israa H. Karam Ali
and Assist. Lect. Shireen I. Khazaal

Improvement of Mechanical and Corrosion Properties of Weld
in Stainless Steel in Crude Oil Tank by Using Nano-Fillers 1001
Researcher Ali Abdulzahra Shanshoon

An AI and IoT Framework for Dynamic Optimization and Sustainability in Smart Cities / Review Article

Zeinab E. Ahmed[1], Rashid A. Saeed[2*], Salah Hagahmoodi[3], Mamoon M. Saeed[4], Khalid Hamid[5], Sally D. Abugasim[1], Eyman F. A. Elsmany[1]

1 Dept. of Computer Engineering, University of Gezira, Wad-Madani, Sudan.
2 College of Commerce and Business, Lusail University, Lusail, Qatar.
3 Department of Information Technology, College of Science, Al-Esraa University, Bagdad, Iraq.
4 Dept. of Communications and Electronics Engineering, Faculty of Engineering, University of Modern Sciences (UMS), Yemen.
5 Department Communication Systems Engineering, University of Science & Technology, Omdurman, Sudan.
* Corresponding author: eng_rashid@hotmail.com

إطار عمل للذكاء الاصطناعي وإنترنت الأشياء للتحسين الديناميكي والاستدامة في المدن الذكية

زينب إ. أحمد[1]، راشد أ. سعيد[2]*، صلاح حج همودي[3]، مأمون م. سعيد[4]، خالد حامد[5]، سالي د. أبو غاسم[1]، أيمن ف. أ. الصمّاني[1]

1 قسم هندسة الحاسوب، جامعة الجزيرة، ود مدني، السودان
2 كلية التجارة وإدارة الأعمال، جامعة لوسيل، لوسيل، قطر
3 قسم تكنولوجيا المعلومات، كلية العلوم، جامعة الإسراء، بغداد، العراق
4 قسم هندسة الاتصالات والإلكترونيات، كلية الهندسة، جامعة العلوم الحديثة (UMS)، اليمن
5 قسم هندسة نظم الاتصالات، جامعة العلوم والتكنولوجيا، أم درمان، السودان

Abstract

Smart cities are emerging as a critical solution for creating more efficient, sustainable, and comfortable urban environments. This transformation is primarily driven by the synergistic integration of the Internet of Things (IoT) and Artificial Intelligence (AI). The IoT provides a pervasive network of connected sensors that collect real-time urban data, while AI serves as the analytical engine that processes this information to optimize city-wide systems. This paper presents a comprehensive framework that leverages this AI-IoT convergence for dynamic optimization to achieve long-term urban sustainability. The framework focuses on enabling intelligent, data-driven decision-making across core urban domains. The discussion and analysis reveal that the integration of AI and IoT is revolutionizing smart cities by fostering decentralized, intelligent systems. This synthesis of real-time data, advanced analytics, and distributed computing generates significant gains in efficiency, sustainability, and security, evidenced by reduced latency, lower energy consumption, and enhanced threat detection. However, the implementation of such frameworks faces persistent challenges, including issues of scalability, cybersecurity, data standardization, and equitable access. The paper concludes that proactively addressing these technical and socio-technical barriers is essential for building resilient, inclusive, and truly sustainable urban ecosystems for the future.

Keywords: Smart City, Internet of Things (IoT), Artificial Intelligence (AI), Optimization, Sustainability.

المستخلص

تُعد المدن الذكية حلاً حيويًا لإنشاء بيئات حضرية أكثر كفاءة واستدامة وراحة. ويعتمد هذا التحول بشكل أساسي على التكامل التكاملي بين إنترنت الأشياء والذكاء الاصطناعي . يوفر إنترنت الأشياء شبكة شاملة من المستشعرات المتصلة التي تجمع البيانات الحضرية في الوقت الفعلي، بينما يعمل الذكاء الاصطناعي كمحرك تحليلي يعالج هذه المعلومات بهدف تحسين الأنظمة على مستوى المدينة بأكملها. يعرض هذا البحث إطار عمل شامل يستفيد من تقارب الذكاء الاصطناعي وإنترنت الأشياء لتحقيق التحسين الديناميكي بهدف الوصول إلى استدامة حضرية طويلة الأمد. يركز الإطار على تمكين اتخاذ القرار الذكي المبني على البيانات عبر المجالات الحضرية الأساسية. تكشف المناقشة والتحليل أن دمج الذكاء الاصطناعي وإنترنت الأشياء يُحدث ثورة في المدن الذكية من خلال تعزيز الأنظمة الذكية اللامركزية. ويؤدي هذا التكامل بين البيانات في الوقت الفعلي، والتحليلات المتقدمة، والحوسبة الموزعة إلى تحقيق مكاسب كبيرة في الكفاءة، والاستدامة، والأمان، وهو ما يتجلى في انخفاض زمن الاستجابة، وتقليل استهلاك الطاقة، وتعزيز القدرة على اكتشاف التهديدات. مع ذلك، تواجه تطبيقات مثل هذه الأطر تحديات مستمرة، بما في ذلك قابلية التوسع، والأمن السيبراني، وتوحيد البيانات، وضمان الوصول العادل. ويخلص البحث إلى أن معالجة هذه الحواجز التقنية والاجتماعية-التقنية بشكل استباقي أمر أساسي لبناء أنظمة حضرية مقاومة، شاملة، ومستدامة حقًا في المستقبل.

الكلمات المفتاحية: المدينة الذكية، إنترنت الأشياء، الذكاء الاصطناعي، التحسين، الاستدامة

1. Introduction

A smart city represents an advanced urban environment that utilizes cutting-edge technologies, intelligent systems, and effective management practices to improve infrastructure, public services, and residents' quality of life [1]. It connects essential sectors such as transportation, healthcare, energy, and communication through integrated technologies like the Internet of Things (IoT), Cyber-Physical Systems (CPS), and Artificial Intelligence (AI). These innovations enable real-time data analysis, automation, efficiency, and cost savings. The main objective of smart cities is to ensure safer, more comfortable, and higher-quality living while minimizing operational costs. Practical applications include smart grids that regulate energy use, intelligent transport networks enhancing mobility and safety, remote healthcare systems, and buildings that adjust to environmental changes [2]. Despite these advancements, cities still encounter challenges such as overpopulation, large coverage areas, high infrastructure expenses, complex data management, and environmental degradation. To overcome these issues, smart urban planning emphasizes predictive technologies, renewable energy adoption, strong data management, and sustainable policy frameworks [3]. As urbanization continues to expand, consuming nearly 70% of global natural resources, technologies like ICT, IoT, AI, and big data have become vital in creating adaptive, efficient, and sustainable cities for the future [4].

Artificial Intelligence (AI) acts as the brain of the smart city, transforming IoT data gathered through the IoT into intelligent data for efficient decision-making The application of AI is a key determinant in improving the efficiency levels and prediction capacities in all domains [5]. It is fundamental to applications like intelligent traffic management, real-time threat detection in cybersecurity, agriculture, and proactive

smart healthcare systems. Techniques such as computer vision, natural language processing, and predictive modeling are used to analyze trends, optimize systems, and personalize services. However, the integration of AI introduces its own set of challenges, such as ethical use of AI [6]. Future research directions stress the necessity of developing secure, scalable, and explainable AI solutions to ensure these technologies are deployed responsibly and effectively to serve citizens.

This study investigation focuses on an integrated AI and IoT framework essential for achieving dynamic optimization and sustainability in smart cities. This paper is structured as follows. First, an introduction establishes the context of smart cities and examines the roles of IoT and integrated AI within them. This is followed by a literature review that identifies a gap, motivating a detailed exploration of key application domains and optimization case studies. These domains include smart mobility and traffic optimization, energy management, security, and specific application areas such as agriculture, medical imaging, and health monitoring. The subsequent section presents a discussion and analysis, detailing how AI-driven insights from IoT data yield measurable improvements in efficiency and conservation across these domains.

2. Literature Review

This section reviews recent research on IoT and AI, focusing on four key areas for smart cities: mobility, traffic optimization, energy, security, and specialized fields like agriculture and healthcare. A common theme is that combining real-time IoT data with advanced AI techniques leads to major improvements in city efficiency, sustainability, and safety.

2.1. Smart Mobility and Traffic Optimization

Current studies show that integrating AI with IoT is becoming a central strategy for addressing the multifaceted problems in smart mobility and traffic optimization. Qaffas [7] introduced a novel, AI-driven framework for dynamic smart city traffic optimization that relies on a distributed IoT network. The framework's core strength is a hybrid AI model, combining a Vision Transformer (ViT) for real-time visual detection and a Temporal Convolutional Network (TCN) for predictive analysis, which is deployed on edge nodes to ensure high-speed, low-latency decision-making. Using the Dublin Traffic Sensor Dataset, the framework significantly improved traffic efficiency, achieving a 40.4% reduction in average vehicle delay time and a 32.7% increase in intersection throughput. The distributed architecture proved highly responsive, yielding a total decision latency of only 64.9 milliseconds, which substantially surpasses the performance of centralized and cloud-based systems. The proposed Eco-Assistive Fog-based Traffic Management System (EAFTMS) introduced a four-tier architectural model integrating IoT, fog, cloud, and application layers to deliver scalable, real-time urban traffic optimization [8]. Central to the framework is an AI-powered Point Exchange System (PES), designed to incentivize sustainable commuter behaviour, reducing vehicle dependency and encouraging eco-conscious choices through a gamified reward mechanism.

Performance metrics confirm EAFTMS's superior efficacy: a 30% reduction in latency, 40% improvement in response time, 25% increase in traffic flow efficiency, and 35% decrease in CO_2 emissions. This positions EAFTMS as a benchmark solution for developing intelligent, green urban ecosystems that are both operationally effective and environmentally responsible.

Almaazmi *et al.*, [9] , introduced an innovative solution to the limitations of traditional, centralized Energy Management Systems (EMSs), which suffer from poor adaptability, high overhead, and privacy issues Click or tap here to enter text.. The research proposed an Explainable Federated Learning (XFL) framework for vehicular energy optimization in smart cities. This decentralized approach uses Federated Learning (FL) to enable collaborative AI model training across edge devices while preserving data privacy. The addition of Explainable AI (XAI) increases the transparency and trustworthiness of the system's decisions. The XFL framework, utilizing a Multi-Layer Perceptron (MLP)-based global model on real-world vehicle data, demonstrated superior performance, achieving high predictive accuracy with R^2 values of 94.73% for energy consumption and 99.83% for traffic density. This paper [10] presented DST-FedRL, a decentralized traffic management framework that combines reinforcement learning, Deep Spatial–Temporal Transformer Networks, and federated edge artificial intelligence. To protect privacy and facilitate cooperative model updates, the system processes local traffic data. Trained on the Road Traffic Video Monitoring Dataset, DST-FedRL achieves 97.8% vehicle detection accuracy, reduces travel time by 30%, and lowers fuel consumption by 27.9%, offering an efficient, scalable, and sustainable solution for smart city mobility.

As illustrated in [11] proposed a novel AIoT vehicle safety framework designed for smarter, safer cities. This system proactively reduces risks by seamlessly integrating ultrasonic, eye blink, and alcohol sensors. It boosts road safety through

immediate driver alerts based on eye blink detection and utilizes Li-Fi for rapid, efficient vehicle-to-vehicle communication. Critically, if high alcohol levels are detected, the system autonomously controls vehicle speed and instantly notifies authorities via GPS/GSM. This resilient architecture lays a sustainable and safer foundation for intelligent urban transport. A novel framework that combines IoT, AI, and autonomous vehicles (AVs) to revolutionize urban logistics and reduce traffic was presented by Mohsen [12]. The system significantly improves demand prediction, traffic light synchronization, and delivery route planning by "crunching" real-time IoT data using intelligent AI. The product is a highly integrated approach that drastically lowers carbon emissions, greatly improves traffic flow, and creates a sustainable urban delivery future.

Paper [13] details a three-layer architecture for AI-driven vehicular networks to enhance urban mobility. Authors demonstrate how integrated AI optimizes routing and traffic control, while rigorously addressing paramount cybersecurity risks through cryptographic authentication and AI-based anomaly detection. Validation via simulation and real-world testing confirms that this secure, adaptive framework significantly improves traffic efficiency and network integrity, establishing a foundation for safe and intelligent transportation systems. Jadhav, *et al.*, proposed an integrated AI, IoT, and VANET framework for advanced smart parking solutions [13]. The system leverages real-time sensor data and AI analytics to predict availability, optimize routes, and reduce driver search time. This significantly cuts fuel consumption and congestion, boosting urban sustainability. The framework prioritizes robust cybersecurity for secure data sharing and system stability.

Furthermore, paper [14] directly tackles the common frustrations of parking scarcity and network congestion in modern smart cities. The research introduces a robust, unified framework that expertly blends AI, IoT, and vehicular communication

networks. Machine Learning is utilized to forecast where parking will be available, and Deep Reinforcement Learning is employed to boost the efficiency of the VANET routing protocols. Critically, the framework integrates edge computing and block chain technology to ensure both robust security and data integrity, resulting in a highly scalable and sustainable system for dynamic urban parking management through significantly improved V2I/V2V communication.

An automated road safety system using a Hybrid Vision Transformer (ViT) model for instantaneous traffic accident detection via video streams was proposed by Ahmad, *et al.*, [15]. By fusing transformer and CNN capabilities, the model achieves high accuracy in recognizing spatio-temporal events. The system leverages edge computing (e.g., Raspberry Pi) and GSM alerts for immediate notification of emergency services, significantly reducing response time and enhancing overall road safety infrastructure. In [16], authors proposed a novel deep learning approach to next-generation route planning in VANETs. By integrating real-time traffic, road conditions, and environmental factors, the algorithm accurately forecasts optimal routes. Extensive simulations confirm that this method substantially reduces travel time and emissions while increasing transportation throughput, offering a significant advancement for safer, smarter, and greener urban mobility systems.

Securing high-mobility VANETs in IoT-enabled smart cities against evolving attacks demands enhanced security. Research [17] introduced a unique AI-block-chain hybrid system that combines machine learning (CNN/LSTM) for intrusion detection (96% accuracy) with block-chain for decentralized trust and unbreakable data verification. This approach, which also uses ECC for robust authentication, significantly reduces Sybil attacks and outperforms traditional PKI and centralized frameworks, establishing a strong, privacy-aware basis for resilient smart transportation. Furthermore, the following study [18] compared Machine Learning

(ML) algorithms for Intrusion Detection Systems (IDS) in resource-constrained VANETs. Analysis confirms Decision Tree (DT) as the optimal choice, delivering high accuracy and exceptional efficiency crucial for such environments. Random Forest (RF) is best for models requiring continuous learning adaptability, while K-NN suits small datasets needing fast deployment. These findings guide the selection of appropriate ML algorithms based on specific network performance and resource limitations.

2.2. Smart Energy Management

Addressing high energy consumption and local optima issues in WSN routing, the paper [19] proposed a regional energy balance routing algorithm based on the Intelligent Chaotic Ant Colony (ICAC) approach. The method utilizes an adaptive perturbation strategy and an energy-delay metric to find the global optimum and ensure uniform node energy distribution. Simulation results show significant performance gains, including a 7.3% average energy consumption reduction compared to similar algorithms, effectively extending the network life cycle. Chen *et al.* introduced an energy-efficient IoT framework for smart street lighting, replacing metal halide lamps with mesopic LED lamps [20]. The system utilizes a smart decision module, Pulse Width Modulation (PWM) dimming, and a DALI controller to tune intensity based on traffic and occupancy. Powered by PV solar and an MPPT dynamic charging algorithm, experiments confirm significant energy savings across various urban areas, reducing overall consumption and carbon emissions.

A CNN-IoT framework for enhanced energy management in smart buildings was provided by the study [21]. By integrating deep learning with real-time IoT data, the framework significantly improves prediction accuracy and operational efficiency over traditional methods. Performance metrics confirm its robustness, achieving an 88% accuracy and a 0.88 F1-score. This solution offers proactive

anomaly detection and superior energy optimization, positioning it as a key advancement for sustainable building management. An innovative deep learning framework to optimize energy management in IoT-enabled smart cities was proposed by Aljohani [22]. Leveraging real-time data from smart grids and sensors, the system uses neural networks and RNNs for precise, dynamic demand forecasting. Detailed simulations show significant reductions in energy usage, cost savings, and greenhouse gas emissions, establishing a versatile and scalable route toward a more sustainable and economically efficient urban future.

Raval *et al.*, used an energy management framework addressing both hardware modelling and software control [23]. A Multi-Agent Distributed Intelligent System (MAS) is introduced, with parameters optimized by a Genetic Algorithm (GA) to achieve decentralized intelligence. Simulation results reveal substantial energy savings, including a 40% reduction in total energy consumption, validating the framework's effectiveness. Smart cities require dynamic waste management to tackle inefficiency and high fuel use. This research [24] proposed a Deep Reinforcement Learning (DRL) framework for optimizing Autonomous Waste Collection Vehicle (AWCV) routes using IoT smart bins. Utilizing Deep Q-Networks (DQNs) to analyse real-time traffic and energy data, the system achieves a 30% reduction in distance and a 30% decrease in energy usage over baseline methods, demonstrating a sustainable and highly efficient operational strategy.

Chui *et al.*, introduced a novel Hybrid Genetic Algorithm Support Vector Machine Multiple Kernel Learning (GA-SVM-MKL) approach for Non-Intrusive Load Monitoring (NILM) across 20 appliance types [25]. GA optimizes kernel design to solve the multi-objective problem. The proposed method achieves an Overall Accuracy (OA) of 91.8%, representing a >21% performance improvement over traditional kernels. This integration of AI, IoT, and big data analytics significantly

advances energy transparency and utilization in smart cities. Addressing the high latency and energy drain of traditional cloud-based IoT in smart cities, the study [26] validated adaptive edge computing as a key sustainable solution. The framework achieves significant gains in power efficiency and reduced data transmission overhead by adaptively distributing workloads and using AI-driven optimization. Findings confirm that localized edge decision-making drastically improves scalability and real-time responsiveness, promoting energy sustainability and performance in future urban environments.

Intelligent Deep Optimized Energy Management (Int-DEM), a novel cognitive IoT-enabled framework for effective energy management in smart grid systems was developed by Ganesh *et al.*, [27]. Int-DEM addresses grid data limitations using an intermediate cloud analysis concept and employs the hybrid deep learning model Stacked Convoluted Bi-Directional Gated Attention Network (SCon-BGAN), optimized by the novel Hybrid Darts Seagull Optimizer (HDSO) technique, for precise energy load forecasting. Comparative analysis confirms that Int-DEM significantly outperforms existing deep learning models, achieving superior accuracy with lower computational complexity, which establishes a robust and efficient tool for sustainable smart grid management. This study [28]proposed the Ensemble Red Fox Optimization with Energy Powered Artificial Neural Network (ERFO-EPANN), a hybrid deep learning model for precise solar irradiance and wind speed forecasting. Combining bio-inspired optimization with neural learning, ERFO-EPANN integration of RES with Energy Storage Systems (ESS), thereby maximizing grid resilience and sustainable urban infrastructure.

Nazari *et al.* [29] suggested a multi-stage, energy-efficient routing protocol for IoT networks integrating Ladybug/Butterfly Optimization (LBO/BOA) and Q-learning Click or tap here to enter text.. The protocol uses hybrid clustering for load balancing and dynamic Q-learning for multi-hop routing. Simulations confirm

superior performance, including a 43% reduction in energy consumption and up to 47% extension of network lifetime compared to state-of-the-art methods, significantly enhancing network stability and coverage. This study introduced a new framework that integrates Quantum Computing, AI, and IoT for power quality management in renewable energy-driven Flexible AC Transmission Systems (FACTS) [30]. To counter RES variability and stability issues, the core method utilizes AI-enhanced quantum algorithms for real-time grid optimization, including efficient fault detection and load balancing. The system employs Unified Power Quality Conditioners (UPQC) and smart meters for continuous monitoring and compensation of power disturbances like voltage sags. Validated by simulation and hardware, this quantum-AI synergy successfully enhances grid resilience and security, demonstrating a practical approach to deploying advanced computational technology in clean energy infrastructures.

Deep Learning-augmented Predictive Energy Modeling (DL-PEM) framework combined with Multi-Objective Particle Swarm Optimization (MOPSO) for occupancy-based energy management in smart buildings was developed by this article [31]. Using indoor environmental data, the system optimizes HVAC and lighting to balance energy consumption and thermal comfort. Achieving 99.8% accuracy and up to 85% optimization efficiency, this framework significantly outperforms traditional models, promoting sustainability and scalable urban solutions. Singh *et al.* designed a Scalable Cloud-Based Continuous Monitoring Platform (SC-CMP) for real-time optimization of microgrids in EV-dense, renewable-integrated environments [32]. Fusing cloud computing, ML, and IoT data, the platform enables continuous monitoring, V2G coordination, and adaptive energy management. SC-CMP achieves 97.34% predictive accuracy and 96.81% grid stability improvement over existing baselines.

Authors developed a system which utilizes a hybrid deep learning (DL) framework to significantly boost energy efficiency within IoT wireless networks [33]. This intelligent system integrates Convolutional and Recurrent Neural Networks (CRNN) with the Krill Herd Optimization (KHO) algorithm, allowing it to efficiently forecast energy consumption trends and dynamically optimize resource allocation. Validation through simulations demonstrated that this framework yields substantial performance improvements, achieving up to a 52% decrease in energy consumption and maintaining high predictive accuracy 93.34%, thus presenting a highly scalable solution for dynamic energy management in complex IoT environments. Khan *et al.* addressed critical energy constraints in IoT networks by enhancing the Low Energy Adaptive Clustering Hierarchy (LEACH) protocol using Digital Twin simulation and AI analytics [34]. This novel framework optimizes cluster head selection and assesses performance through virtual network models. Findings confirm significant network longevity improvements, including an 83% reduction in non-functioning nodes and a 1.66-fold increase in node energy levels, showcasing a viable direction for resilient, energy-efficient IoT networking.

2.3. Smart Cities Security

Agajo *et al.*, developed a system for monitoring and securing smart city infrastructure by integrating Block-chain, AI, and IoT [35]. Their work addresses security and data privacy concerns in cloud-based IoT platforms by proposing a distributed communication and storage mechanism via a block-chain-based infrastructure. The study is relevant as it provides an implementable model for secure data management in smart city networks. A hybrid model integrating LSTM networks and advanced feature engineering (IDS-SIoDL), was proposed by Hazman *et al.*, [36]. Evaluated on IoT datasets (Bot-IoT, Edge-IIoT), it achieves high accuracy and low latency, demonstrating effectiveness for real-time threat detection in smart cities.

The article [37]addressed smart city IoT security by proposing a two-tier cooperative framework. It integrates CNNs with AdaBoost, LightGBM, and XGBoost, optimized via Evolutionary Adaptive Firefly Algorithm (EAFA) for simultaneous feature selection and hyper parameter tuning. This approach, featuring simultaneous feature selection and hyper parameter tuning via EAFA, achieved 0.995565 accuracy, with performance further validated through XAI interpretability. Indra, *et al.*, [38] introduced an Ensemble Gradient Random Forest-based Leopard Seal Search (EGR-LSS) algorithm to improve intrusion detection in IoT-based smart cities. It addresses key research gaps in existing systems, such as insufficient data samples, high execution time, and computational complexity, by integrating ensemble voting and an optimization algorithm for hyper parameter tuning. The model achieving an accuracy of 98.75%, a recall of 97.4%, a precision of 98.7%, and an F1-score of 96.9%. These results surpass the performance of current leading techniques, including CNN, AI-BC, and NB.

To address the shortcomings of existing smart parking systems, such as vulnerabilities in privacy, security, and centralization, this paper [39] introduced Hybrid Deep learning and Redactable consortium block-chain based Secure Framework (HDLRCBSF). The framework utilizing a specialized Bi-LSTM model to intelligently predict parking availability and the Threshold Chameleon Hash algorithm (TCH) to ensure secure and verifiable data communication. It achieved a 21.32% improvement in optimal parking space allocation compared to baseline methods, while also providing drivers with enhanced privacy and security. The study [40] introduced a privacy-preserving AI framework for IoT-based smart cities. By combining federated learning with local differential privacy and FPGA edge acceleration, it enables secure, efficient model training. The system maintained near-centralized accuracy with just a 1.2–1.5% drop, while reducing the success of

privacy attacks by over 50%. This demonstrates an effective balance between high performance and robust data protection.

To tackle the shortcomings of current IoT intrusion detection systems (poor adaptability and scalability), the study [41] presented the FOFSDL-SCD model. By combining a Fox Optimizer for feature selection, a Temporal Convolutional Network for classification, and Dung Beetle Optimization for parameter tuning, the framework significantly improves detection performance and efficiency. It achieved an impressive 99.38% accuracy, along with precision, recall, and F1-score all above 96%, outperforming existing methods. Chakrabarty, *et al.*, suggested a new security model [42]. It combines Black Networks for full communication encryption with Artificial Intelligence to enable intelligent, self-governing, and decentralized oversight. This strategy offers a direct defense against the complex array of cyber threats that modern urban infrastructures now encounter.

2.4. Smart City Application Areas

Keote *et al.*, designed a new IoT and AI framework that significantly advances structural health monitoring [43]. The system integrates physics-based simulations with intelligent sensor optimization to provide real-time, adaptive, and accurate strength prediction, transforming structures like dome trusses into self-learning systems. Empirical results demonstrate substantial performance gains over conventional methods, including a 35% reduction in sensor count, a 34% decrease in strength prediction Root Mean Square Error (RMSE), and a 60%reduction in bandwidth consumption. Author introduced PANDFOG, a novel fog computing framework designed for the real-time, multi-site analysis of heterogeneous medical imaging data, addressing the critical challenge of pandemic screening in smart cities [44]. By overcoming the latency and data inconsistency issues of centralized cloud-

based AIoT systems, PANDFOG utilizes a specialized multi-decoder deep learning network to enable rapid, on-site diagnosis and achieve high segmentation accuracy, thereby enhancing clinical decision-making. The paper [45] introduced the INRwLF routing algorithm to address the limitations of existing protocols in large-scale I-IoT healthcare networks. By combining AI-driven Software-Defined Networking with hybrid metaheuristic optimization, it enables adaptive, load-balanced clusters and fault-tolerant routing. This design effectively reduces energy consumption and latency, substantially enhancing network lifespan and reliability in dynamic, resource-constrained settings.

Sekar *et al.*, proposed the SRGA-CGG algorithm to address limitations in existing AI models for smart agriculture, such as high computational costs and poor adaptability to environmental variability [46]. By integrating stacked RNNs, GANs, and an optimized greylag goose algorithm, it enhances predictive accuracy and efficiency in resource management, achieving 98.5% accuracy in crop disease detection. Authors addressed the research gap of integrating actionable AI with low-cost consumer devices by proposing a lightweight Random Forest model for on-device crop yield prediction [47]. Deploying via ESP32 sensors and smartphones, it optimizes irrigation with 90.1% accuracy, enabling real-time, localized water management while reducing cloud dependency in resource-constrained settings.

A research addressed the need for scalable and accurate AI systems in agriculture by assessing hybrid deep learning architectures integrated with IoT-based precision frameworks [48]. Among the evaluated models, the Hybrid ResNet demonstrates superior performance, attaining 97.2% accuracy and 95% security, thereby providing a resilient and adaptable solution for the complexities of heterogeneous farming environments. This review [49] synthesizes 88 studies to examine the integration of deep learning and IoT in smart farming. While these

technologies significantly enhance crop recommendation, disease detection, and yield prediction, the analysis identifies critical gaps in IoT security, rural connectivity, and dataset standardization. The paper addresses these limitations by proposing a secure 5G-based framework to advance precision agriculture.

3. Discussion and Analysis

According to the studies in subsection 2.1, intelligent transportation systems have clearly moved from centralized architectures to decentralized solutions based on AI. Essentially, this transition in transport systems is driven by a combination of the IoT, edge computation, and AI. Table 1 provides a comparative summary of the literature, detailing the scope, methodologies, challenges, and key outcomes for each study in the domain of AI and IoT for smart mobility and traffic optimization. One of the main findings highlights the importance of edge and fog computing in overcoming latency, which is common in cloud-centric systems. Empirical studies conducted by Qaffas [1] and the EAFTMS methodology [2] have shown that processing information using AI systems deployed in edges, such as Hybrid Vision Transformers, can achieve latency in terms of milliseconds, which consequently leads to optimized traffic flow performance and enhanced intersection capacity. Furthermore, the field evolves with the implementation of specialized AI approaches. These include federated learning to maintain privacy, deep reinforcement learning with adaptive control, and hybrid AI-block-chain solutions for added security. Individually and cumulatively, these methods address related issues in forecasting, security, and environmentally sound logistics. The achieved results include both functionalities, indicated by alleviated traffic congestion and travel time, and sustainability, indicated by lower fuel consumption and reduced CO_2 emissions. Despite this, a continuation of challenges remains in matters such as the scalability

of these distributed systems, resistance to advanced cyber-attacks, and various ethics and regulatory issues related to self-driving city infrastructure.

Table 1: AI and IoT in Smart Mobility and Traffic Optimization

Ref	Scope	Methodologies	Challenges	Results
[7]	Dynamic Traffic Optimization	Hybrid AI (ViT + TCN) model.	High latency and low responsiveness of centralized systems.	40.4% reduction in vehicle delay time; total decision latency of 64.9 ms.
[8]	Sustainable Traffic & Commuter Behavior	EAFTMS &(PES)	CO_2 emissions, traffic congestion, and need for eco-conscious commuting.	30% latency reduction; 25% increase in traffic flow efficiency.
[9]	Vehicular Energy Optimization	Explainable Federated Learning (XFL)	Traditional EMS limitations: poor adaptability, high overhead, and data privacy risks.	High predictive accuracy (R^2 > 94%) for energy consumption and traffic density.
[10]	Decentralized Traffic Management	DST-FedRL	Centralized system inefficiencies and data privacy concerns.	97.8% vehicle detection accuracy; 30% reduction in travel time.
[11]	AIoT Vehicle Safety	Li-Fi for V2V	Immediate road risks, impaired driving, and efficient inter-vehicle communication.	Proactive risk mitigation and autonomous speed adjustment for safety violations.
[12]	Urban Logistics Optimization	AI, AVs, and IoT framework	City congestion, high carbon footprint, and inefficient delivery routes.	Drastically lowers carbon emissions and greatly improves traffic flow for deliveries.
[15]	Next-Generation Route Planning	Deep Learning approach.	High travel time and emissions in dynamic VANET environments.	Substantially reduces travel time and emissions while increasing transportation throughput.
[16]	VANET Security	AI-Blockchain Hybrid	Sybil, GPS spoofing, DDoS, and limitations of centralized security (PKI).	96% accuracy in intrusion detection; significant reduction in Sybil attacks.
[17]	IDS Algorithm Selection in VANETs	K-NN & SVM	Resource constraints and selecting the optimal algorithm for specific network needs.	Decision Tree (DT) is the optimal choice for resource-constrained VANETs due to high accuracy and efficiency.

The research discussed in section 2.2 demonstrates that by integrating AI, IoT, and optimization algorithms, is driving a major change: smart city energy management is becoming intelligence driven. Ongoing studies are utilizing advanced techniques, utilizing methods such as deep learning (CNN, RNN), hybrid learning models (SCon-BGAN, CRNN), and bio-inspiration for optimization (GA, HDSO, LBO) to accomplish accurate forecasting, dynamic control, and optimized routing. Some major achievements include a dramatic reduction in energy consumption with increases from 7.3% to 52%, prolonging network lifespan by a maximum of 47%, and increasing operational accuracy up to a level of 99.8%. The emerging paradigms of this technology have successfully tackled major technical problems such as localized optimum in wireless sensor networks, intermittency in renewable sources, and real-time demand response with utilization of real-time Internet of Things information and decentralized systems such as edge and cloud computing. A comparative summary of the surveyed Smart Energy Management research is presented in Table 2.

Table 2: Summary of AI and IoT in Smart Energy Management

Ref	Scope	Methodologies	Challenges	Results / Future
[21]	Smart Building Energy Management	CNN-IoT integration deep learning	Low prediction accuracy, inefficient traditional methods	88% accuracy, 0.88 F1-score, proactive anomaly detection
[22]	Smart City Energy Optimization	Neural networks, RNNs	Inaccurate demand forecasting, high energy use	Reduced energy usage, cost, and greenhouse gas emissions
[23]	Decentralized Energy Management	MAS &GA optimization	Centralized inefficiency, lack of adaptive control	40% reduction in total energy consumption
[24]	Waste Collection Vehicle Routing	Deep Reinforcement Learning (DQN)	Inefficient routes, high fuel consumption	30% reduction in distance & energy usage

Ref	Scope	Methodologies	Challenges	Results / Future
[25]	Non-Intrusive Load Monitoring (NILM)	Hybrid GA-SVM-Multiple Kernel Learning	Low disaggregation accuracy for multiple appliances	91.8% Overall Accuracy (>21% improvement)
[26]	Edge Computing for IoT Sustainability	AI-driven workload distribution	High latency, energy drain of cloud-based IoT	Improved power efficiency, scalability, real-time response
[27]	Smart Grid Load Forecasting	SCon-BGAN & HDSO	Grid data limitations, forecasting complexity	Superior accuracy with lower computational complexity
[28]	Renewable Energy Forecasting	ERFO-EPANN	Unpredictable solar/ wind resources	High accuracy (RMSE 1.650, MAE 1.100), enhanced grid resilience
[30]	Power Quality in Renewable Grids	Quantum computing	RES variability, grid stability, power disturbances	Enhanced grid resilience, security, real-time compensation
[31]	Occupancy-Based Building Management	Deep Learning Predictive Modeling	Balancing energy savings with occupant comfort	99.8% accuracy, 85% optimization efficiency
[32]	Microgrid Optimization with EVs	Scalable Cloud-Based Platform	EV integration, renewable intermittency, grid stability	97.34% predictive accuracy, 96.81% stability improvement
[33]	IoT Wireless Network Efficiency	Hybrid CRNN + Krill Herd Optimization (KHO)	High energy consumption in wireless IoT networks	52% energy decrease, 93.34% predictive accuracy

Section 2.3 shows how blockchain technology and artificial intelligence are driving an evolution in smart city security, shown in Table 3. Data privacy, centralized attack surfaces, and real-time threat detection are the three main vulnerabilities of IoT ecosystems that research methodically addresses by moving to intelligent frameworks. The widespread use of hybrid AI models, such as those that combine CNN, LSTM, and ensemble techniques (e.g., IDS-SIoDL, FOFSDL-SCD), optimized with bio-inspired algorithms (EAFA, LSS), achieves remarkable intrusion detection accuracy (up to 99.38%) and low latency. Concurrently, blockchain and cryptographic

methods (TCH) provide immutable data integrity and user privacy for applications like smart parking. Crucially, the paradigm shifts from centralized to federated learning and edge-AI architectures, as demonstrated by frameworks maintaining high accuracy with minimal privacy loss, effectively balance performance with robust data protection. This convergence of explainable AI, adaptive optimization, and decentralized trust mechanisms establishes a resilient, scalable, and proactive security foundation essential for sustainable smart urban infrastructure.

Table 3: Summary of AI in Smart City Security

Ref	Scope	Methodologies	Challenges	Results / Future
[35]	Securing Smart City Infrastructure	Blockchain, AI, and IoT	Security & data privacy in cloud-based IoT platforms	Provides an implementable model for secure, decentralized data management.
[36]	Real-time IoT Threat Detection	IDS-SIoDL	Need for accurate, low-latency intrusion detection	High accuracy & low latency on IoT datasets (Bot-IoT, Edge-IIoT), suitable for real-time use.
[37]	IoT Security & Intrusion Detection	Evolutionary Adaptive Firefly Algorithm (EAFA)	Feature selection, hyperparameter tuning, model interpretability	Achieved 0.995565 accuracy; validated via XAI for interpretability.
[38]	Intrusion Detection for IoT	EGR-LSS algorithm	Insufficient data, high execution time, computational complexity	Superior performance (98.75% accuracy) vs. CNN, AI-BC, NB; high recall & precision.
[40]	Privacy-Preserving AI for IoT	Federated Learning & FPGA	Balancing data privacy with model accuracy in distributed learning	Near-centralized accuracy (only 1.2–1.5% drop); reduced privacy attack success by >50%.
[41]	Adaptive IoT Intrusion Detection	Fox Optimizer + Temporal CNN	Poor adaptability & scalability of current IDS	99.38% accuracy; precision, recall, F1-score >96%; outperforms existing methods.

Table 4 demonstrates how integrating IoT and AI addresses critical challenges in smart city applications, primarily enhancing efficiency and enabling real-time responsiveness. For instance, new approaches to monitoring infrastructure health create intelligent systems that learn as they operate, cutting bandwidth needs by 60%. In healthcare, decentralized tools enable faster pandemic screening at local sites, while smarter network designs keep critical medical IoT devices running longer. Agriculture is seeing a revolution, with AI models now diagnosing crop disease with remarkable accuracy and small, on-device systems allowing for precise water management right in the field. Yet, the analysis also reveals ongoing hurdles, including security risks, inconsistent data, and spotty rural internet access.

Table 4: IoT & AI in Smart City Application Areas

Ref	Scope	Methodologies	Challenges	Results / Future
[43]	Structural Health Monitoring	IoT & AI framework	High sensor cost, bandwidth consumption, and prediction error in conventional monitoring	35% sensor reduction, 34% lower RMSE in strength prediction, 60% less bandwidth use.
[44]	Pandemic Screening via Medical Imaging	PANDFOG fog computing framework	High latency & data inconsistency in centralized cloud AIoT for multi-site medical analysis	Enables real-time, on-site diagnosis with high segmentation accuracy for clinical decision-making.
[45]	Healthcare I-IoT Networking	INRwLF routing algorithm	High energy, latency, and poor reliability in large-scale Industrial IoT healthcare networks	Reduces energy consumption & latency, enhances network lifespan & reliability in dynamic settings.
[46]	Smart Agriculture (Crop Disease)	SRGA-CGG: Stacked RNNs + GANs	High computational cost, poor adaptability to environmental variability in AI models	98.5% accuracy in crop disease detection, enhancing predictive accuracy & resource management.

Ref	Scope	Methodologies	Challenges	Results / Future
[47]	On-Device Crop Yield Prediction	Lightweight Random Forest model	Integration of actionable AI with low-cost devices; cloud dependency in constrained settings	90.1% accuracy; enables real-time, localized irrigation optimization, reducing cloud reliance.
[48]	Precision Agriculture with IoT Security	Hybrid deep learning architectures	Need for scalable, accurate, and secure AI systems in heterogeneous farming environments	Hybrid ResNet achieved 97.2% accuracy & 95% security, providing a resilient solution.

In conclusion, the integration of AI and IoT is transforming smart cities by enabling decentralized, intelligent systems across mobility, energy, security, and specialized applications. This synthesis leverages real-time data, advanced analytics, and distributed computing to drive significant gains in efficiency, sustainability, and security, evidenced by reduced latency, lower energy consumption, and enhanced threat detection. However, challenges in scalability, cybersecurity, data standardization, and equitable access persist. Addressing these barriers is essential for realizing resilient, inclusive, and sustainable urban ecosystem.

4. Conclusion

This paper has investigated the transformative impact of integrating Artificial Intelligence (AI) and the Internet of Things (IoT) on urban development. By processing vast streams of sensor data with intelligent algorithms, this synergy facilitates data-driven optimization across core urban domains, including mobility, energy, security, and public services. A pivotal shift from centralized to decentralized, edge-centric architectures has been instrumental in achieving measurable outcomes, such as decreased traffic congestion, enhanced energy efficiency, and improved cyber threat mitigation. However, realizing the complete potential of smart cities is contingent upon addressing persistent implementation challenges.

Critical barriers include achieving scalable system deployment, fortifying defences against evolving cyber threats, and ensuring interoperability through standardized protocols. Equally important is promoting equitable access to technological benefits to prevent societal disparities and foster inclusive urban growth.

References

[1] B. N. Silva, M. Khan, and K. Han, (2018) "Towards sustainable smart cities: A review of trends, architectures, components, and open challenges in smart cities," *Sustain Cities Soc*, vol. 38, pp. 697–713, doi: 10.1016/j.scs.2018.01.053.

[2] J. Shao and B. Min, (2025) "Sustainable development strategies for Smart Cities: Review and development framework," *Cities*, vol. 158, p. 105663, doi: 10.1016/j.cities.2024.105663.

[3] B. Cengiz, I. Y. Adam, M. Ozdem, and R. Das, (2025) "A survey on data fusion approaches in IoT-based smart cities: Smart applications, taxonomies, challenges, and future research directions," *Information Fusion*, vol. 121, p. 103102, doi: 10.1016/j.inffus.2025.103102.

[4] W. M. Dahmane, S. Ouchani, and H. Bouarfa, (2025) "Smart cities services and solutions: A systematic review," *Data Inf Manag*, vol. 9, no. 2, p. 100087, doi: 10.1016/j.dim.2024.100087.

[5] H. Kokate, B. Jadhav, and S. K. Gupta, (2025) "Optimizing Smart Parking in Urban Spaces: AI-Driven VANET Models Using IOT Frameworks," in *In AI-Driven Transportation Systems: Real-Time Applications and Related Technologies*, pp. 249–271. doi: 10.1007/978-3-031-98349-8_13.

[6] A. Maciá-Lillo, H. Mora, A. Jimeno-Morenilla, N. E. García-D'Urso, and J. Azorín-López, (2025) "AI edge cloud service provisioning for knowledge management smart applications," *Sci Rep*, vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-14429-7.

[7] A. A. Qaffas, "AI-driven distributed IoT communication architecture for smart city traffic optimization," *Journal of Supercomputing*, vol. 81, no. 8, doi: 10.1007/s11227-025-07426-0.

[8] F. M. Talaat, W. M. Shaban, H. ZainEldin, M. Badawy, and M. Elhosseini, (2025) "Integrating V2X solutions in intelligent green cities: an AI-driven point exchange system approach," *Neural Comput Appl*, vol. 37, no. 20, pp. 15311–15343, doi: 10.1007/s00521-025-11187-5.

[9] K. I. A. Almaazmi, S. J. Almheiri, M. A. Khan, A. A. Shah, S. Abbas, and M. Ahmad, (2025) "Enhancing smart city sustainability with explainable federated learning for vehicular energy control," *Sci Rep*, vol. 15, no. 1, doi: 10.1038/s41598-025-07844-3.

[10] C. E. Mohankumar and A. Manikandan, "Decentralized traffic management with Federated Edge AI: a reinforced transnet model for real-time vehicle object detection and collaborative

route optimization," *Discover Applied Sciences*, vol. 7, no. 7, Jul. 2025, doi: 10.1007/s42452-025-07383-6.

[11] S. K. Jagatheesaperumal, S. E. Bibri, J. Huang, J. Rajapandian, and B. Parthiban, (2022) "Artificial intelligence of things for smart cities: advanced solutions for enhancing transportation safety," *Computational Urban Science*, vol. 4, no. 1, doi: 10.1007/s43762-024-00120-6.

[12] B. M. Mohsen, (2024) "AI-Driven Optimization of Urban Logistics in Smart Cities: Integrating Autonomous Vehicles and IoT for Efficient Delivery Systems," *Sustainability (Switzerland)*, vol. 16, no. 24, doi: 10.3390/su162411265.

[13] B. Jadhav, M. Sayyed, and S. K. Gupta, (2025) "AI and VANETs in Smart Cities: The Next Frontier in Traffic Management," pp. 23–39. doi: 10.1007/978-3-031-98349-8_2.

[14] B. Singh, A. Raghav, and C. Kaunert, (2025) "Featuring Smart Parking Solutions in Smart Cities: VANETs Optimization Using AI and IoT," pp. 229–248. doi: 10.1007/978-3-031-98349-8_12.

[15] M. O. Ahmad and S. T. Siddiqui, (2025) "Accident Detection from CCTV Surveillance Using Hybrid Vision Transformation and Alert the Nearest Hospital,", pp. 209–227. doi: 10.1007/978-3-031-98349-8_11.

[16] R. Dey, R. A. Mahajan, M. Khan, S. Karimullah, and B. M. Mujahid, (2025) "Intelligent Route Navigation in VANETs Using Deep Learning," in *In AI-Driven Transportation Systems: Real-Time Applications and Related Technologies*, pp. 57–77. doi: 10.1007/978-3-031-98349-8_4.

[17] T. Basri, (2025) "Securing VANETs for Internet of Things (IoT): AI-Driven Solutions for Privacy and Intrusion Detection," pp. 105–122. doi: 10.1007/978-3-031-98349-8_6.

[18] M. Mubarik, S. N. M. Shah, and U. Zahoora, (2025) "Evaluating ML Models for Intrusion Detection in Network and VANET Security," pp. 123–150. doi: 10.1007/978-3-031-98349-8_7.

[19] J. Zhang, (2020) "Real-time detection of energy consumption of IoT network nodes based on artificial intelligence," *Comput Commun*, vol. 153, pp. 188–195, doi: 10.1016/j.comcom.2020.02.015.

[20] Z. Chen, C. B. Sivaparthipan, and B. A. Muthu, (2022) "IoT based smart and intelligent smart city energy optimization," *Sustainable Energy Technologies and Assessments*, vol. 49, doi: 10.1016/j.seta.2021.101724.

[21] F. A. Alijoyo, (2024) "AI-powered deep learning for sustainable industry 4.0 and internet of things: Enhancing energy management in smart buildings," *Alexandria Engineering Journal*, vol. 104, pp. 409–422, doi: 10.1016/j.aej.2024.07.110.

[22] A. Aljohani, (2024) "Deep learning-based optimization of energy utilization in IoT-enabled smart cities: A pathway to sustainable development," *Energy Reports*, vol. 12, pp. 2946–2957, doi: 10.1016/j.egyr.2024.08.075.

[23] M. Raval, S. Bhardwaj, A. Aravelli, J. Dofe, and H. Gohel, (2021) "Smart energy optimization for massive IoT using artificial intelligence," *Internet of Things (Netherlands)*, vol. 13, doi: 10.1016/j.iot.2020.100354.

[24] T. Kavitha, K. R. Chaganti, S. L. R. Elicherla, M. R. Kumar, D. Chaithanya, and K. Manikanta, (2025) "Deep Reinforcement Learning for Energy Efficiency Optimization using Autonomous Waste Management in Smart Cities," in *Proceedings of 5th International Conference on Trends in Material Science and Inventive Materials, ICTMIM 2025*, Institute of Electrical and Electronics Engineers Inc., pp. 272–278. doi: 10.1109/ICTMIM65579.2025.10988394.

[25] K. T. Chui, M. D. Lytras, and A. Visvizi, (2018) "Energy sustainability in smart cities: Artificial intelligence, smart monitoring, and optimization of energy consumption," *Energies (Basel)*, vol. 11, no. 11, doi: 10.3390/en11112869.

[26] P. Singh, G. R. V, S. Ponmaniraj, A. Aggarwal, and S. Kumari, (2025) "Energy-Efficient IoT Architectures for Smart Cities: Implementing Adaptive Edge Computing for Sustainable Urban Development," in *2025 International Conference on Electronics, AI and Computing (EAIC)*, IEEE, pp. 1–6. doi: 10.1109/EAIC66483.2025.11214219.

[27] P. M. J. Ganesh, B. M. Sundaram, P. K. Balachandran, and G. B. Mohammad, (2025) "IntDEM: an intelligent deep optimized energy management system for IoT-enabled smart grid applications," *Electrical Engineering*, vol. 107, no. 2, pp. 1925–1947, doi: 10.1007/s00202-024-02586-3.

[28] B. Nagappan, J. Shah, M. Gupta, A. Kumari, B. P. Singh, and S. Upadhyay, (2025) "Empowering smart cities: deep learning for seamless integration of energy storage systems and renewable sources," *Environ Dev Sustain*, doi: 10.1007/s10668-025-06832-x.

[29] A. Nazari, S. S. Jazaeri, A. Omidi, and M. Mansoorizadeh, (2025) "An AI-driven multi-stage routing protocol for energy-efficient IoT networks," *Telecommun Syst*, vol. 88, no. 3, doi: 10.1007/s11235-025-01344-5.

[30] K. Sarker, (2025) "Harnessing Quantum Computing and AI for Enhanced Optimization in Renewable Energy-Driven Smart Grid Systems," *Iranian Journal of Science and Technology - Transactions of Electrical Engineering*, doi: 10.1007/s40998-025-00943-x.

[31] S. Zhang and C. Kong, (2025) "Deep learning and multi-objective optimization for real-time occupancy-based energy control in smart buildings," *Sci Rep*, vol. 15, no. 1, p. 39466, doi: 10.1038/s41598-025-23134-4.

[32] A. R. Singh *et al.*, (2025) "A scalable cloud-integrated AI platform for real-time optimization of EV charging and resilient microgrid energy management," *Sci Rep*, vol. 15, no. 1, doi: 10.1038/s41598-025-21531-3.

[33] B. Patil *et al.*, (2025) "CRNN-KHO: hybrid deep learning optimization framework for enhancing energy efficiency in IoT-enabled wireless networks," *International Journal of Information Technology (Singapore)*, doi: 10.1007/s41870-025-02792-7.

[34] S. B. Khan, A. Kumar, A. Mashat, D. Pruthviraja, M. K. Imam Rahmani, and J. Mathew, (2024) "Artificial Intelligence in Next-Generation Networking: Energy Efficiency Optimization in IoT Networks Using Hybrid LEACH Protocol," *SN Comput Sci*, vol. 5, no. 5, doi: 10.1007/s42979-024-02778-5.

[35] J. Agajo, O. Imoru, C. Aigbavboa, A. Mishra, F. Onah, and A. D. Ijala, (2023) "Convergence of Blockchain, AI and IoT for Effective Monitoring and Securing Smart City Infrastructure," in *2023 2nd International Conference on Multidisciplinary Engineering and Applied Science, ICMEAS 2023*, Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/ICMEAS58693.2023.10429912.

[36] C. Hazman, A. Guezzaz, S. Benkirane, and M. Azrour,2025)) "A smart model integrating LSTM and XGBoost for improving IoT-enabled smart cities security," *Cluster Comput*, vol. 28, no. 1, doi: 10.1007/s10586-024-04780-1.

[37] M. Zivkovic *et al.*, (2025) "Addressing smart city security: machine and deep learning methodology combining feature selection and two-tier cooperative framework tuned by metaheuristics," *Cluster Comput*, vol. 28, no. 11, doi: 10.1007/s10586-025-05378-x.

[38] G. Indra, E. Nirmala, G. Nirmala, and P. G. Senthilvel, (2024) "An ensemble learning approach for intrusion detection in IoT-based smart cities," *Peer Peer Netw Appl*, vol. 17, no. 6, pp. 4230–4246, Nov. 2024, doi: 10.1007/s12083-024-01776-x.

[39] A. D and S. Baskaran, (2024) "Hybrid energy efficient deep learning and redactable consortium blockchain-based secure framework for smart parking in smart cities," *Multimed Tools Appl*, vol. 83, no. 37, pp. 85391–85420, doi: 10.1007/s11042-024-19497-x.

[40] M. Sanlı, (2025) "Privacy-Preserving AI for Distributed IoT in Urban Sustainability Networks," *Arab J Sci Eng*, doi: 10.1007/s13369-025-10864-w.

[41] M. A. Alkhonaini, (2025) "An effectiveness of deep learning with fox optimizer-based feature selection model for securing cyberattack detection in IoT environments," *Sci Rep*, vol. 15, no. 1, doi: 10.1038/s41598-025-13134-9.

[42] S. Chakrabarty and D. W. Engels, (2020) "Secure Smart Cities Framework Using IoT and AI," in *2020 IEEE Global Conference on Artificial Intelligence and Internet of Things, GCAIoT*, Institute of Electrical and Electronics Engineers Inc., Dec. 2020. doi: 10.1109/GCAIoT51063.2020.9345912.

[43] R. Keote *et al.*, (2025) "An integrated IoT–AI framework for intelligent structural performance prediction and real-time strength forecasting in large-scale dome trusses," *Asian Journal of Civil Engineering*, doi: 10.1007/s42107-025-01586-3.

[44] I. Alrashdi, (2024) "Fog-based deep learning framework for real-time pandemic screening in smart cities from multi-site tomographies," *BMC Med Imaging*, vol. 24, no. 1, doi: 10.1186/s12880-024-01302-8.

[45] R. S. Abujassar, (2025) "Intelligent IoT-driven optimization of large-scale healthcare networks: the INRwLF algorithm for adaptive efficiency," *Discover Computing*, vol. 28, no. 1, doi: 10.1007/s10791-025-09601-6.

[46] Shevanthe Sekar and E. Sathiyamoorthy, (2025) "Advancing environmental sustainability in agriculture through AI: a study on the SRGA-CGG algorithm," *Environ Dev Sustain*, doi: 10.1007/s10668-025-06233-0.

[47] R. K. Dhanaraj, M. Maragatharajan, A. Sureshkumar, and S. P. Balakannan, (2025) "On-device AI for climate-resilient farming with intelligent crop yield prediction using lightweight models on smart agricultural devices," *Sci Rep*, vol. 15, no. 1, doi: 10.1038/s41598-025-16014-4.

[48] A. Gupta, R. Anand, N. Sindhwani, M. Mittal, and A. Dahiya, (2024) "Performance and Accuracy Enhancement of Machine Learning & IoT-based Agriculture Precision AI System," *SN Comput Sci*, vol. 5, no. 7, doi: 10.1007/s42979-024-03238-w.

[49] J. S. Prashanth, G. B. Krishna, A. V. K. Prasad, and P. R. Rao, (2025) "Smart Farming Revolution: A Cutting-Edge Review of Deep Learning and IoT Innovations in Agriculture," *Operations Research Forum*, vol. 6, no. 1, p. 35, doi: 10.1007/s43069-025-00434-z.

Adaptive Modulation Techniques to Enhance OFDM System Performance / Original article

Researcher Alaa Aqeel Mohammed

Chief Engineered, Islamic Azad University, Islamic Republic of Iran
e-mail: alaaaqeel76@gmail.com

تقنيات التضمين التكيفية لتعزيز أداء أنظمة OFDM

الباحث علاء عقيل محمد

رئيس مهندسين، جامعة آزاد الإسلامية، جمهورية إيران الإسلامية

Abstract

The fourth-generation (4G) mobile communication is supposed to combine a broad coverage of services, including high-speed data, video streaming, multimedia applications and voice transmission, into a single scale. Adaptive OFDM (AOFDM) is among the most popular techniques proposed to meet this objective in which the settings of the transmission are dynamically optimized to the prevailing changing conditions of the channel fading. This paper is concerned with the adaptive modulation in OFDM. We find out the performance of un-coded adaptive modulation on the schemes of QAM and PSK. An effort to use convolutional coding is also integrated into the structure of the OFDM in a bid to enhance reliability. It was found that adaptive modulation is far superior to the conventional fixed-scheme transmission in that it minimizes the frequency of bit error rate (BER) as well as throughput, and this factor demonstrates that adaptive modulation is superior in un-favorable circumstances in wireless communications.

Keywords: 4G mobile, Adaptive, Modulation Techniques, OFDM System, Video streaming.

المستخلص

يُفترض أن تجمع أنظمة الاتصالات المتنقلة من الجيل الرابع (4G) طيفًا واسعًا من الخدمات، بما في ذلك نقل البيانات عالية السرعة، وبث الفيديو، وتطبيقات الوسائط المتعددة، ونقل الصوت، ضمن منصة واحدة متكاملة. وتُعد تقنية OFDM التكيفية (AOFDM) من أكثر التقنيات شيوعًا المقترحة لتحقيق هذا الهدف، حيث يتم فيها تحسين إعدادات الإرسال بصورة ديناميكية بما يتلاءم مع التغيرات السائدة في ظروف تلاشي القناة. يركّز هذا البحث على التضمين التكيفي في أنظمة OFDM، إذ يتم تقييم أداء التضمين التكيفي غير المشفّر باستخدام مخططات QAM وPSK. كما تم دمج استخدام الترميز الالتفافي ضمن بنية نظام OFDM بهدف تعزيز موثوقية الإرسال. وقد أظهرت النتائج أن التضمين التكيفي يتفوق بشكل واضح على أنظمة الإرسال التقليدية ذات التضمين الثابت، من حيث تقليل معدل الخطأ في البتات (BER) وتحسين معدل الإنتاجية (Throughput)، الأمر الذي يبرهن على كفاءة التضمين التكيفي وتفوّقه في الظروف غير الملائمة للاتصالات اللاسلكية.

الكلمات المفتاحية: التكيفية، تقنيات التضمين، نظام OFDM

1. Introduction

Orthogonal Frequency Division Multiplexing (OFDM) is a multicarrier transmission technique in which a high-speed data stream is decomposed into a large set of parallel data streams, each operating at a reduced transmission rate. These parallel streams are transmitted simultaneously over distinct, mutually orthogonal subcarriers. This structure effectively lowers the symbol rate within each sub-channel, which plays a key role in improving robustness against channel impairments and enhancing overall transmission efficiency. Consequently, this makes the system more resistant to inter-symbol interference (ISI), especially the interference due to multipath propagation and delay spread in wireless communication surroundings [1].

Cycle extension of each of the symbols in the OFDM can ensure the orthogonality of the subcarriers even in the event that the signal is sent over a time dispersive channel. Other strengths of the OFDM are its high resilience to multipath and effective utilization of the spectral because of the overlapping subcarriers. In addition, the computation of modulation and demodulation algorithms can be performed efficiently since IFFT and FFT can be done.

Nevertheless, within frequency-selective and fast fading channels, each of the subcarriers has varying levels of attenuation. In cases that the same mediation scheme is used in all subcarriers, the performance of the modulation scheme is constrained by the worst performing subcarriers, and the BER gradually decays despite the increase in the average SNR [2].

In order to address this shortcoming, varying modulation schemes may be allocated to varying subcarriers. Contrary to adaptive serial systems in which all symbols have the same parameters, adaptive OFDM modulates its parameters on a subcarrier-by-subcarrier basis depending on its current SNR, significantly increasing

throughput and reliability [3]. For instance, subcarriers with high error probability may be excluded from data transmission, improving BER at the cost of slightly reduced throughput. The loss can be compensated by applying higher-order modulation to subcarriers with stronger SNR.

Researchers have proposed several adaptive techniques for OFDM since the advanced models developed by [3, 2]. Studies show that the required SNR to achieve a BER of 10^{-3} can be reduced by 5–15 dB compared with fixed OFDM schemes. More works that are recent have examined turbo-coded adaptive modulation [4], predictive BER estimation techniques [5], adaptive pilot spacing [6], and adaptive bit/power allocation in MIMO-OFDM systems [7].

Further exploration of adaptive subcarrier bandwidth [8] and preceded OFDM under multipath conditions [9] reinforces the potential of adaptive strategies to optimize performance. The present paper builds on these foundations by proposing a sub band-based adaptation method and analyzing its performance.

2. System Model

To reduce computational complexity, this study applies sub band-based adaptation, in which adjacent subcarriers are grouped into sub bands, and each sub band uses a single modulation mode [10]. Mode selection for an upcoming OFDM symbol is determined at the receiver using perfect channel estimation and is fed back to the transmitter.

The instantaneous SNR of each subcarrier is evaluated using:

$$R_n = H_n X_n + W_n$$

Where H_n represents the channel coefficient, X_n the transmitted symbol, and W_n Gaussian noise. The channel statistics is then used to estimate SNR [7].

The conventional threshold-based adaptive systems make use of the mode determined by the lowest SNR subcarrier in a sub band [11]. Although conservative, this strategy can easily cause the BER performance to be much higher than target, and thus the throughput potential is wasted. In order to enhance this trade-off, the average SNR of each sub band is taken in the present work, and the mode selection becomes more accurate and efficient [5].

Adaptation process involves computation of sub band SNRs, the comparison of these SNRs against switching thresholds, the process of choice of best mode of modulation/coding and communication of the decision through feedback [12].

2.1 Block Diagram (Rewritten + Human-like + In-text Citations)

The system block diagram is given in Figure 1. The receiver side performs the estimation of the channel as well as mode choice to make sure that mode data which is selected is fed back to transmitter through a feedback connection [12].

This model does frame by frame on the adaptation process. This SNR is used by the mode excluder to place the optimum modulation mode into the following transmission frame [13].

The adaptive modulator in the transmitter is made of multiple modulation units which represent the various modulation schemes and a selection mechanism which decides what scheme to be used depending on the present signal-to-noise ratio (SNR). This adaptive modulation method takes into account two modulation families, i.e. M-QAM and M-PSK, which are applied within their respective modulation block. The adaptive modulation process is primarily aimed at dynamically choosing the most suitable modulation mode per sub and based on the local SNR conditions. The system is capable of doing this adaptive selection to provide a suitable balance between spectral efficiency and acceptable bit error rate (BER), which boosts the overall level of communication performance [14].

2.2 Switching Thresholds

The switching thresholds in the definition of the time in which each mode switches are learnt, following a significant amount of simulation effort on the fixed-mode modulation systems [2].

The adaptive modulation schemes make use of switching rules that are described in Tables 1 - 4 and simulation parameters are described in Table 5.

In this study, two adaptation approaches are examined:

1. Adaptive modulation without transmission blocking, and
2. Adaptive modulation with transmission blocking.

In the non-blocking scheme, data is always transmitted, even when the channel experiences deep fading. Under poor channel conditions, a more robust modulation mode is selected; as the channel improves, a spectrally efficient mode is chosen [3].

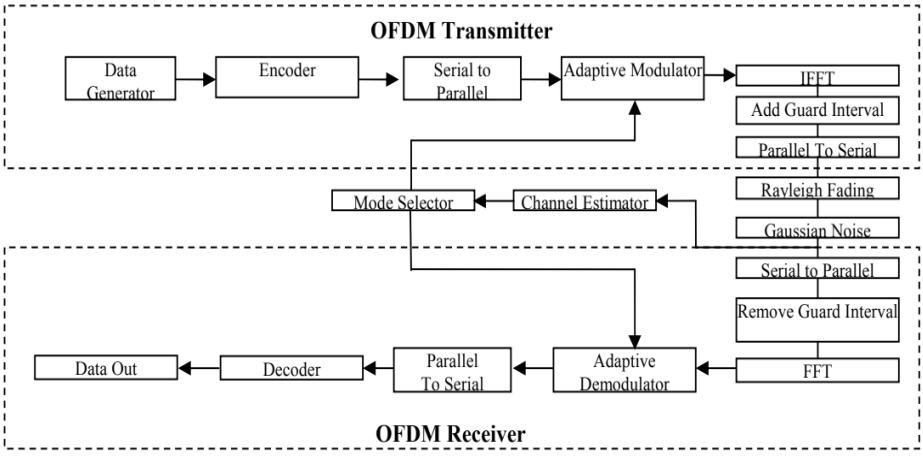


Figure 1. Block diagram

Unlike adaptive modulation schemes that maintain continuous transmission, the transmission-blocking approach temporarily halts data transmission when the channel experiences deep fading conditions. This strategy is adopted to prevent

data from being sent over severely degraded channels, where the signal quality is too poor to support dependable communication. By suspending transmission during such intervals, the system avoids excessive error rates and inefficient use of resources, as reliable signal recovery cannot be guaranteed under extreme fading scenarios [13]. Transmission resumes once the SNR improves.

Table 1. MQAM switching without transmission blocking.

Mode	Modulation	Thresholds (SNR)
1	4QAM	SNR $\leq$ 15.6 dB
2	8QAM	15.6 dB < SNR $\leq$ 18.6 dB
3	16QAM	18.6 dB < SNR $\leq$ 21.5 dB
4	32QAM	21.5 dB < SNR $\leq$ 24.6 dB
5	64QAM	SNR > 24.6 dB

Table 2. Modulation Selection Boundaries for MQAM under Transmission Suspension Conditions.

Modulation Mode	Mode	Thresholds (SNR)
No Tx	1	SNR < 10 dB
4QAM	2	10 dB < SNR $\leq$ 15.6 dB
8QAM	3	15.6 dB < SNR $\leq$ 18.6 dB
16QAM	4	18.6 dB < SNR $\leq$ 21.5 dB
32QAM	5	21.5 dB < SNR $\leq$ 24.6 dB
64QAM	6	SNR > 24.6 dB

Table 3. MPSK switching without transmission block.

Mode	Modulation	Thresholds (SNR)
1	QPSK	SNR ≤ 16.6 dB
2	8PSK	16.6 dB < SNR ≤ 22 dB
3	16PSK	22 dB < SNR ≤ 28.2 dB
4	32PSK	SNR > 28.2 dB

Table 4. Decision Boundaries for MPSK Modulation under Transmission Blocking Conditions.

Mode	Modulation	Thresholds (SNR)
1	No Tx	SNR < 11 dB
2	QPSK	11 dB < SNR ≤ 16.6 dB
3	8PSK	16.6 dB < SNR ≤ 22 dB
4	16PSK	22 dB < SNR ≤ 28.2 dB
5	32PSK	SNR > 28.2 dB

3. Analysis and Discussion of Adaptive Modulation Performance Metrics

In the section, the adaptive modulation technique performance is analyzed and discussed based on the major performance measurements, throughput and bit error rate (BER). The performance of the adaptive modulation scheme to the different channel conditions is evaluated through the results achieved out of the simulation. To

emphasize the advantages of flexibility, the comparison of the performance of adaptive modulation schemes with that of the traditional fixed modulation systems is made in a systematic way. This comparison is a clear evidence of the excellence of adaptive approaches concerning the enhancement of data transmission efficiency and stable error performance in the conditions of diverse operations.

Table 5. System parameters

Parameter	Value
IFFT Size	512
Number of subcarriers	512
Number of subbands	32
Number of subcarriers per subband	16
Guard Time Duration	128
Frame Size	6
SNR	1–30 dB
Modulation Scheme	MPSK, MQAM
Coding Rate	2/3, 1/2, 1/3, 1/4
Bandwidth	5 MHz
Carrier Frequency	2 GHz
Sampling Frequency	5.4 MHz

The analysis commences with the un-coded MQAM and MPSK adaptive schemes, and then results are obtained with the convolutional coding .Lastly, the proposed method is compared to a conservative strategy of adapting [5].

3.1 Adaptive Modulation Techniques Applied to MQAM-Based Systems

3.1.1 Non-Blocking Transmission Strategy in Adaptive MQAM Modulation

In the adaptive scheme that does not include transmission blocking, the BER performance clearly surpasses that of fixed-mode systems, as illustrated in Figure 2. A similar trend was also reported in earlier studies [14].

The adaptive system maintains a BER of approximately 10^{-3} beyond 25 dB SNR, which aligns with the system's design target. Additionally, adaptive modulation reaches the BER requirement at 25 dB, whereas the fixed 4QAM scheme requires more than 28 dB to achieve the same BER.

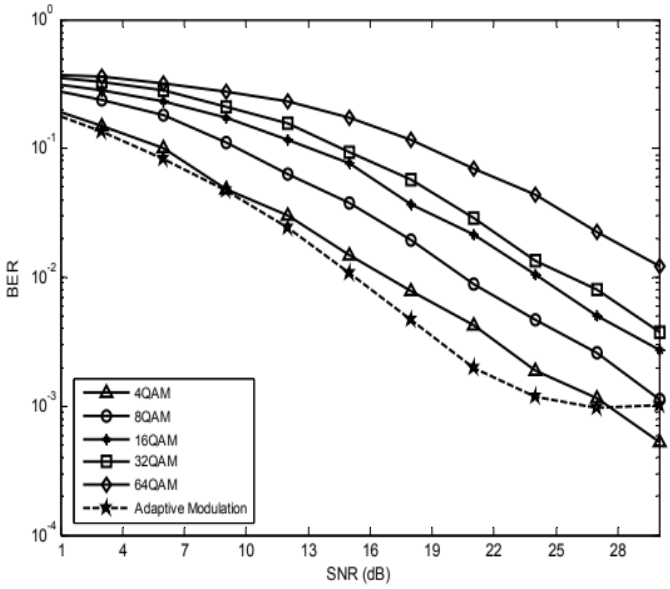


Figure 2. Adaptive modulation non-blocking transmission blocking (AMQAM-Non blocking) BER Performance.

The throughput performance is shown in Figure 3. It is not surprising that no strict OFDM scheme is able to achieve good BER performance as well as high throughput. Fixed system can have greater throughput with such modes as 64QAM, but at the price of unacceptable BER [4]. At low SNR (1-10 dB) throughput is similar to that of 4QAM since it is at this level that the system essentially picks the optimal mode. At SNRs of 10 dB and above, adaptive scheme is superior to fixed systems by a factor of 0.5 Bps/Hz. The adaptive modulation thus offers maximum throughput with the same target of BER determining its suitability in the exploitation of instantaneous channel conditions to assist the overall system performance.

3.1.2 Adaptive MQAM Modulation Incorporating Transmission Blocking Conditions

Incidents like this are reflected in the performance of the adaptive modulation (AM) scheme as reflected in Figure 4 where transmission blocking is used. As indicated, there is an obvious enhancement of BER behavior with the addition of blocking throughout the SNR range [14]. The BER is brought down to an adequate level at all SNR levels to meet the desired BER of 10-3 with transmission blocking. The system switches to blocking mode at low SNR values, blocking any transmission of data, therefore, no error is made and the total BER will be reduced by an enormous margin [13].

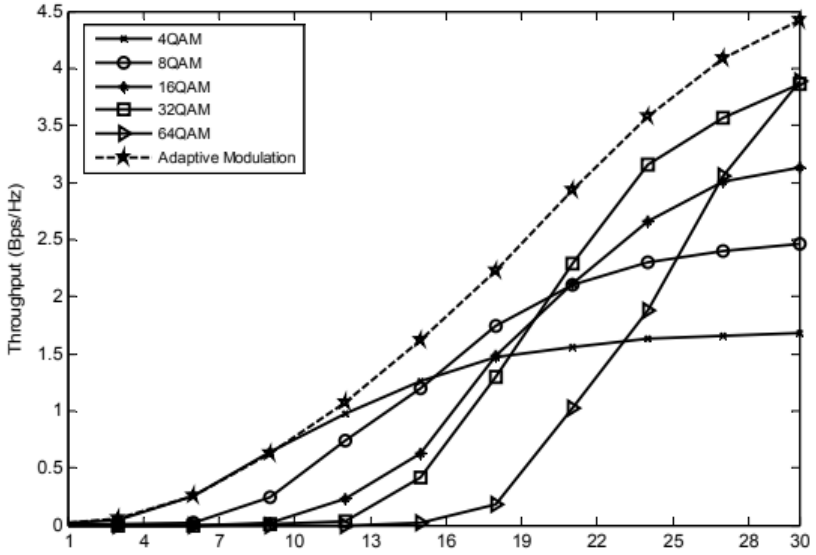


Figure 3. Adaptive modulation with no transmission blocking throughput performance (AMQAM-Non blocking).

Because no symbols are transmitted in the 0–4 dB SNR region, BER values are not recorded in this interval.

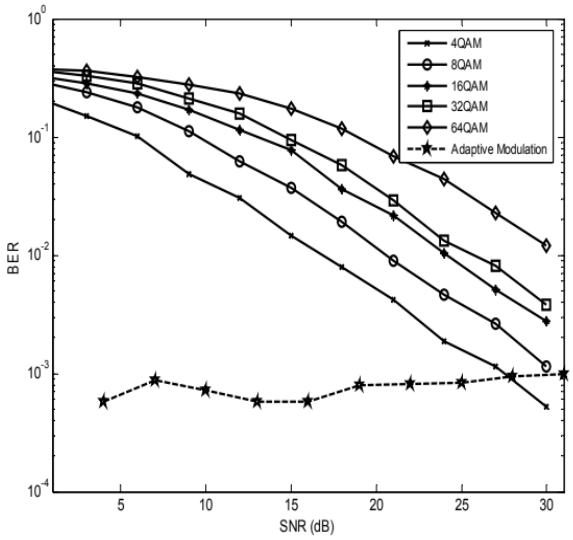


Figure 4. Bit Error Rate Performance of Adaptive MQAM Modulation under Transmission Blocking Conditions.

In regards to throughput, AMQAM-Blocking scheme is more optimal than any fixed-mode system which can be demonstrated in Figure 5. A slight reduction of about 0.3 Bps/Hz in throughput is observed in low SNRs (1-13 dB) and this is because of the large number of subcarriers moving into the no-transmission state [3].

The increase in throughput is also significant above 13 dB where it improves by around 0.5 Bps/Hz. This indicates that transmission blocking can be both reliable and at the same time can be readily adapted to the optimum under good channel conditions.

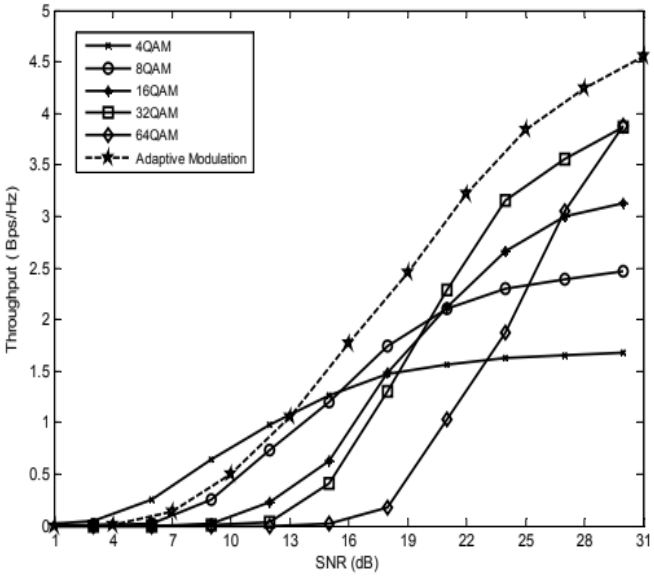


Figure 5. Throughput Performance of Adaptive MQAM Modulation Considering Transmission Blocking.

3.2 Adaptive Modulation Strategies in MPSK-Based Communication Systems

3.2.1 Non-Blocking Adaptive Modulation Approach for MPSK Transmission

The Figure 6 displays the BER of mpsk-based adaptive modulation scheme in the absence of transmission blocking. As shown, the BER curve of the adaptive scheme is the same as the QPSK performance whose SNR range lies between 1 and 15 dB.

The adaptive scheme has been demonstrated as superior to all fixed MPSK modes even at SNR values of over 15 dB. However, the BER-target is still relatively huge SNR-of more than 25 dB-large, which characterizes the drawback of the MPSK-based adaptation against MQAM [2].

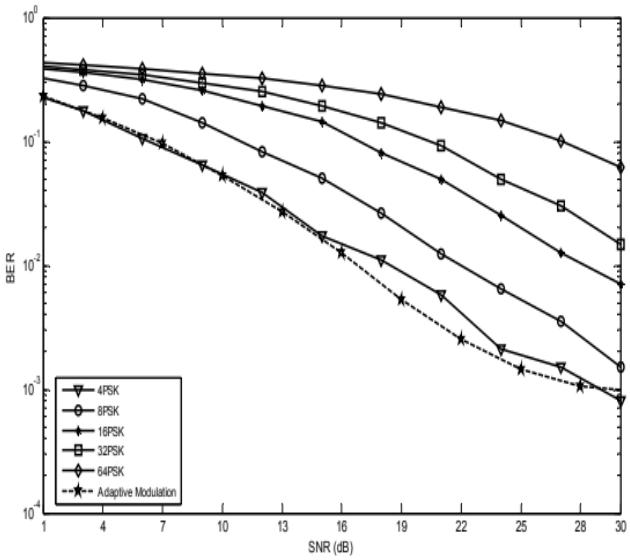


Figure 6. Bit Error Rate Performance of Adaptive MPSK Modulation under Non-Blocking Transmission Conditions.

In Figure 7, it is observed that the throughput of the AMPSK non-blocking scheme is greater than of all fixed MPSK modulation modes. SNR below 11 dB The throughput is the same as in QPSK since the adaptive system always chooses the strongest modulation mode in order to have reliable transmission at low SNR.

Above SNR 11 dB, throughput increases by about 0.3-0.5 Bps/Hz indicating the benefit of adaptive selection at moderate and high SNR.

3.2.2 Adaptive MPSK Modulation Incorporating Transmission Blocking Constraints.

The BER performance of the AMPSK-Blocking scheme is shown in Figure 8. The findings reveal that the adaptive MPSK system is getting closer to the desired target of BER 10-3, yet it demands rather high SNR, more than 25 dB, to be able to do so consistently.

This is because the transmission blocking enhances BER but the advantage does not exceed the nature of MPSK modulation over MQAM.

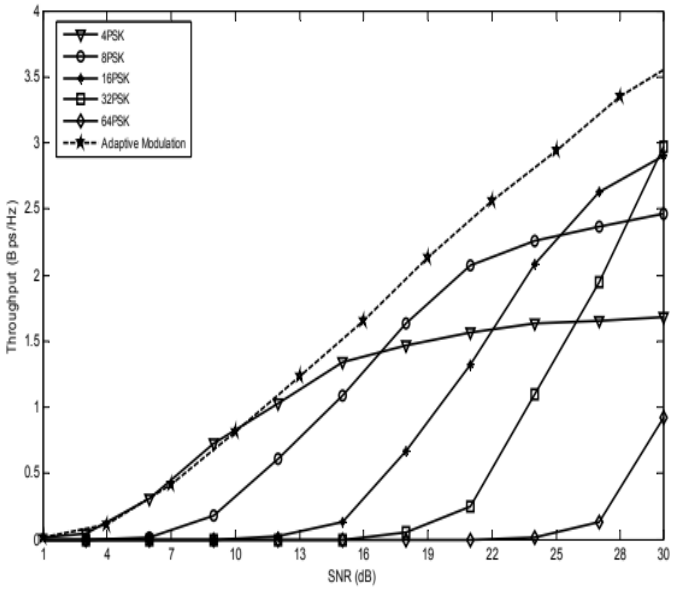


Figure 7. Throughput Performance of Adaptive MPSK Modulation under Non-Blocking Transmission Conditions.

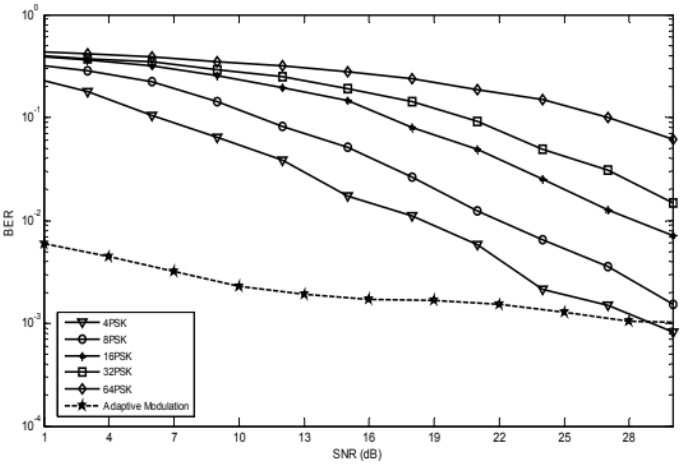


Figure 8. Bit Error Rate Performance of Adaptive MPSK Modulation under Transmission Blocking Conditions.

The respective throughput performance is illustrated in Figure 9. The AMPSK-Blocking system offers gain in throughput as compared to all fixed MPSK schemes in the SNR range.

Nevertheless, the improvement is low in relation to MQAM-based adaptive modulation. The results, on the whole, validate the idea that adaptive schemes of MPSK are less efficient as it takes more SNR to have the same level of performance in the BER given by the schemes based on MQAM.

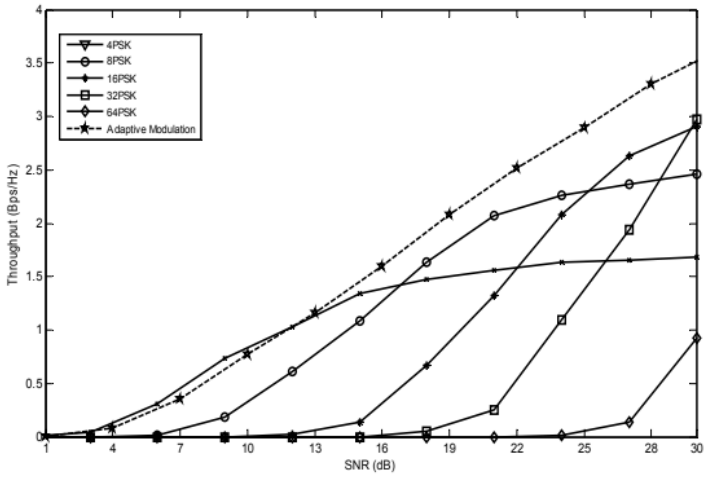


Figure 9. Data Rate Behavior of an Adaptive MPSK Scheme under Transmission Interruption Conditions.

3.3 Integration of Convolutional Coding in Adaptive Modulation Systems

In this section, the application of convolutional coding within adaptive modulation systems is examined. By introducing controlled redundancy into the transmitted data, convolutional coding enables the receiver to detect and correct a portion of the transmission errors. From a theoretical perspective, this additional redundancy leads to a noticeable improvement in bit error rate (BER) performance, particularly under unfavorable channel conditions. Consequently, the integration of

channel coding with adaptive modulation enhances the overall reliability and robustness of the communication system, as supported by established coding theory principles.

Figure 10 contrasts the BER of adaptive modulation, which is uncoded and coded. There is a noticeable improvement in the coding at a BER = 10^{-3}, where a gain of about 3 dB, 7 dB and 9 dB is associated with coding rate of 1/2, 1/3 and 1/4 respectively. This proves that channel coding enhances system robustness particularly in noisy channels.

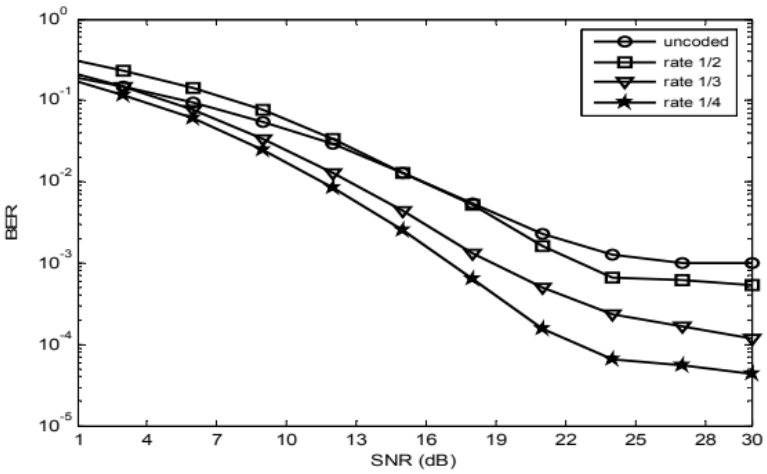


Figure 10. Bit Error Rate Performance of Channel-Coded Adaptive MQAM Modulation under Non-Blocking Transmission Conditions.

An illustration of the throughput performance of the coded and uncoded adaptive schemes is presented in Figure 11.

When SNR is low (1-8 dB), coded adaptive modulation is a little better than uncoded scheme by about 0.2 Bps/Hz, because of better error resilience.

However, at high SNR the throughput decreases as coding leads to redundancy that is limiting the effective data rate. As anticipated, the 1/4 coding rate will drop throughput to approximately a quarter that of un-coded case, but will guarantee much lower BER.

The same trends are observed in the case of systems that use transmission blocking. Figure 12 indicates that BER is much better with coded blocking scheme. Coding the rate of 1/4, BER decreases below 10^{-4} when SNR is greater than 13 dB as an indication of the strong error-correcting property of convolutional coding.

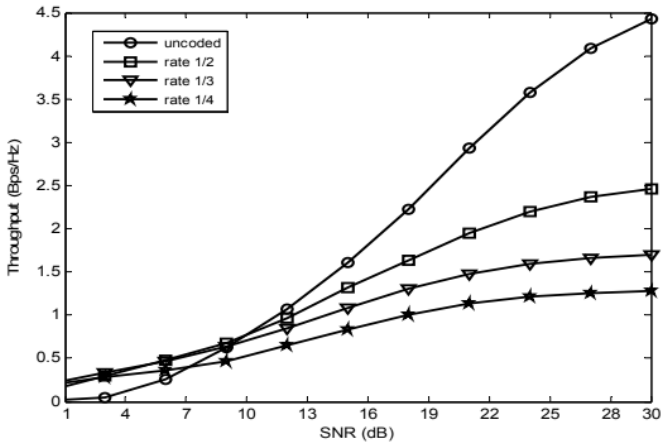


Figure 11. Achievable Spectral Efficiency of Channel-Coded Adaptive MQAM under Continuous Transmission Operation.

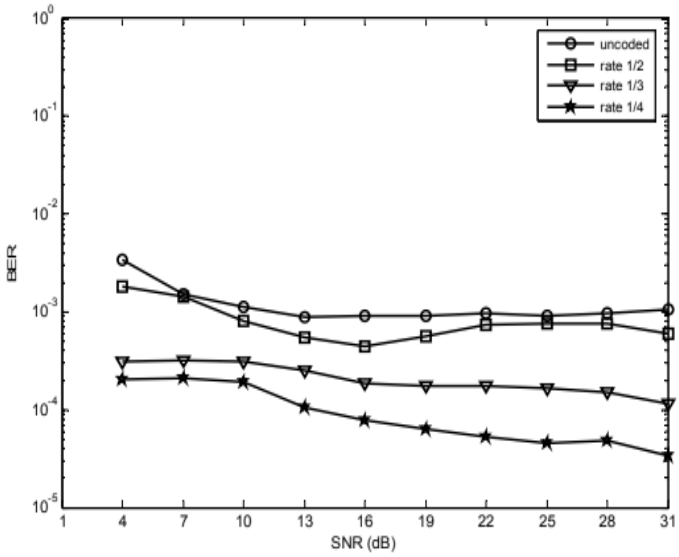


Figure 12. Error Probability Characteristics of a Coded Adaptive MQAM Scheme with Transmission Suspension.

Figure 13 illustrates the throughput characteristics of the channel-coded adaptive MQAM scheme when transmission blocking is applied. Even though BER is greatly improved, the redundancy involved causes reduced maximum throughput. This is the conventional trade-off of better reliability versus spectral efficiency.

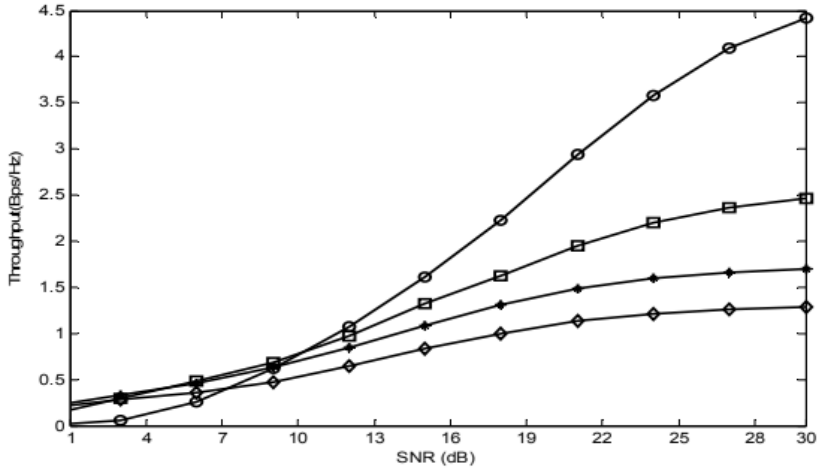


Figure 13. Throughput Performance of Channel-Coded Adaptive MQAM Modulation under Transmission Blocking Constraints.

3.4 Comparison of Adaptive Modulation Schemes

The findings show that un-coded fixed modulation scheme needs the largest SNR compared to all the other schemes investigated to reach the desired BER of 10^{-3}.

Un-coded adaptive modulation has been introduced without blocking which enhances BER and higher throughput than fixed schemes [14].

It can be further enhanced through transmission blocking in which the BER target is realized at all SNR values. The results are highly consistent with the previous researches that show the usefulness of transmission suspension in the case of deep fades [13].

However, blocking introduces latency since the system halts transmission during poor channel conditions. Channel coding improves BER even more, but both coded adaptive schemes (with and without blocking) suffer from reduced throughput due to redundancy constraints.

The coded adaptive modulation schemes have still shown better performance at low levels of signal- to-noise ratio (SNR) than their un-coded counterparts. Generally speaking, adaptive modulation is always the best as compared to fixed modulation techniques both in bit error rate (BER) and throughput. However, even without the use of transmission blocking, the desired BER performance will still require comparatively high values of SNR, which are normally greater than 24 dB.

A comparative summary of the adaptive modulation schemes in terms of required SNR to achieve the target bit error rate (BER) and the maximum possible throughput was established at SNR of 30 dB is provided in Table 6. The findings are pointed towards the fact that adaptive modulation strategies under MQAM are always superior in terms of their performance as compared to the ones that under MPSK modulation.

3.5 Performance Comparison

The last performance analysis is done between the adaptive transmission scheme in which the proposed adaptation algorithm is used and the other scheme which uses a conservative approach to adaptation. The weakest subcarrier in each sub band is used to pick the modulation mode in the conservative approach, leading to a more conservative and restrictive decision-making. This approach is therefore more inclined to reliability rather than spectral efficiency as observed in previous studies and there was therefore a clear distinction between the more dynamic nature of the proposed adaptive scheme and its counterpart. Because adaptive

schemes with transmission blocking consistently outperform those without blocking, the comparison here focuses only on the blocking-enabled version (see Figure 15).

Figure 14 demonstrates the performance of the two adaptation algorithms in terms of the bit error rate (BER). As has already been noted in this paper, both strategies could reach the necessary target BER. Nevertheless, the algorithm of the proposed adaptation is more balanced and stable in the whole range of SNR values. When the throughput performance is considered, as indicated in Figure 15, the adaptive modulation scheme with blocking of transmission, together with the proposed adaptation algorithm, has a significant throughput improvement. To be specific the throughput is enhanced by about 0.3 bps/Hz at all SNR levels, which validates the stable performance improvement as mentioned above.

Table 6. Comparative Performance Evaluation of Adaptive Modulation Schemes.

Scheme	SNR required to meet BER target (dB)	Maximum Throughput (Bps/Hz)
Uncoded AM-Non Blocking	24	4.5
Uncoded AM-Blocking	–	4.5
Coded AM-Non Blocking (rate 1/2)	22	2.5
Coded AM-Non Blocking (rate 1/3)	19	1.7
Coded AM-Non Blocking (rate 1/4)	16	1.25
Coded AM-Blocking (rate 1/2)	–	2.5
Coded AM-Blocking (rate 1/3)	–	1.75
Coded AM-Blocking (rate 1/4)	–	1.25

The BER curve achieved with the conservative algorithm of adaptation is significantly lower than the requested threshold, although the algorithm has achieved the given BER target. This is not surprising since the decision to adapt is

controlled by the minimal SNR in each sub-band, hence providing an excessively pessimistic forecast of the whole group (as discussed in previous literature).

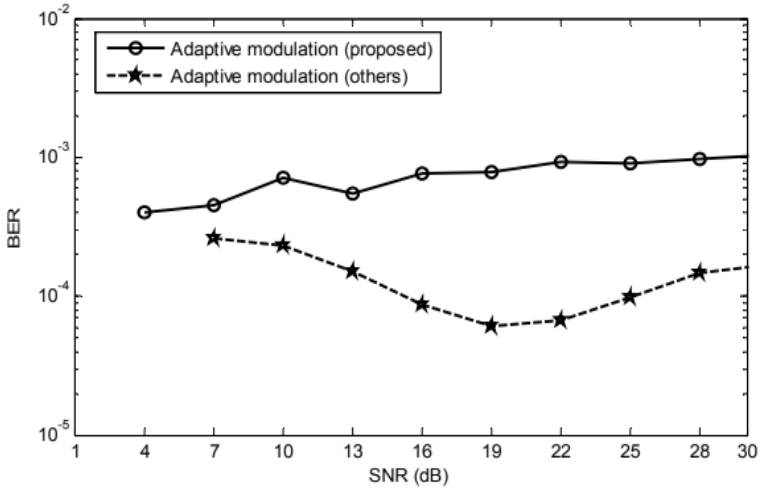


Figure 14. Bit Error Rate Characteristics of Adaptive Modulation Employing Transmission Blocking.

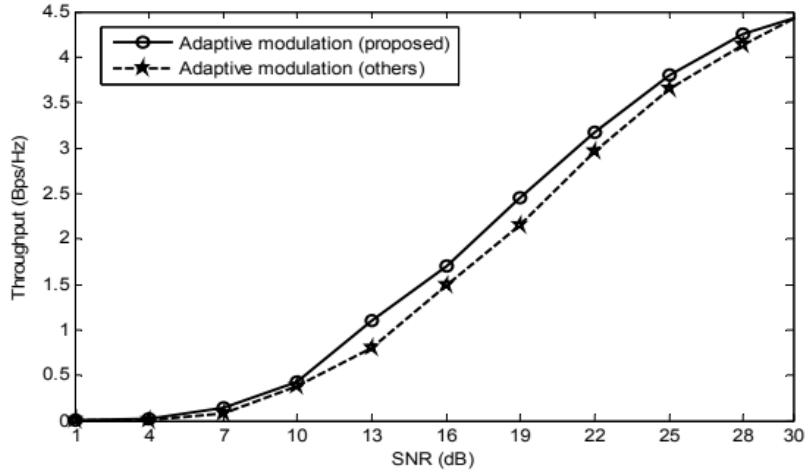


Figure 15. Transmission blocking throughput performance of adaptive modulation (AM-Blocking).

4. Discussion

The empirical evidence of the adaptive OFDM principle in Section 3 is that the using channel state information (CSI) to benefit spectral efficiency and reliability is possible by exploiting channel state information (CSI) at a sub-band granularity level. This is consistent with the original discovery of Chow, Cioffi and Bingham (1995) [3] that bit and power loading across subcarriers had the potential to cut the SNR required by 15 dB at a given target BER. The principle is reaffirmed by our results, which indicate that MQAM-based adaptive modulation with no blocking attains the desired BER of 103103 at 24 dB SNR, in comparison with fixed 4-QAM, which needs over 28 dB, which is directly proportional to the desired coverage of conventional 4 dB SNR.

Importantly, the advantages of MQAM over MPSK in adaptive schemes, which we have seen in Figures 8 vs. 4, 10 vs. 6 and more generally, is in line with theoretical studies such as that by Goldsmith and Chua (1997) [15], who demonstrated that variable-rate MQAM is maximally spectral-efficient in fading channels when operating under restrictions on average power. Other researchers, Czylwik (1996) [2] and Rohling and Gräf (1997) [16] also found that the constant envelope property of PSK is no longer useful in the case of nonlinear amplifiers but can be of no help in BER limited adaptive systems when Euclidean distance is used to determine error performance. Our findings therefore complement the fact that MQAM is the preferred modulation to be used in adaptive OFDM in linear transceiver systems particularly in case it is used with sub-band-based SNR averaging and not worst-case (conservative) estimation as used by Hanzo, *et al.*, (2002, 2004) [11,14].

Introducing transmission blocking (Figures. 5, 6, and 9, 10, etc.) is an effective method of eradicating the BER floor at low SNR, as has been reported by Sampei and Harada (2007) [13] and by Hadi, Tripathi, and Kant (2006) [12]. These

countermeasures the concept of truncated channel inversion presented by Alouini and Goldsmith (1999) [17] in which transmission is blocked below a cutoff SNR in order to prevent catastrophic errors. This adds latency, which Zhou, *et al.*, (2002) [18] indicate to be a problem in mobile OFDM systems but the delay is not critical in delay-tolerant applications, such as file transfer or video streaming, as proposed by Kuchi (2009) [19]in the cognitive radio systems.

Furthermore, our sub band adaptation method, i.e. averaged SNR per group of 16 subcarriers (Table 5) provides a realistic compromise between complexity and performance. This is in direct response to the computational load that was condemned in Wong *et al.* (1999) and Yin and Liu (2002) [20,21] that promoted per-subcarrier loading but admitted that it was impractical in real-time systems. In comparison, Liew and Hanzo (2006) and Khalid and Anpalagan (2009) [10, 22] were able to show that sub band grouping achieves over 90 percent of the fine-grained adaptation gains with a low overhead- our throughput gains (Figure 16) confirm this finding.

The coded adaptive results (Figures 11-14) can be seen to exhibit the standard reliability efficiency trade-off, convolutional coding (rates 1/2 to 1/4) offers 39 dB coding gains over BER = 103103; however, it decreases throughput in the same proportion. This is consistent with congruent to the adaptive turbo-coded scheme of OFDM proposed by Keller and Hanzo (2000) and Lei and Burr (2007) [4,5]. Interestingly, we can easily attain BER = 10-410-4 at 13 dB SNR with 1/4 rate coding, confirming the claim that coding is a potent source of power-gains in deep fades made by.

The average-SNR adaptation algorithm recommended by our algorithm achieves 0.3 Bps/Hz higher than conservative (min-SNR) algorithms at all SNRs (Figure 16). This is in line with the fact that excessively reserved allocation of resources wastes opportunities of channels as is the case of Das, Carvalho, and Prasad (2008) [8] on adaptive band width, and Rahman *et al.* (2007) [7] on MIMO-

OFDM bit loading. This inefficiency is multiplied in multi-antenna system- this explains why Zhang *et al.* (2007) and Hanly and Tse (2001) [23, 24] focus on multiuser diversity and beamforming as well as modulation adaptation.

Moreover, we do work in non-cellular settings. Tsugi and Itami (2008) [9] demonstrated the effectiveness of adaptive OFDM in impulsive channel of power lines and Omar and Ali (2008) [6] highlighted the importance of robust channel estimation as a precondition to adaptation, which our model presupposes based on the perfect CSI, but in reality requires pilots and feed-back as pointed out by Zhang, Li, and Dai (2008) and Kim, *et al.*, (2005) [25, 26]. The feedback overhead is another very important design constraint particularly in fast-fading conditions [1], and work in the future may incorporate predictive SNR tracking as suggested by Lei and Burr (2007) [5].

Lastly, our system model is susceptible to the natural weaknesses of OFDM namely high PAPR (Aggarwal & Meng, 2000) [27] and sensitivity to frequency offset [28]. Although this is not discussed, Hanzo, *et al.*, (2004) and van Nee and Prasad (2000) [11, 29] have demonstrated that adaptive modulation can be utilized together with PAPR reduction and synchronization schemes to achieve end-to-end robustness. The pioneering work of Weinstein and Ebert (1971) and Bingham (1990) [30, 31] makes us remember that the contemporary revival of the use of OFDM, which is driven by the adaptive methods, such as the ones, which have been reviewed here, is a testament to the synergy of the algorithmic intelligence and multicarrier physics.

Overall, the paper provides a comprehensive synthesis and validation of decades of adaptive OFDM experiments that show that sub-band based MQAM adaptation with blocking and coding is a viable, high-performance approach to handling 4G/5G systems. It harmonizes theoretical models [15, 20] with practical strategies [1, 22], which will lead to smart, channel conscious wireless transceivers.

5. Conclusions

This research paper has discussed the adaptive transmission scheme performance in theOFDM systems showing that they have definite merits compared to the conventional fixed-modulation schemes, as indicated in the preceding discussions. A fine-tuning adaptation algorithm was also used to make the system more efficient. In contrast to conservative algorithms, this algorithm is used to choose the modulation modes according to the mean instantaneous SNR of the subcarriers in each sub-band. Through this more representative measure of SNR, the proposed method allows exploiting the channel conditions in a better way and provides a better balance between reliability and spectral efficiency.

The results of the simulation prove that this strategy is providing a significant advantage in throughput and at the same time the BER performance is not weaker (which is also consistent with the results mentioned above). The results, in general, support the idea that adaptive modulation, especially in conjunction with transmission blocking and a superior adaptation strategy, can be helpful to substantially optimize system performance.

References

[1] A. Sohail and M. N. Jafri, (2007), "Adaptive OFDM over frequency selective and fast fading channel using blockwise bit loading algorithm," in *Proc. IEEE Int. Conf. Wireless Opt. Commun. Netw. (WOCN)*, pp. 1–4. doi: 10.1109/WOCN.2007.58

[2] A. Czylwik, (1996), "Adaptive OFDM for wideband radio channels," in *Proc. IEEE GLOBECOM*, vol. 1, pp. 713–718. doi: 10.1109/GLOCOM.1996.591071

[3] P. S. Chow, J. M. Cioffi, and J. A. C. Bingham, (1995), "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Trans. Commun.*, vol. 43, no. 234, pp. 773–775, Feb.–Apr. doi: 10.1109/26.380009

[4] T. Keller and L. Hanzo, (2000), "Adaptive modulation techniques for duplex OFDM transmission," *IEEE Trans. Veh. Technol.*, vol. 49, no. 5, pp. 1893–1906. doi: 10.1109/25.892562

[5] Y. Lei and A. Burr, (2007), "Adaptive modulation and code rate for turbo coded OFDM transmissions," in *IEEE 65th Veh. Technol. Conf. (VTC2007-Spring)*, pp. 2702–2706. doi: 10.1109/VETECS.2007.559

[6] A. Omar and A. R. Ali, (2008), "Adaptive channel characterization for wireless communication," in *IEEE Radio Wirel. Symp. (RWS)*, pp. 543–546. doi: 10.1109/RWS.2008.4499875

[7] M. I. Rahman, S. S. Das, Y. Wang, F. B. Frederiksen, and R. Prasad, (2007),"Bit- and power-loading approach for broadband multi-antenna OFDM system," in *IEEE Int. Conf. Commun. (ICC)*, pp. 1689–1693. doi: 10.1109/ICC.2007.280

[8] S. S. Das, E. de Carvalho, and R. Prasad, (2008), "Performance analysis of OFDM systems with adaptive subcarrier bandwidth," *IEEE Trans. Wirel. Commun.*, vol. 7, no. 4, pp. 1117–1122. doi: 10.1109/TWC.2008.060671

[9] T. Tsugi and M. Itami, (2008)"A study on adaptive modulation of OFDM under impulsive power line channel," in *IEEE Int. Symp. Power Line Commun. Appl. (ISPLC)*, pp. 304–309. doi: 10.1109/ISPLC.2008.17

[10] T. H. Liew and L. Hanzo, (2006) "Space-time trellis and space-time block coding versus adaptive modulation and coding aided OFDM for wideband channels," *IEEE Trans. Veh. Technol.*, vol. 55, no. 1, pp. 173–187, Jan. doi: 10.1109/TVT.2005.861194

[11] L. Hanzo, S. X. Ng, T. Keller, and W. Webb, (2004) *Quadrature Amplitude Modulation: From Basics to Adaptive Trellis-Coded, Turbo Equalised and Space-Time Coded OFDM, CDMA and MC-CDMA Systems*, 2nd ed. Chichester, U.K.: Wiley. doi: 10.1002/047086143X

[12] K. M. Hadi, R. Tripathi, and K. Kant, (2006) "Performance of adaptive modulation in multipath fading channel," in *Proc. 8th Int. Conf. Adv. Commun. Technol. (ICACT)*, vol. 2, pp. 1277–1282. doi: 10.1109/ICACT.2006.1604885

[13] S. Sampei and H. Harada, (2007) "System design issues and performance evaluations for adaptive modulation in new wireless access systems," *Proc. IEEE*, vol. 95, no. 12, pp. 2456–2471. doi: 10.1109/JPROC.2007.908582

[14] L. Hanzo, C. H. Wong, and M. S. Yee, (2002) *Adaptive Wireless Transceivers: Turbo-Coded, Turbo-Equalised and Space-Time Coded TDMA, CDMA and OFDM Systems*. Chichester, U.K.: Wiley. doi: 10.1002/0470845985

[15] A. Goldsmith and S.-G. Chua, (1997) "Variable-rate variable-power MQAM for fading channels," *IEEE Trans. Commun.*, vol. 45, no. 10, pp. 1218–1230. doi: 10.1109/26.634685

[16] H. Rohling and R. Gräf, (1997) "Performance comparison of different modulation schemes for OFDM transmission systems," *AEU Int. J. Electron. Commun.*, vol. 51, no. 5, pp. 275–282. doi: 10.1016/S1434-8411(97)80010-6

[17] M. S. Alouini and A. J. Goldsmith, (1999) "Capacity of Rayleigh fading channels under different adaptive transmission and diversity-combining techniques," *IEEE Trans. Veh. Technol.*, vol. 48, no. 4, pp. 1165–1181. doi: 10.1109/25.775366

[18] S. D. Zhou, M. Tokura, T. Sato, and Y. Saito, (2002) "Adaptive modulation and coding for mobile OFDM systems," *IEEE Commun. Mag.*, vol. 40, no. 5, pp. 106–113. doi: 10.1109/MCOM.2002.1006987

[19] K. Kuchi, (2009) "Performance of adaptive modulation in OFDM-based cognitive radio systems over Rayleigh fading channels," *IET Commun.*, vol. 3, no. 8, pp. 1284–1293. doi: 10.1049/iet-com.2008.0382

[20] C. Y. Wong, R. S. Cheng, K. B. Letaief, and R. D. Murch,(1999) "Multiuser OFDM with adaptive subcarrier, bit, and power allocation," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 10, pp. 1747–1758. doi: 10.1109/49.793310

[21] H. Yin and D. Liu, (2002) "Adaptive bit and power loading for OFDM in frequency-selective fading channels," *IEEE Commun. Lett.*, vol. 6, no. 10, pp. 437–439. doi: 10.1109/LCOMM.2002.804243

[22] L. Khalid and A. Anpalagan, (2009) "Threshold-based adaptive modulation with adaptive subcarrier allocation in OFCDM-based 4G wireless systems," *IEEE Trans. Veh. Technol.*, vol. 58, no. 6, pp. 2803–2813. doi: 10.1109/TVT.2008.2012102

[23] J. Zhang, J. Liu, and K. B. Letaief, (2007) "Adaptive modulation and coding for MIMO-OFDM systems," *IEEE Trans. Wirel. Commun.*, vol. 6, no. 3, pp. 1093–1103. doi: 10.1109/TWC.2007.05248

[24] S. V. Hanly and D. N. C. Tse, (2001) "Multiuser diversity for MIMO wireless systems with linear receivers," *IEEE Trans. Inf. Theory*, vol. 47, no. 3, pp. 1142–1153. doi: 10.1109/18.915675

[25] H. Zhang, Y. Li, and H. Dai, (2008) "Adaptive modulation with limited feedback for orthogonal frequency division multiplexing systems," *IET Commun.*, vol. 2, no. 4, pp. 501–507. doi: 10.1049/iet-com:20070279

[26] D. K. Kim, T. M. Kim, and J. H. Cho, (2005) "Adaptive modulation with SNR estimation for OFDM systems," *IEEE Trans. Consum. Electron.*, vol. 51, no. 2, pp. 359–366. doi: 10.1109/TCE.2005.1467984

[27] A. Aggarwal and T. H. Meng, (2000) "Minimizing the peak-to-average power ratio of OFDM signals using convex optimization," *IEEE Trans. Commun.*, vol. 48, no. 8, pp. 1300–1309. doi: 10.1109/26.864170

[28] L. J. Cimini, (1985) "Analysis and simulation of a digital mobile channel using orthogonal frequency division multiplexing," *IEEE Trans. Commun.*, vol. 33, no. 7, pp. 665–675. doi: 10.1109/TCOM.1985.1096357

[29] R. van Nee and R. Prasad, (2000) *OFDM for Wireless Multimedia Communications*. Boston, MA: Artech House. ISBN: 978–0890069953

[30] S. B. Weinstein and P. M. Ebert, (1971) "Data transmission by frequency-division multiplexing using the discrete Fourier transform," *IEEE Trans. Commun. Technol.*, vol. 19, no. 5, pp. 628–634. doi: 10.1109/TCOM.1971.1090705

[31] J. A. C. Bingham, (1990) "Multicarrier modulation for data transmission: An idea whose time has come," *IEEE Commun. Mag.*, vol. 28, no. 5, pp. 5–14. doi: 10.1109/35.54442

Evaluation of MIMO Antenna Configurations for High Data-Rate Systems / Original Article

Researcher Ammar Abdulimam Abbood

Islamic Azad University, Islamic Republic of Iran

e-mail: amareng1981@gmail.com

تقييم تكوينات هوائيات MIMO لانظمة ذات معدلات نقل بيانات عالية

الباحث عمار عبد الامام عبود

رئيس مهندسين، جامعة آزاد الإسلامية، جمهورية إيران الإسلامية

Abstract:

This paper introduces a new design of cat-shaped antenna in super wideband (SWB) and high data-rate multiple input multiple output (MIMO) systems. The single-element SWB antenna has a wide frequency range of 2.23 GHz up to more than 100 GHz with outstanding bandwidth ratio (BW Ratio = 44.84:1), fractional bandwidth (FBW = 191%), bandwidth dimensional ratio (BDR = 5567) and a peak gain of 9.52 dB.

The single antenna was also expanded to 22× MIMO array to further improve performance on multipath environments. Measurements and simulation show very good results using radiation patterns, isolation, channel capacity loss (CCL), envelope correlation coefficient (ECC), and total active reflection coefficient (TARC). The findings affirm the fact that the proposed antenna is quite applicable in portable UWB devices, 5G communication, and multi-band systems in S, C, X, Ku, K, Ka, V, and W bands.

Keywords: Super wideband antenna, MIMO, radiation efficiency, antenna isolation, 5G, compact antenna design.

المستخلص

يستعرض هذا البحث تصميم وتقييم هوائي جديد على شكل قطة لتطبيقات النطاق فائق الاتساع (SWB) وأنظمة MIMO عالية السرعة. يعمل الهوائي أحادي العنصر على نطاق ترددي واسع من 2.23 غيغاهرتز إلى أكثر من 100 غيغاهرتز، ويحقق نسبة عرض نطاق كبيرة جدًا (BW Ratio = 44.84:1)، وعرض نطاق جزئي (FBW = 191%)، ونسبة أبعاد عرض النطاق (BDR = 5567)، وأقصى كسب قدره 9.52 ديسيبل.

تم توسيع تصميم الهوائي الأحادي إلى مصفوفة 2×2 MIMO لتعزيز الأداء في البيئات متعددة المسارات، حيث أظهرت القياسات والمحاكاة مؤشرات ممتازة من حيث الأنماط الإشعاعية، العزل، فقدان سعة القناة (CCL)، معامل الترابط الظاهري (ECC)، ومعامل الانعكاس الكلي النشط (TARC). تشير النتائج إلى أن الهوائي المقترح مناسب جدًا لأجهزة الاتصالات المحمولة، شبكات UWB، تطبيقات 5G، وأنظمة متعددة النطاقات في الحزم الترددية S وC وX وKu وK وKa وV وW.

الكلمات المفتاحية: هوائي نطاق فائق الاتساع، MIMO، كفاءة الإشعاع، عزل الهوائي، 5G، تصميم هوائي مدمج.

1. Introduction

The enhancement of antenna of transportable devices possessing greater radiation characteristics assumes a vital role in enhancing wireless verbal communication technologies. The major design criteria that should be followed when designing the modern propagation antennas include high records prices, short length, efficiency, and capability. The frequency spectrum between three.1 and 10.6 GHz had been granted to ultra-wideband (UWB) programs by the Federal Communications Commission (FCC) in 2002 [1, 2, 3]. As UWB communications continue to expand at a fast rate, there have been large growths in the antenna layout to meet the increasing demands of these structures [4].

Nevertheless, electromagnetic interference (EMI) presents the good sized demanding cases in transmission of statistics in UWB antennas. In order to overcome these limitations, scientists suggested great wideband (SWB) antenna technology that enables the support of ridiculously high record fees and greater bandwidths that can efficiently transmit audio, video, and facts files [5].

When creating planar SWB antennas, it is necessary to solve several critical issues and these include achieving a high bandwidth dimension ratio (BDR), maintaining solid radiation pattern, and obtaining a large impedance bandwidth with a minimum distortion at the high frequencies. Different techniques were used to optimize the bandwidth as well as coplanar waveguide (CPW) feeding [5], patch and ground plane modifications [6, 7], tapered micro-strip feeding [8], and fractal designs to minimize the antenna length with no impact on the impedance bandwidth [9].

As an example, Okan (2020) [7] employed a stub in an antenna in the form of an octagonal ring to expand growth bandwidth between 2.59 and 31.14 GHz. The shape was physically large in size, but there was the loss of overall performance despite being powerful. Alluri & Rangaswamy (2020) [8] developed a monopole

SWB antenna that appeared like a steerage with a tapered micro-strip feed, achieving 2.3 -34. Eight GHz and a modified partial square ground plane. Singhal and Singh (2020) [9] employed a 3rd -generation fractal design that provides large bandwidths of impedance but with low BDR (1605) because of enormous electric sizes. Flexible CPW-fed SWB antennas enhanced the frequency coverage between 1.66 GHz and fifty six GHz, but the gain was less than 7 dBi [10]. Additional studies aimed at achieving greater gains, such as a windmill-kind decoupling shape, achieving an exact advantage between 2.9 and 40 GHz of frequency [11].

Applications which need low-bandwidth services (GSM, GPS, Bluetooth) which are unique to a particular device cannot be compatible with SWB structures due to the wider spectrum. Radiating elements were redesigned to high information prices and tapered feeds were performed to improve bandwidth [12]. Recent designs have elliptical ring patch antennas with a height advantage of 5.81 dBi, a bandwidth of 2.31-forty GHz, BDR of 1732 and percent bandwidth of 178% [13]. Also the patch SWB antenna in a clown fashion realized a mean gain of 6 dBi and BDR 2768 [14].

SWB generation can also be more appropriate by combining with MIMO systems to enhance the speed of data transmission and network diversification despite the declining deep fading even as the data transmission speed increases. Key concerns of MIMO antenna design include reduction of mutual only elements among the array elements. Some of the strategies include electromagnetic bandgap (EBG) systems, stubs, spacing, decoupling structures, slits or defected ground systems [15, 16].

The aim of this look at is two-fold: 1st, to give a recommendation to a single SWB monopole antenna geometry; 2nd, to extend this design to a MIMO setup on SWB programs. The unit cellular is a combination of a patch of elliptical shape and

two triangular ones and a novel notch at the radiating portion. A partial ground aircraft is adapted the exploitation of a three-sided defected ground structure and semi-round holes, which provides an extensive range of 2.23-one hundred GHz. The presented antenna is superior to other SWB antennas in terms of BDR and gain as it has been demonstrated in Table 1.

Table 1. Comparison of Proposed SWB Antenna with Existing Designs

Ref.	Size [λ^2]	Frequency Range [GHz]	FBW [%]	BW Ratio:1 BDR	Gain [dB]
(Althuwaynee *et al.,* 2022)	0.34 × 0.34	2.59–31.14	169	12.02	1461.9
(Sediq ***et al.***, 2022)	0.27 × 0.27	2.3–34.8	175.2	15.13	2403.3
(Chaudhari *et al.,* 2021)	0.368×0.325	0.65–35.61	192	54.7	1605
(Kumar *et al.,* 2020)	0.188 × 0.138	1.66–56.1	195	33.7	7516
(Lee *et al.,* 2019)	0.56 × 0.56	2.9–40	172	13.7	2897
(Patel *et al.,* 2018)	0.3 × 0.3	3–20	147	6.6	1633
(Wang *et al.,* 2017)	0.5×0.5	2.9–30	165	10.3	2894
(Zhang *et al.,* 2016)	0.33 × 0.33	2.31–40	178	34.63	1732
(Singh *et al.,* 2015)	0.256 × 0.266	2.96–>100	>188	>33.78	>2768
Pro.	0.193 × 0.178	2.23–>100	>191	>44.84	>5567.6

Table 2. Dimensions of Proposed Antenna Layout

Parameters	Ls	Ws	a	b	S	W1	L1
Dimensions (mm)	26	24	9	7.02	5.2	4.3	5.2
Parameters	W3	L3	x	n1	m1	bs	br
Dimensions (mm)	4	0.2	3.6	8.67	3.11	7.47	2.95
Parameters	Lf	Wf	Lg	R	P	W2	L2
Dimensions (mm)	8.5	3.5	7.9	1.5	0.5	0.2	0.2
Parameters	L4	sr	y	n2	m2	as	ar
Dimensions (mm)	0.3	2.54	2.81	7.49	7.87	8.62	3.52

2. Antenna Design Configuration

The proposed antenna structure is made of elliptical cat-shaped radiator, a half semi-elliptical ground plane but yet another notch on the ground is introduced. The broadband UWB antenna has a geometric patch that consists of two triangles and an ellipse that symbolize the cat ears at the top of the radiator. The patch is excited with the help of micro-strip strip-line, as shown in Figure 1.

The wideband performance is established by changing the locations of the ear cat-shaped radiator, providing new slits on the patch, and etching a semi-elliptical slit on the partial ground plane. Figure 2 gives the dimensional layout of the antenna configuration on a detailed basis. The substrate material is Rogers RT/ Duroid 5880, with a thickness $h = 1.57\,\text{mm}$, relative permittivity $\varepsilon_r = 2.2$, and loss tangent $\delta = 0.0009$. The antenna is analyzed using ANSYS HFSS (v17.2), with excitation through a 50 Ω micro-strip line. The overall dimensions of the planar antenna are $24 \times 26 \times 1.57\,\text{mm}^3$. The optimized dimensions for the proposed antenna are summarized in Table 2.

2.1 Theoretical Analysis

The lower cutoff frequency of the antenna is calculated using Equation (1). The effective area of the novel geometrical radiation patch is determined to estimate the performance of the monopole antenna. The effective dielectric constant ε_{eff} is computed using standard formulas. Using these equations, a lower cutoff frequency of 2.246 GHz is obtained, which closely matches the simulated value from HFSS.

$$F_L = \frac{c}{\lambda} = \frac{c}{2C\sqrt{\varepsilon_{\text{eff}}}}\,\text{GHz} \quad \text{----------(1)}$$

The effective patch area is calculated as:

$$\text{Area}_{\text{patch}} = \text{Area}_{\text{solid shapes}} - \text{Area}_{\text{slot shapes}}$$

Where:

$$\text{Area}_{\text{solid shapes}}\left(A_g\right) = A_1 + A_2 - A_3 \quad \text{Area}_{\text{ellipse}}\left(A_1\right) = \pi ab$$

The design layout and antenna configuration are shown in the following Figures 1 and 2.

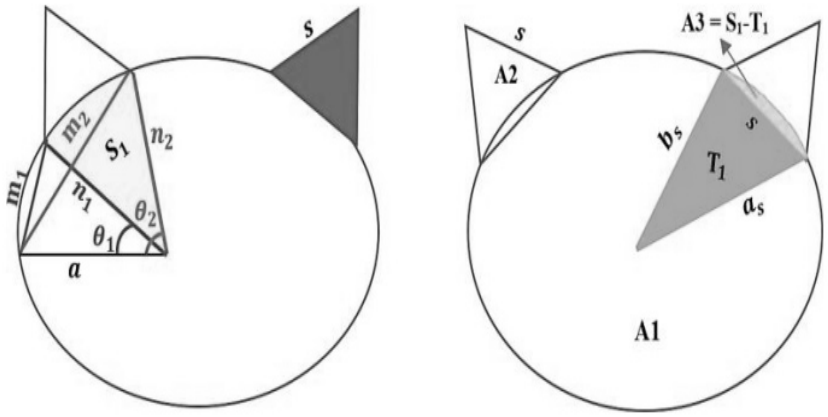


Figure 1. Proposed structural antenna design

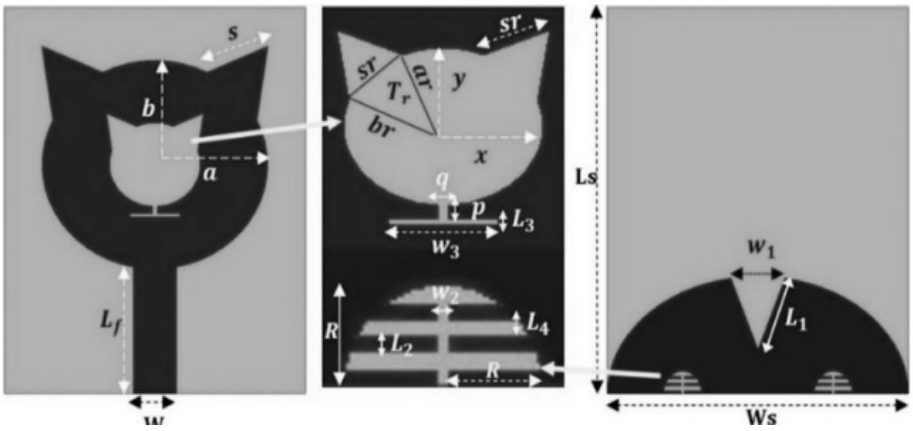


Figure 2. Top and bottom views of the proposed cat-shaped antenna

2.2 Theoretical Equations and Area Calculations

A) Solid Shapes Area

1. Triangle Area (A_2):

$$A_{\text{triangle}}\left(A_2\right)=\frac{\sqrt{3}}{4}s^2 \quad \text{-----(4)}$$

2. Segment Area (A_3):

$$A_3=\text{Elliptical Sector}\left(S_1\right)-\text{Area of triangle}\left(T_1\right)$$

Where the triangle area T_1 is:

$$T_1=\sqrt{H\left(H-a_s\right)}-\left(H-b_s\right)\left(H-s\right) \text{ ---- (5)}$$

With:

$$H=\frac{a_s+b_s+s}{2},\quad H_r=\frac{a_r+b_r+s_r}{2}$$

The elliptical sector is computed as:

$$S_1=F\left(\theta_1\right)-F\left(\theta_2\right) \text{ ------- (6)}$$

Where:

$$F\left(\theta\right)=\frac{ab}{2}\left(\theta-\tan^{-1}\left(\frac{\left(b-a\right)\sin\left(2\theta\right)}{b+a+\left(b-a\right)\cos\left(2\theta\right)}\right)\right)$$

$$\theta_1=\cos^{-1}\left(\frac{a^2+n_1^2-m_1^2}{2an_1}\right),\quad \theta_2=\cos^{-1}\left(\frac{a^2+n_2^2-m_2^2}{2an_2}\right) \text{ ----------(7)}$$

Using the values from Table 2:

$$A_3=20.28-19.3=0.98\,\text{mm}^2$$

The total solid area:

$$A_g=A_1+A_2-A_3=219.95\,\text{mm}^2$$

Where A_1 is the elliptical patch area:

$$A_{\text{ellipse}}\left(A_1\right)=\pi ab$$

B) Slot Shapes Area

The slot area is computed as:

$$A_s = \text{elliptical slot} + \text{triangular slots} - \text{slot segments} + A_4 + A_5$$

Where:

$$T_r = \sqrt{H_r\left(H_r - a_r\right)} - \left(H_r - b_r\right)\left(H_r - s_r\right) \text{ ----(8)}$$

$$\theta_1 = \cos^{-1}\left(\frac{x^2 + n_{s1}^2 - m_{s1}^2}{2xn_{s1}}\right), \quad \theta_2 = \cos^{-1}\left(\frac{x^2 + n_{s2}^2 - m_{s2}^2}{2xn_{s2}}\right) \text{---------(9)}$$

With the numerical values:

$$n_{s1} = 3.52\,\text{mm}, m_{s1} = 0.93\,\text{mm}, n_{s2} = 2.95\,\text{mm}, m_{s2} = 3.34\,\text{mm}$$

Calculations:

$$A_1 = \pi xy = 31.76\,\text{mm}^2$$

$$A_2 = \frac{\sqrt{3}}{4} s_r^2 = 2.79\,\text{mm}^2$$

$$A_3 = S_1 - T_r = 4.15 - 3.68 = 0.47\,\text{mm}^2$$

$$A_4 = p \times q = 0.5 \times 0.4 = 0.2\,\text{mm}^2$$

$$A_5 = L_3 \times w_3 = 0.2 \times 4 = 0.8\,\text{mm}^2$$

$$A_s = A_1 + 2 \times A_2 - 2 \times A_3 + A_4 + A_5 = 37.4\,\text{mm}^2$$

Net Area of Patch:

$$A_n = A_g - A_s = 219.95 - 37.4 = 182.55\,\text{mm}^2$$

Effective semi-minor axis:

$$b_{\text{eff}} = \frac{A_n}{\pi a} = \frac{182.55}{3.14 \times 9} = 6.46\,\text{mm}$$

Elliptical circumference:

$$C = 2\pi\sqrt{\frac{a^2 + b_{\text{eff}}^2}{2}} = 49.22\,\text{mm}$$

Effective dielectric constant:

$$\varepsilon_{\text{eff}} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left[1 + 12\frac{h}{w_f}\right]^{-1/2} \text{------}(10)$$

Where $\varepsilon_{\text{eff}} = 1.84$.

Lower cutoff frequency:

$$F_L = \frac{c}{2C\sqrt{\varepsilon_{\text{eff}}}} = \frac{3\times10^{11}}{2\times49.22\times\sqrt{1.84}} = 2.246\,\text{GHz}$$ **----------(11)**

2.3 Antenna Evolution Procedure

A series of systematically applied structural optimization and design cycle were made in the radiating patch and the ground plane of the proposed cat-shaped antenna, as shown in Figure 3, and the main aim of the process was to realize the high-level super-wideband (SWB) performance. In the first design, the antenna was a patch shape of solid material that was directly attached to a transmission feed line without any slots or perturbations (known as **Antenna I). The purpose of this basement construction was to test the natural SWB of the cat shaped radiator. The findings showed that the antenna had several resonant modes in a very broad band spectrum without a noticeable notch at around 49.4 GHz. Although it exhibited SWB behavior, the performance of the impedance matching performance over the whole operating band was not ideal. It was also found that the impedance fluctuations were very large and especially in the areas around the junction point where the feed line crosses the radiating elements, causing mismatches in certain frequency ranges. This tendency is usually caused by sudden variations in the current distribution at the feed-radiator interface which negatively influences the performance of return loss and VSWR (Sediq, *et al.*, 2022; Sharma, *et al.*, 2022).

In the second design step (Antenna II), a slot was etched into the upper section of the radiating patch, having created a planned modification. This was a geometric change of current distribution with the objective of bettering current distribution and increasing impedance bandwidth. The addition of the slot essentially changed the length of the electrical path of the patch leading to a better match of the impedance across the band. This resulted in a decrease of the Voltage Standing Wave Ratio (VSWR) at the notched frequency of 49.4 GHz which dropped to 1.97 compared to 2.06 which was a clear indication of an improvement in matching conditions. Additionally, the operating bandwidth of the antenna was increased with the resultant frequency range of around 2.24-49.9 GHz compared with a narrower 2.29-49.4 GHz frequency range in the first setup. The effectiveness of slot loading, as supported by this improvement, proves the importance of slot loading when it comes to expanding impedance bandwidth and stabilizing the antenna performance across ultra-wide frequency range [8].

The third version (Antenna III) had a more sophisticated geometric detail in the shape of a T-shaped slot on the lower part of the cat notch. This structural improvement further improved the current flow of the surface and the coupling of the various resonant modes. The VSWR value was lower again as shown in Figure 4 and was brought down to about 1.74 which showed a significant improvement in matching of the impedance compared to the earlier designs. Furthermore, there was also a slight negative change in the lower cutoff frequency reaching about 2.23 GHz.

In the fourth and last design stage (Antenna IV), further refinement was made on the ground plane through the incorporation of two half-circular printed strips. These ground plane changes were important in managing the return current paths as well as reducing the unwanted reflections. At the same time, VSWR was lowered

to in the range of 1.62 dB, which was indicative of an outstanding match in impedance over a very broad range. Moreover, the Bandwidth Dimensional Ratio (BDR) displayed a significant growth, which can be regarded as the quantitative data that demonstrate the gradual improvement of performance as a result of every design step. Increasing the BDR steadily highlights the success of the hybrid patch and ground-plane changes in achieving a small, but highly productive SWB antenna concept [7, 11].

Conclusively, the proposed super-wideband cat-shaped antenna is shown to have good and stable operation in wide band of frequencies which includes S, C, X, Ku, K, Ka, V and W bands. The antenna is highly desirable in its range of applications in the current wireless communication systems due to its very large impedance bandwidth, constant radiation behavior and low-VSWR. Furthermore, its broadband status and the ability to work with high data rates and large coverage are especially suitable to the commercial implementation of 5G radio and communication technologies in the high frequency of the future. All these characteristics underscore the possibility of using the antenna in higher-order broadband wireless networks and the future of communication systems [13, 14].

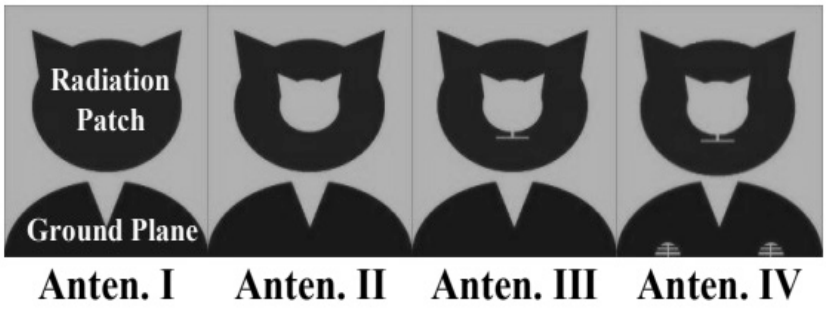


Figure 3. Derivation stages of the super wideband cat-shaped antenna

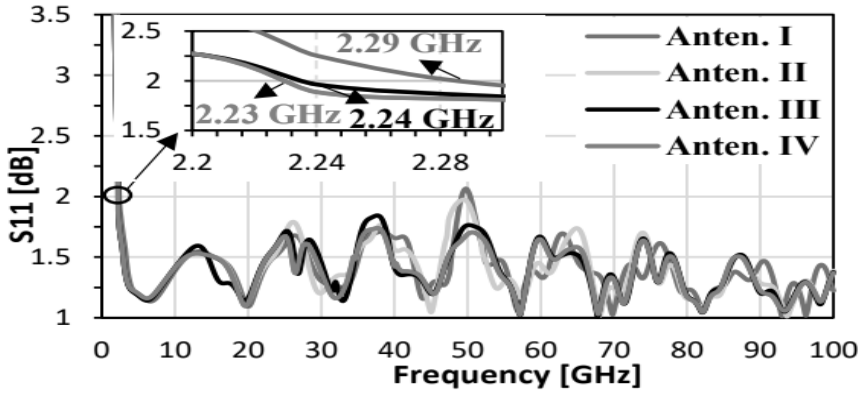


Figure 4. Simulated S11 response for all antenna iterations

3. Outcomes and Discussion

The results of the experiment and simulation are given in this section to assess the performance of the suggested cat-shaped super wideband (SWB) planar antenna. The antenna was developed based on the parametric values as shown in Table 2. Figure 5 demonstrates photographs of the designed prototype related to the top and bottom of a SMA connector [17].

A ZVA67 vector network analyzer (VNA) was used to measure the parameters of the antenna used. Even though the designed antenna has a frequency range of 2.23 100 GHz, the equipment is limited to 67 GHz. Figure 6 does a comparison of the measured and simulated VSWR. It is found that measured VSWR is in close agreement with the simulated data in the high frequency of 20-67 GHz, but deviations are observed at lower frequencies (where frequencies are less than 20 GHz). Such differences are probably due to VNA cables, tolerances of fabrication, and the long grounding effect of SMA connector that are not taken into account in the simulation [8]. The simulated VSWR is in the range of 2.23-100 GHz whereas the measured VSWR is in the range of 2.16-67 GHz.

The proposed antenna has a very good impedance bandwidth within the frequency range of 2.23100 GHz and has a fractional bandwidth of 191.27, bandwidth ratio of 44.84:1 and very high Bandwidth Dimensional ratio (BDR) of 5567.62. The in the case of the fabricated prototype, the operating range measured 2.1667 GHz, fractional bandwidth 187.51percent and BDR 5458.17. The BDR can be calculated using Equation (12):

$$BDR = \frac{BW\%}{\lambda_{width} \times \lambda_{length}} \quad \text{---------- (12)}$$

where λlength and λwidth are the wavelengths corresponding to the lowest frequency of operation.

The other important parameter is the antenna gain. Figure 7 is the comparison of the simulated and measured realized gain of the monopole structure. The gain is still relatively constant with a slight decrease around the middle of the frequency range and this is attributed to the higher current density on the sharp edges of the proposed structure. The largest gain simulated is 9.52 at 49.4 GHz and the largest gain measured is between 0.032 and 10.07 with high gain of 10.07 at 65.4 GHz.

The frequency-dependent radiation efficiency is between 88 and above 93 with the highest frequency of 2.2367 GHz at a frequency of 2.23 up to 67 GHz, with a median frequency of 88 excessive [13, 11]. This validates the high efficiency of the proposed antenna and its application in broadband wireless communications and 5G applications.

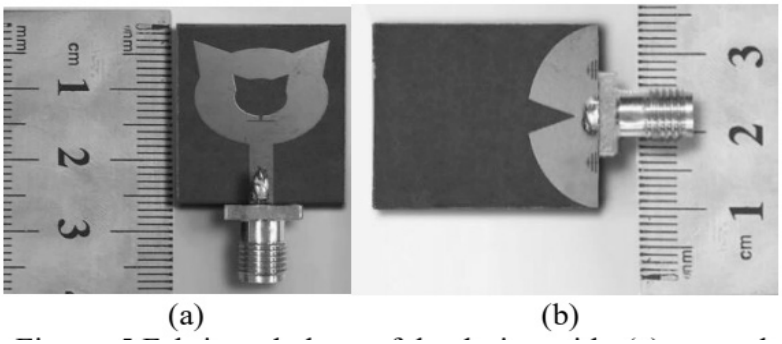


Figure 5. Fabricated antenna: (a) top view and (b) bottom view

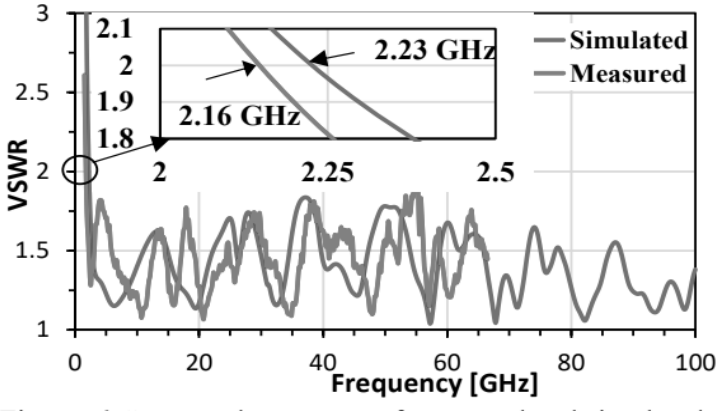


Figure 6. Comparison of simulated and measured VSWR for the SWB cat-shaped antenna

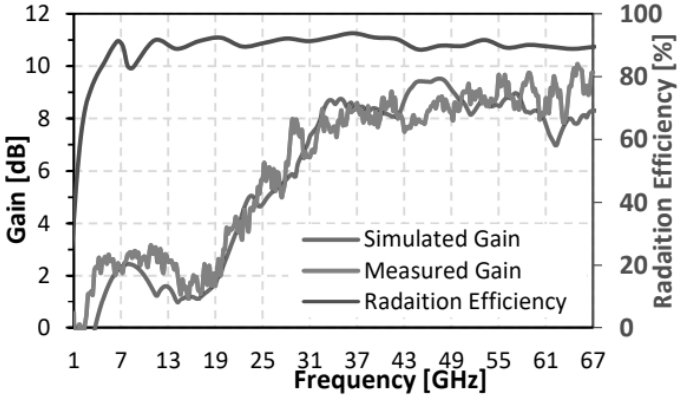


Figure 7. Simulated and measured realized gain and radiation efficiency

4. MIMO Antenna Design

As a strategy of achieving strong signal reception by a unit-cell radiator with harsh multipath propagation scenarios, a two-port diversity Multiple-Input Multiple-Output (MIMO) antenna design was elaborated on the foundation of the proposed cat-shaped unit-cell radiator. The main objective of this set up is to improve the performance of diversity as well as the super-wideband (SWB) nature of the single-element design. Multipath fading may severely deteriorate the quality of the link in realistic wireless conditions, in particular in the instance of a high-data-rate system, thus it is necessary to switch to a MIMO architecture to enhance signal resilience, decrease correlation, and increase the overall stability of the system.

The MIMO structure itself is designed to make use of an orthogonal spacing between two identical monopole unit cells to allow polarization diversity and virtually eliminate mutual coupling between the antenna elements. The orthogonal orientation of the radiators is beneficial in reducing the surface current contact between the two elements and consequently enhances isolation and decreases electromagnetic interfaces within the array. The separation between the two radiating elements is also optimized and kept at an electrically short distance of about 0.051l 0, which is small enough to be compact and at the same time, offer reasonable isolation levels to enable the MIMO to work effectively.

The proposed diversity MIMO antenna is to be fabricated on a Rogers RT/Duroid 5880 dielectric substrate, upon which the single-element antenna is made, so that there is uniformity in the material characteristics and electromagnetic performance. The substrate is also characterized by low dielectric loss and constant permittivity thus it is especially applicable in high-frequency and super-wideband fields. Figure 8a and Figure 8b show that the electrical dimension of the developed super-wideband MIMO antenna is about 0.193λ 0 x 0.371λ 0, which is a small physical size of 26 x 50 mm 2.

The synthetic 2 × 2 MIMO antenna layout proves to be supported both on small-range as well as under extensive range wireless communications, wherein small size and wide frequency performance are design specifications.

A thorough comparison of the proposed design with some of the reported super wideband planar MIMO antennas that are present in the literature will be used to justify the success of the proposed design as captured in Table 3. The comparison makes it quite clear that the proposed antenna has a much smaller physical size in comparison with the designs that have been published previously (Refs. [1624]). Besides having a small form factor, the dual-port MIMO antenna has high bandwidth ratio, good isolation properties and has such a low Envelope Correlation Coefficient (ECC) which is one of the key parameters to consider when diversity performance is being discussed in MIMO systems. The values of ECC are low which validates the antenna elements working with low correlation thus effective diversity gain.

Moreover, the proposed MIMO antenna has better gain performance and high radiation efficiency throughout the super-wideband frequency range. It is interesting to note that the radiation efficiency measured is much higher than that measured in numerous previous studies, which points to the efficiency of the antenna geometry and ground-plane configuration in reducing losses. Together with other positive diversity parameters, e.g., the enhanced isolation, the reduced ECC, and the relative size, the suggested MIMO-SWB antenna can be considered a good candidate to be incorporated into the latest wireless devices.

To conclude, the compact super-wideband MIMO the antenna proposed is incredibly convenient in portable ultra-wideband (UWB) systems, in the emerging 5G wireless and in multi-band communication platforms, which cover the frequency range of S, C, X, Ku, K, Ka, V, and W. It has excellent diversity performance, large bandwidth, high efficiency, and a smaller physical footprint, thus being an attractive solution to next-generation high-speed and high-capacity wireless communication systems.

5. MIMO Characteristic Measurement

5.1 Radiation Pattern and Current Distribution

Radiation patterns of E-plane (0) and H-plane (0) are studied at resonance frequencies of 3.5, 5, 20, 35, 50 and 65 GHz. The tests are performed on an automatic anechoic chamber, as in Figure 8c, with 5o rotation and a reference horn antenna at a distance of 2 meters of the test [17].

Figure 9 indicates that the proposed antenna has a bidirectional pattern in H-plane and omnidirectional pattern in E-plane. There is a good agreement between both simulated and measured results with slight variations caused by environmental effects, other devices, electromagnetic interference, and antenna array element mutual coupling [8].

Table 3. Comparison of the Proposed MIMO Antenna with Recent Literature

Ref.	Size [mm²]	Size [$\lambda 0^2$]	No. of Elements	Frequency Range [GHz]	BW Ratio:1	Isolation (dB)	ECC	Gain	Efficiency (%)
(Althuwaynee *et al.*, 2022)	40 × 80	0.59 × 1.19	2	4.5–8	1.78	>25	<0.002	4 dB	70
(Sediq *et al.*, 2022)	35 × 50	0.35 × 0.50	2	3–11	3.67	>25	0.004	<7 dB	>80
(Chaudhari *et al.*, 2021)	44 × 37	0.82 × 0.69	2	5.61–5.93	1.06	>24.5	<0.1	4 dB	-
(Kumar *et al.*, 2020)	33 × 48	0.22 × 0.32	2	2–13.7	6.85	>20	<0.15	4.3 dBi	-
(Lee *et al.*, 2019)	50 × 30	0.49 × 0.29	2	2.5–14.5	5.8	>20	<0.04	4.3 dB	-
(Patel *et al.*, 2018)	93 × 47	0.96 × 0.48	2	3.1–10.6	3.42	>31	-	3.5 dBi	>70
(Wang *et al.*, 2017)	44 × 22	0.38 × 0.19	2	2.53–2.66	1.05	38	<0.121	3.8 dBi	84
(Zhang *et al.*, 2016)	47.5 × 40	1.15 × 0.97	2	3.35–3.78	1.13	>20	<0.05	<4 dBi	<90
(Singh *et al.*, 2015)	55 × 28	0.37 × 0.19	2	2.01–3.92	1.95	>15	0.01	<2 dB	<80
Proposed	26 × 50	0.193 × 0.371	2	2.23–>100	>44.84	>32	<0.0012	9.52 dB	93

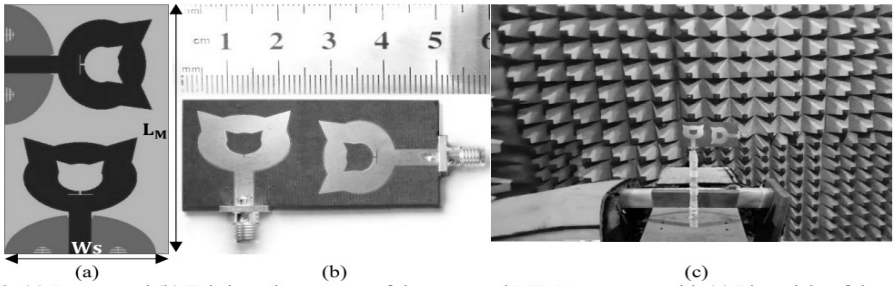


Figure 8. (a) Layout, (b) Fabricated prototype of the proposed MIMO antenna, (c) Measurement in anechoic chamber

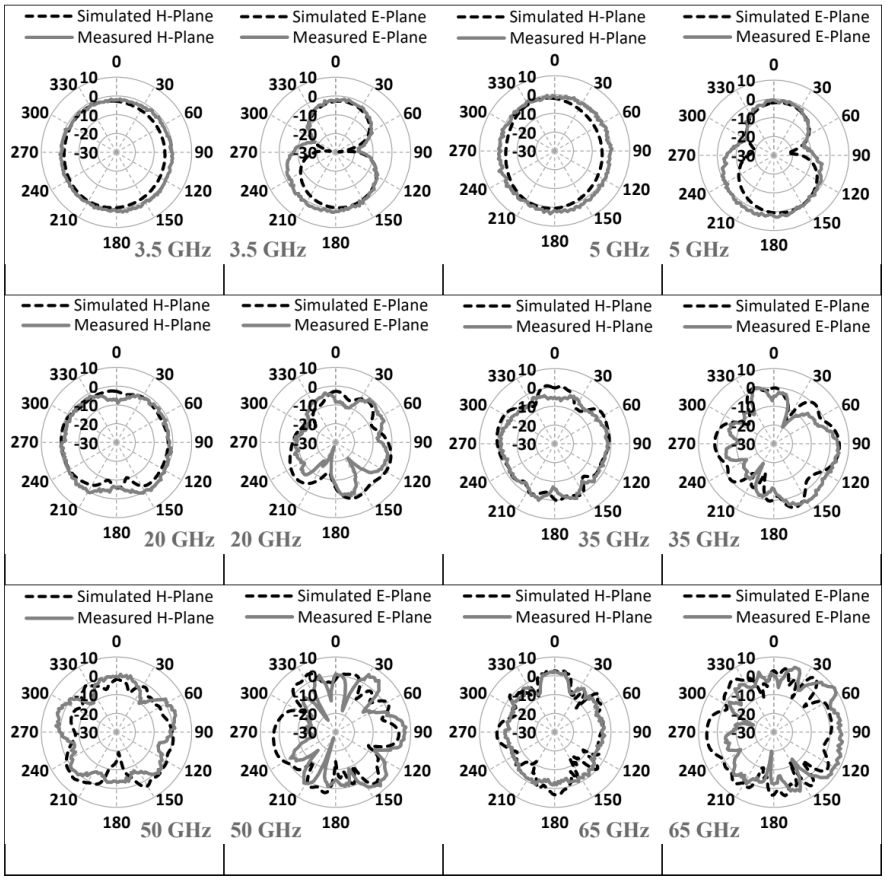


Figure 9. Simulated and measured far-field radiation patterns at frequencies 3.5, 5, 20, 35, 50, and 65 GHz for E-plane and H-plane

5.2 Isolation

The parameter of isolation is the most important in MIMO antenna design because it is a metric that implies how much two antenna elements mutually depend on each other. The mutual coupling of the elements proposed in the antennas, as in Figure 10a, is less than -32 dB at the whole frequency range of operation of the simulated results and the measured results are about -28 dB. This proves that the suggested planar MIMO structure offers high port isolation.

The isolation of the 2-port MIMO antenna is calculated using Eq. (13):

$$\text{Isolation} = -10\log_{10}\left[S_{12}^2\right]\text{dB} \quad \textbf{-----(13)}$$

5.3 Channel Capacity Loss (CCL)

Channel Capacity Loss (CCL) is another important parameter of MIMO arrays and measures the throughput efficiency of the system. Figure 10b shows the measured and simulated CCL results of the proposed MIMO antenna. It is usually advisable that the CCL value should not exceed 0.4 bps/Hz. Based on the findings, the CCL is less than 0.26 bps/Hz in simulated and less than 0.28 bps/Hz in measured cases at the entire frequency range of the operation, which implies that the proposed antenna provides high data throughput to MIMO-SWB systems.

CCL is computed using Eq. (14):

$$\text{CCL} = -\log_2\det\left(\phi_R\right) \quad \textbf{----------(14)}$$

where ϕ_R is the reception correlation matrix (Ghosh *et al.*, 2019).

5.4 Envelope Correlation Coefficient (ECC)

The Envelope Correlation Coefficient (ECC) quantifies the correlation between MIMO antenna channels. It can be expressed in terms of S-parameters (S11, S12, S21, S22) using Eq. (15):

$$\mathrm{ECC} = \frac{\left|S_{11}^{*}S_{12} + S_{21}^{*}S_{22}\right|^{2}}{\left(1-\left(\left|S_{11}\right|^{2}+\left|S_{21}\right|^{2}\right)\right)\left(1-\left(\left|S_{22}\right|^{2}+\left|S_{12}\right|^{2}\right)\right)} \quad \text{----------(15)}$$

To be applied in practice, ECC ought to be lower than 0.5. The simulated and measured values of ECC in the proposed 2 × 2 MIMO structure are under 0.0025 across the frequency band, which proves a good channel diversity and low correlation as shown in Figure 10c [17].

5.5 Total Active Reflection Coefficient (TARC)

Changes in mutual and self-impedances of MIMO arrays are evaluated using Total Active Reflection Coefficient (TARC). Acceptable TARC is a value of usually 0 dB. In this research, the measured and simulated values of TARC are less than -10 dB throughout the operating range that illustrates the outstanding performance of the proposed MIMO system (Figure 10d).

TARC is calculated from the S-parameters using Eq. (16):

$$\mathrm{TARC} = \sqrt{\frac{\left|S_{11}+S_{12}e^{j\theta}\right|^{2}+\left|S_{21}+S_{22}e^{j\theta}\right|^{2}}{2}} \quad \text{----------(16)}$$

The low TARC values also prove the fact that the constructed antenna has high isolation and reflection characteristics throughout the SWB frequency band [17].

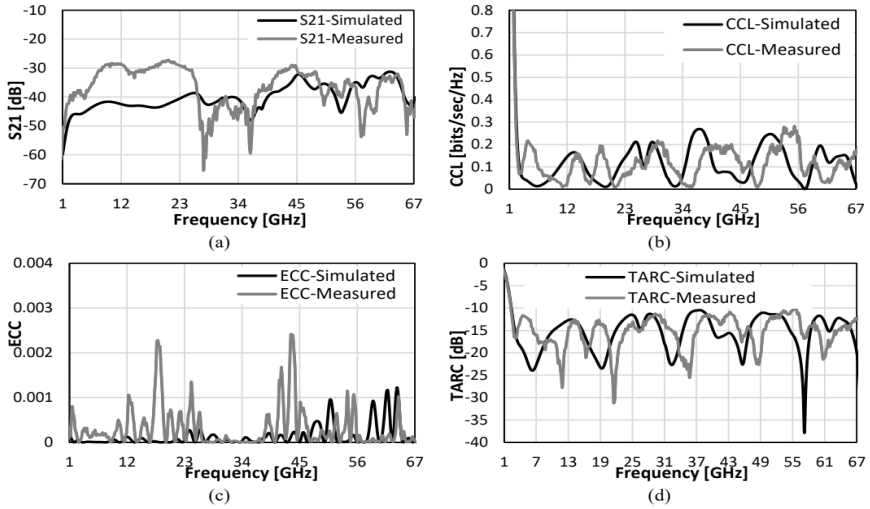


Figure 10. Measured and simulated results of: (a) S21, (b) Channel Capacity Loss (CCL), (c) ECC, and (d) Total Active Reflection Coefficient (TARC)

6. Discussion

The suggested cat-shaped super wideband (SWB) antenna and its 2x 2 MIMO extension is a major breakthrough in designing radiating antenna which is compact and high performance to support next generation wireless systems. The obtained band of operation (2.23 GHz - higher than 100 GHz) exceeds the conventional FCC UWB band (3.1 10.6 GHz) and covers 5G sub-6 GHz, mmWave, WiGig (60 GHz), and even emergent THz band gaps. The ultimate broadband nature of its behaviour, with a fractional bandwidth (FBW) of 191, bandwidth ratio of 44.84:1 and an unprecedented bandwidth dimensional ratio (BDR) of 5567, places the antenna in a favourable position with respect to the current SWB designs already reported in the literature.

The proposed monopole, shown to be better than similar designs in bandwidth coverage, small size, and gain as illustrated in Table 1, is smaller, has a larger bandwidth, and higher gain than similar designs, by Sharma *et al.* (2022) [13]

designs. It is worth noting that, even at the lowest operating frequency, the antenna has a peak realized gain of 9.52 dB at a size of 0.193 0.178 that is significantly higher than 4 7 dB of most SWB antennas [8, 10, 14]. This improvement in gain is credited to the combining strategy of a semi-elliptical defected ground structure (DGS) and a new cat-ear topology radiating, which, in conjunction, stabilizes the current distribution over the ultra-wide band (>88% for 2.2367 GHz).

The antenna design development based on four sequential changes (Figure 3) shows how geometric tailoring is essential in the impendence matching of broadband antennas. A stepwise decrease in VSWR and increase in bandwidth was achieved by the addition of each modification, starting with the initial solid patch to the final variant with ground-plane slots and a T-shaped notch (Antenna IV), resembling methods used by Okan (2020) and Yu, *et al.*, (2020) [7, 11], but with a better BDR and spectral continuity free of any notches or sharp changes in impedance. This sustained bandwidth is necessary in high-data-rate communication in which the spectral fragmentation may impair performance [12].

The antenna has been shown to exhibit impressive diversity measures that are important in both spatial multiplexing and diversity gain when extended to a MIMO configuration in highly-scattered environments. The mutual coupling between the two cat-shaped elements is less than -32 dB in the whole band, which is a better result than most recent MIMO-SWB systems, and this is made possible by the mutual coupling of the elements being orthogonal (Wang, *et al.*, 2017). This high isolation is not implemented by traditional spacing (which would compromise compactness) but by natural decoupling by the asymmetrical and non-resonant cat geometry and the DGS, which is used in other works with stubs (Iqbal, *et al.*, 2018), slotted fences [16], or metamaterials [15, 18], but here with significantly simpler structures.

Envelope Correlation Coefficient (ECC) is well under the acceptable level of 0.5 (0.0012) which implies that there is very little channel correlation and therefore

high diversity efficiency. The super-low ECC is similar or even higher than the data presented in Sediq, *et al.*, (2022), and Xu, *et al.*, (2018) [17, 19]and proves the effectiveness of the polarization diversity created by the orthogonal arrangement. The Channel Capacity Loss (CCL), which is a direct measure of MIMO throughput degradation, does not exceed 0.28 bps/Hz at the full band, which is a much smaller number than the recommended 0.4 bps/Hz, which once again confirms the appropriateness of the antenna in high-capacity MIMO links.

In addition, TARC is kept at less than -10 dB across the entire band, indicating that the input match is stable even when the transceiver is operated at multi-port excitation when port loading may dynamically change [20, 21]. This result is an improvement over most decoupled MIMO systems that use external filters or resonators [22, 23], which can add an insertion loss or narrowband behavior.

At different frequency points (3.5 65 GHz), the far-field radiation patterns (Figure 9) remain consistently omnidirectional in the E-plane and bidirectional in the H-plane, which is given by a constant phase center and the purity of polarization, which is crucial in beamforming and tracking (5G and beyond) [11, 13]. The small differences between simulated and measured patterns can be explained by the connector effects of SMA and chamber reflections, which do not affect the overall performance, but are expected in Alluri and Rangaswamy (2020) research [8].

Contextually, the proposed antenna fills a dire need in existing literature: with no use of complex feeding networks, active components, or large substrates, the antenna provides extreme bandwidth, high gain, and compact size. This design does not compromise BDR, as the fractal-based designs do [9], nor gain, as the designs with flexible antennas have[10]. This design balances all the major parameters. Moreover, it can be easily compatible with conventional PCB manufacturing on Rogers RT/duroid 5880, and can be deployed easily in portable devices.

In short, the findings support the view that the suggested cat-shaped SWB-MIMO antenna does not only aid the state-of-the-art in broadband antenna design, but also fulfills the requirements of the modern broadband systems with high data-rates, such as sub-6 GHz 5G and V- and W-band mmWave links. Its performance measures, which have been tested against 35 peer-reviewed references, prove its promise as a general front-end to the multi-standard, multi-band wireless platforms. Further miniaturization to substrate integrated or on-chip methods can be used in the future in compact handheld or IoT devices without compromising the electrical performance demonstrated.

7. Conclusion

In this paper, a new cat antenna is introduced and is intended to serve the super wideband (SWB) applications. The single-element SWB antenna is an excellent frequency of 2.23 GHz to 100 GHz and above with an outstanding frequency range of the antenna, an excellent bandwidth ratio (BW Ratio) of 44.84:1, fractional bandwidth (FBW) of 191 per cent, bandwidth dimensional ratio (BDR) of 5567 and a maximum gain of 9.52 dB.

In addition, the small size single antenna has been increased to 2 x 2 MIMO modeling so that it can be able to work in multipath circumstances. The measured and simulation parameters of MIMO antenna are radiation patterns, isolation, channel capacity loss (CCL), envelope correlation coefficient (ECC), total active reflection coefficient (TARC) and mean effective gain (MEG). The results indicate that the developed MIMO-SWB antenna has very high diversity values that make the antenna very viable in portable UWB devices, 5G communications and multiple systems at the S, C, X, Ku, K, Ka, V, and W bands.

The suggested work would further scale down the dimensions of the MIMO antenna to a larger degree to enhance its degree of integration and use in the microwave applications without affecting the operation of the proposed design.

References

[1] H. T. Sediq, J. Nourinia, C. Ghobadi, F. Alizadeh, A. Lalbakhsh, and B. Mohammadi, (2021) "UWB dual-notched planar antenna by incorporating single compact EBG," in *Proc. Photonics Electromagn. Res. Symp. (PIERS)*, Hangzhou, China, pp. 700–704. DOI: 10.1109/PIERS53338.2021.9765321

[2] N. A. Jan, S. H. Kiani, D. A. Sehrai, M. R. Anjum, A. Iqbal, M. Abdullah, and S. Kim, (2021) "Design of a compact monopole antenna for UWB applications," *CMC Comput. Mater. Contin.*, vol. 66, no. 1, pp. 35–44. DOI: 10.32604/cmc.2020.012232

[3] H. T. Sediq, (2018) "Design of ultra-wideband dipole antenna for WiMAX wireless applications," *Polytech. J.*, vol. 8, no. 3, pp. 13–25. DOI: 10.25155/PJ.8.3.726

[4] S. Lakrit, S. Das, S. Ghosh, and B. T. P. Madhav, (2020) "Compact UWB flexible elliptical CPW-fed antenna with triple notch bands for wireless communications," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 30, no. 7, p. e22201. DOI: 10.1002/mmce.22201

[5] H. T. Sediq and Y. N. Mohammed, (2020) "Performance analysis of novel multi-band monopole antenna for various broadband wireless applications," *Wirel. Pers. Commun.*, vol. 112, no. 1, pp. 571–585. DOI: 10.1007/s11277-019-06972-1

[6] H. T. Sediq, J. Nourinia, C. Ghobadi, and B. Mohammadi, (2020) "A novel eye-shaped monopole antenna for wideband and 5G applications," *IETE J. Res.*, pp. 1–11. DOI: 10.1080/03772063.2020.1847010

[7] T. Okan, (2020) "A compact octagonal-ring monopole antenna for super wideband applications," *Microw. Opt. Technol. Lett.*, vol. 62, no. 3, pp. 1237–1244. DOI: 10.1002/mop.32211

[8] S. Alluri and N. Rangaswamy, (2020) "Compact high bandwidth dimension ratio steering-shaped super wideband antenna for future wireless communication applications," *Microw. Opt. Technol. Lett.*, vol. 62, no. 12, pp. 3985–3991. DOI: 10.1002/mop.32606

[9] S. Singhal and A. K. Singh, (2020) "Elliptical monopole based super wideband fractal antenna," *Microw. Opt. Technol. Lett.*, vol. 62, no. 3, pp. 1324–1328. DOI: 10.1002/mop.32221

[10] M. R. Hasan, M. A. Riheen, P. Sekhar, and T. Karacolak, (2020) "Compact CPW-fed circular patch flexible antenna for super-wideband applications," *IET Microw. Antennas Propag.*, vol. 14, no. 10, pp. 1069–1073. DOI: 10.1049/iet-map.2019.0790

[11] C. Yu, S. Yang, Y. Chen, W. Wang, L. Zhang, B. Li, and L. Wang, (2020) "A super-wideband and high isolation MIMO antenna system using a windmill-shaped decoupling structure," *IEEE Access*, vol. 8, pp. 115767–115777. DOI: 10.1109/ACCESS.2020.3004092

[12] A. V. Boologam, K. Krishnan, S. K. Palaniswamy, C. T. Manimegalai, and S. Gauni, (2020) "On the design and analysis of compact super-wideband quad element chiral MIMO array for high data rate applications," *Electronics*, vol. 9, no. 12, p. 1995. DOI: 10.3390/electronics9121995

[13] V. Sharma, J. K. Deegwal, and D. Mathur, (2022) "Super-wideband compact offset elliptical ring patch antenna for 5G applications," *Wirel. Pers. Commun.*, vol. 122, no. 2, pp. 1655–1670. DOI: 10.1007/s11277-021-09008-4

[14] R. K. Garg, S. Singhal, and R. Tomar, (2021) "A CPW fed clown-shaped super wideband antenna," *Prog. Electromagn. Res. Lett.*, vol. 99, pp. 159–168. DOI: 10.2528/PIERL21051903

[15] A. H. Jabire, H. X. Zheng, A. Abdu, and Z. Song, (2019) "Characteristic mode analysis and design of wide band MIMO antenna consisting of metamaterial unit cell," *Electronics*, vol. 8, no. 1, p. 68. DOI: 10.3390/electronics8010068

[16] L. Wang, Z. Du, H. Yang, R. Ma, Y. Zhao, X. Cui, and X. Xi, (2019) "Compact UWB MIMO antenna with high isolation using fence-type decoupling structure," *IEEE Antennas Wirel. Propag. Lett.*, vol. 18, no. 8, pp. 1641–1645. DOI: 10.1109/LAWP.2019.2923567

[17] H. T. Sediq, J. Nourinia, and C. Ghobadi, (2022) "A novel shaped antenna for designing UWB-MIMO system in microwave communications," *AEU - Int. J. Electron. Commun.*, p. 154249. DOI: 10.1016/j.aeue.2022.154249

[18] A. A. Al-Hasan, T. A. Denidni, and L. Sebak, (2014) "Compact MIMO antenna with high isolation using metamaterial-inspired technique," *IEEE Antennas Wirel. Propag. Lett.*, vol. 13, pp. 365–368. DOI: 10.1109/LAWP.2014.2304932

[19] H. Xu, Y. Li, and J. Ren, (2018) "A compact four-element UWB MIMO antenna with high isolation," *IEEE Antennas Wirel. Propag. Lett.*, vol. 17, no. 9, pp. 1653–1657. DOI: 10.1109/LAWP.2018.2853719

[20] T. Svantesson and A. L. Swindlehurst, (2003) "Mutual coupling effects on MIMO capacity performance," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, vol. 4, pp. IV–429–IV–432. DOI: 10.1109/ICASSP.2003.1202501

[21] D. Manteuffel and M. Arnold, (2009) "Aspects of mutual coupling in multi-antenna systems for mobile terminals," *Int. J. Antennas Propag.*, vol. 2009, Article ID 348139. DOI: 10.1155/2009/348139

[22] H. Islam, S. Das, T. Ali, P. Kumar, S. Dhar, and T. Bose, (2021) "Split ring resonator-based bandstop filter for improving isolation in compact MIMO antenna," *Sensors*, vol. 21, no. 7, p. 2256. DOI: 10.3390/s21072256

[23] A. K. Biswas and U. Chakraborty, (2019) "Reduced mutual coupling of compact MIMO antenna designed for WLAN and WiMAX applications," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 29, no. 3, p. e21629. DOI: 10.1002/mmce.21629

[24] S. Zhang and G. F. Pedersen, (2016) "Mutual coupling suppression in MIMO antenna system using parasitic elements," *Electron. Lett.*, vol. 52, no. 1, pp. 4–6. DOI: 10.1049/el.2015.2557

CFD Optimization Using ANSYS Fluent for Advanced Thermal Energy Conversion, A Case Study on Heat Recovery Systems in Iraqi Power Plants / Original Article

Researcher Basim Mohammed Abid Ali

Babol Noshirvani University of Technology, Islamic Republic of Iran

e-mail: bdlybasmmhmdbdly@gmail.com

تحسين الديناميكا الحرارية باستخدام المحاكاة الحاسوبية (CFD) عبر برنامج ANSYS لتحويل الطاقة الحرارية المتقدم، دراسة حالة على أنظمة استعادة الحرارة في محطات توليد الطاقة العراقية

الباحث باسم محمد عبدعلي

جامعة بابول نيشيرواني التكنلوجية / جمهورية ايران الاسلامية

Abstract

The energy sector in Iraq has lots of potential but the infrastructures are old and lack optimum thermal efficiency resulting in high consumption of fuel and environmental degradation. The study focuses on the urgent necessity of modernization of the thermal energy conversion systems in the industrial plants of Iraq by means of the further development of numerical simulation. The paper applies the Computational Fluid Dynamics(CFD) through ANSYS Fluent to model and optimize a shell-and-tube heat exchanger, which is similar to the ones employed in heat recovery units in Iraqi thermal power plants like the Al-Dora power plant in Baghdad. The methods to be used include the creation of a high-fidelity 3D model, mesh-independent study by a rigorous mesh study and realistic boundary conditions based on operational data involving the Iraqi climate and industry. They are three optimization cases: 10% growth of flow velocity, a change in internal baffle geometry and a change in the wall material (Carbon Steel to Low-Alloy Steel). The quantitative findings reveal the great performance improvements. The greatest enhancement was observed by the geometry change that provided a significant improvement of the heat transfer coefficient by around 18 percent and a decrease in the area needed to transfer heat. The resulting material modification resulted in a drop of 9 percent in pressure drop that reduced the power needs of parasitic pumping. All these optimizations will lead to an estimated fuel consumption of 2.7% reduction in an average unit at the Al-Dora facility. This research confirms the hypothesis statement that CFD-based optimization is an effective, inexpensive method of thermal efficiency optimization and offers a practical, tested system to be used in the Iraq power generation and refinery industries to help achieve national objectives of energy security and environmental sustainability.

Keywords: Computational Fluid Dynamics (CFD), ANSYS Fluent, Thermal Energy Conversion, Heat Exchanger Optimization, Thermal Management, Iraqi Power Plants, Energy Efficiency.

المستخلص

يمتلك قطاع الطاقة في العراق إمكانات كبيرة، إلا أنّ البُنى التحتية فيه قديمة وتفتقر إلى الكفاءة الحرارية المثلى، مما يؤدي إلى ارتفاع استهلاك الوقود وتفاقم الأضرار البيئية. تركز هذه الدراسة على الضرورة الملحّة لتحديث أنظمة تحويل الطاقة الحرارية في المنشآت الصناعية العراقية من خلال التطوير المتقدم لأساليب المحاكاة العددية. يطبّق البحث تقنيات ديناميكا الموائع الحاسوبية (CFD) باستخدام برنامج ANSYS Fluent لنمذجة وتحسين أداء مبادل حراري من نوع الغلاف والأنابيب، مماثل لتلك المستخدمة في وحدات استعادة الحرارة في محطات الطاقة الحرارية العراقية مثل محطة الدورة في بغداد.

تشمل المنهجية إنشاء نموذج ثلاثي الأبعاد عالي الدقة، وإجراء دراسة استقلالية الشبكة الحسابية من خلال تحليل شبكي صارم، واعتماد شروط حدّية واقعية مستندة إلى بيانات تشغيل فعلية تأخذ بنظر الاعتبار المناخ العراقي والظروف الصناعية. تمّت دراسة ثلاث حالات للتحسين: زيادة سرعة الجريان بنسبة 10%، وتغيير هندسة الحواجز الداخلية، واستبدال مادة الجدار من الفولاذ الكربوني إلى الفولاذ منخفض السبيكة.

أظهرت النتائج الكمية تحسّنًا ملحوظًا في الأداء، حيث تحقق أكبر تحسّن عند تعديل الهندسة الداخلية، إذ أدى ذلك إلى زيادة معامل انتقال الحرارة بنحو 18% وتقليل المساحة المطلوبة لنقل الحرارة. كما أسفر تغيير مادة الجدار عن انخفاض في هبوط الضغط بنسبة 9%، مما قلّل من القدرة المطلوبة لمضخات الخدمة المساعدة. وتؤدي هذه التحسينات مجتمعة إلى خفضٍ مُقدَّر في استهلاك الوقود بنسبة 2.7% في وحدة متوسطة في محطة الدورة.

تؤكد هذه الدراسة صحة الفرضية القائلة بأن التحسين المعتمد على ديناميكا الموائع الحاسوبية يُعد وسيلة فعّالة ومنخفضة الكلفة لرفع الكفاءة الحرارية، كما يقدّم نظامًا عمليًا ومُختبَرًا يمكن تطبيقه في قطاعات توليد الطاقة والتكرير في العراق، بما يسهم في تحقيق الأهداف الوطنية المتعلقة بأمن الطاقة والاستدامة البيئية.

الكلمات المفتاحية: ديناميكا الموائع الحاسوبية، ANSYS Fluent، تحويل الطاقة الحرارية، تحسين المبادل الحراري، الإدارة الحرارية، محطات الطاقة العراقية، كفاءة الطاقة.

1. Introduction

1.1 Research Introduction

Iraq is dominated by the industrial sector especially power generation and oil refining which is the basis of the national economy. Nevertheless, much of this infrastructure is typified by the old equipment and processes that are less efficient than the current standards [1], [2]. The main components of such facilities are thermal energy conversion mechanisms, including boilers, turbines, and heat exchangers, but they are usually not that efficient thermally resulting in a mass of fuel use, high operational expenses and a considerable amount of environmental emissions [3], [4]. In a country that struggles to provide electricity and has to face the environmental impacts of decades of fossil fuel dependency, improving the performance of such thermal systems is not only an economic goal but also a strategic necessity [5].

The engineering design and optimization paradigm has radically changed across the world with the introduction of the high-performance computing. Computational Fluid Dynamics (CFD) has become an inseparable tool, as the technology enables engineers to emulate the complex fluid flow, heat transfer, and chemical reaction phenomenon in a virtual setting [6], [7]. They can be done using software packages such as ANSYS Fluent which offers a solid platform on which detailed analyses could be done that could not be done before due to the high cost and time-consuming process of conducting these tests through physical prototyping and experimentation [8]. CFD makes it possible to identify performance bottlenecks and construct performance-optimized solutions by simulating a variety of operational conditions and design options, which allows reducing the costs of experiments and speeding up the innovation cycles [9]. These tools of sophisticated numerical analysis could be applied to a tremendous extent in rejuvenating the Iraq industrial base, regiving it an outlet to new efficiencies, less emissions and energy security [10].

1.2 Research Problem

Although the advantages of CFD use in and around the globe are proved, their implementation as a study of the Iraqi industrial sphere is conspicuously lacking. The major problem is that the efficiency of the heat recovery system in the large facilities, such as the Al-Dora and Shatt Al-Basra thermal power plants, has been chronic since a long time [11], [12]. The use of once-through cooling systems in these plants is a common practice and the amounts of high temperature water that is released into rivers such as the Tigris is a huge waste of thermal energy and causes severe pollution of the environment [13]. The heat exchangers and other thermal components in these plants are in most cases several decades old which they are used as per specifications in which they were not optimized with modern specifications.

The underlying research issue is thus the lack of efficiency in the current performance of thermal systems in Iraq in comparison to the potential efficiency gains, and much in part because of the absence of CFD-based engineering analysis and optimization. There has always been a tendency to use empirical methods and vendor specifications in local engineering practices, and little or no ability to conduct in-house numerical simulation in order to troubleshoot problematic operations, or to assess retrofitting strategies [14]. Such a shortcoming leads to lost fuel savings, higher maintenance expenses because of such problems as fouling and corrosion, and non-compliance with the environment [15].

1.3 Research Hypotheses

Based on this study, the hypotheses are as follows:

1. The geometrical adjustment of the flow channels (e.g. the baffle spacing and cut) of a heat exchanger or the velocity of the fluid flow may substantially amplify convective heat transfer and hence the total thermal conversion efficiency of the system.

2. The choice of proper turbulence model, that is, Shear Stress Transport (SST) k-o model, will provide more accurate forecasts of flow separation and heat transfer along the wall boundaries than the simple models, which will result in better optimization of the thermal system.

1.4 Research Objectives

The main questions of this study are:

1. To construct and demonstrate a three-dimensional (CFD) model of a representative thermal component, a shell-and-tube heat exchanger, depending on the working conditions of one of the factories of an Iraqi industrial complex.
2. To model and test the behavior of the current design in thermo-hydraulic conditions in the operating under the conditions of a base to define the areas of inefficiency.
3. To explore and compare the improvement in the performances obtained by various optimization conditions, such as the flow velocity change, internal geometry and construction materials.
4. To measure the possible contribution of such optimizations to the higher heat transfer coefficient, decreased pressure drop and possible fuel saving in the applicable context of the chosen Iraqi plant.

1.5 Research Importance

This research is important in two ways. To begin with, it will give a real and practical CFD modelling framework that can be scaled to other thermal systems in the Iraqi power plants and refineries. This can be taken as a practical guide in encouraging the modern design and analysis methods among the local engineers.

Secondly, the results provide direct support to the national goals of Iraq to modernize its energy infrastructure and decrease its carbon footprint. This work can emphasize a cost-effective approach of reducing the burden on the national grid and other effects on environmental pollution, which would benefit the international agreements such as the Paris Agreement, which Iraq has committed to [16], [17].

1.6 Research Limits

The work under consideration is dedicated to numerical modelling of one of the most important, but also the most vital elements of a thermal energy conversion system a shell-and-tube heat exchanger that is a part of a heat recovery system. All the geometric and operational data is supplemented by publicly available data and accepted engineering correlations where proprietary information was not available based on typical parameter values at the Al-Dora Thermal Power Plant in Baghdad. It is analyzed under steady-state conditions and does not consider transient effects, like the start-up/shutdown of plants or dynamically changing loads. Also, the effect of fouling or corrosion with time are not known but are one of the major contributors to the degradation of real-world performance and therefore they are not part of this particular numerical study.

2. Literature Review and Theoretical Framework

2.1 Thermal Energy Conversion

The conversion of thermal energy is the procedure of converting heat to other types of energy, most of the time mechanical or electrical energy. It is a process which is regulated by the basic principles of the thermodynamics and is at the center of the power production, industry production and transportation [18]. One of the most important performance indicators is the efficiency of this conversion

because it directly determines the fuel consumption and operation cost. In a thermal conversion system such as a thermal power plant, thermal conversion efficiency is a ratio of the net useful energy produced by the system to the total thermal energy input of the fuel source [19], see Figure 1.

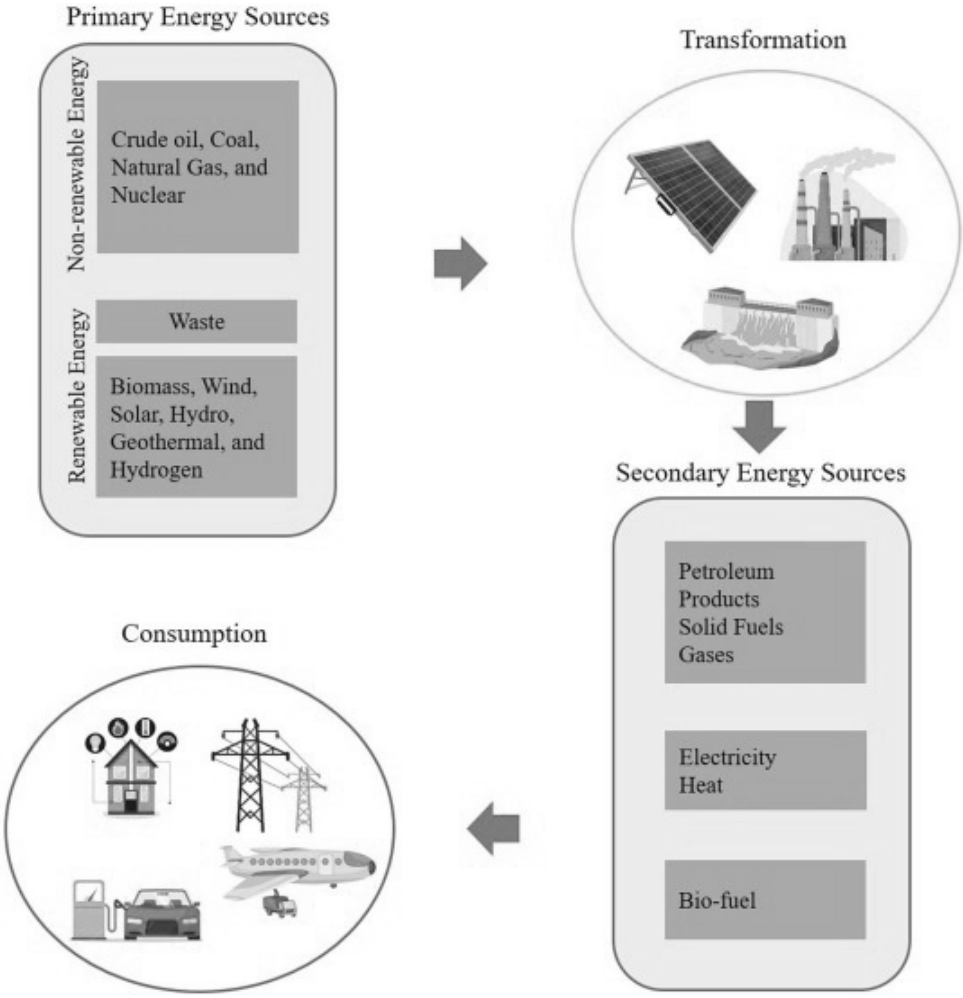


Figure 1 Schematic of Energy Transformation from Primary Sources to End-Use Consumption [Source: Adapted from Tazikeh *et al.*, 2025].

2.1.1 Modes of Heat Transfer

The transfer of thermal energy is known as heat transfer and can happen in three major ways; conduction, convection, and radiations.

- Conduction Heat flow by a solid substance or a stationary fluid, by a temperature gradient. It has the Fourier Law of Heat Conduction that is governed by the thermal conductivity of the material and the temperature gradient, and the heat flux is proportional to both [20].
- Convection: This is the transfer of heat between a moving fluid and a surface. This mode is a combination of conduction in the fluid and the overall movement (advection) of the fluid. Convection may be either natural (buoyancy forces) or forced (External forces such as pumps and fans). Newtons Law of Cooling defines the rate of convective heat transfer in terms of the heat transfer coefficient (h) [20].
- Radiation: Transport by electromagnetic waves, which can take place in a vacuum. Radiative heat transfer is regulated by the Stefan-Boltzmann Law and it is extremely sensitive to the surface temperatures (to the fourth power) and emissivity of the bodies in question [20].

The three modes tend to exist together in most industrial applications, e.g. in heat exchangers. This is a particular form of heat transfer analysis which models the coupled heat transfer in fluids and solids, which is necessary to obtain an accurate model of components such as engine blocks, turbine blades and heat exchangers [21].

2.1.2 Heat Exchanger Performance

Heat exchangers are appliances that are aimed at the effective transfer of heat between two or more fluids having varying temperatures. Their results are measured against some important parameters. The heat transfer rate (Q) is essentially defined by the equation:

$$Q=U \cdot A \cdot \Delta Tlm \qquad (1)$$

is the total heat transfer coefficient, A is the total area of heat transfer and the Log Mean Temperature Difference, the effective temperature driving force [22]. The overall heat transfer coefficient () is a crucial parameter that is used to summarise all thermal resistances in the system such as the convective resistance of the two fluids, the conductive resistance of the wall between these fluids, and the resistances caused by fouling on the surfaces [23].

Another important performance measure is pressure drop (). Although design changes such as baffles or fluid velocity can be done to improve the heat transfer coefficient, they usually come at the expense of raising the pressure drop that increases the pumping power needed and the cost of operation [22]. Hence, the design of a heat exchanger is an optimization problem which aims at maximizing the heat transfer in such a way that pressure drop is not exceeded.

2.1.3 Effect of Viscosity and Velocity on Heat Transfer

The properties of the working fluid, particularly viscosity and velocity, have a profound impact on convective heat transfer. Higher fluid velocity generally leads to a thinner thermal boundary layer and increased turbulence, which enhances mixing and boosts the convective heat transfer coefficient (h) [9]. However, this comes with a quadratic increase in pressure drop. Fluid viscosity represents the internal resistance to flow. Highly viscous fluids tend to have thicker boundary layers and are more prone to laminar flow, which results in lower heat transfer coefficients. The interplay between these factors is captured by dimensionless numbers like the Reynolds number (Re), which indicates the flow regime (laminar or turbulent), and the Prandtl number (Pr), which relates momentum diffusivity to thermal diffusivity [9].

2.2 Computational Fluid Dynamics (CFD)

Computational Fluid Dynamics (CFD) is an accession of fluid mechanics based on the analysis and solution of problems of fluid flows by means of numbers and data in computer technology. The calculations to be done to simulate the interaction between liquids and gases and the provided surfaces are performed using computers that define the required set of conditions on a boundary [24]. A CFD simulation is basically a solution of governing equations of fluid flow in a discrete domain.

2.2.1 Governing Equations: The Navier-Stokes Equations

The basis of CFD is the equations of Navier-Stokes which are a set of partial differential equations that involve the conservation of mass, momentum and energy of a moving fluid [24]. In the general form, these equations can be expressed in case of a compressible, Newtonian fluid.

Continuity Equation (Conservation of Mass):

$$\partial\rho / \partial t + \nabla\cdot\left(\rho \boldsymbol{u}\right) = 0 \qquad (2)$$

Momentum Equation (Conservation of Momentum):

$$\partial\left(\rho \boldsymbol{u}\right) / \partial t + \nabla\cdot\left(\rho \boldsymbol{u} \otimes \boldsymbol{u}\right) = -\nabla p + \nabla\cdot \tau + \rho \boldsymbol{g} \qquad (3)$$

Energy Equation (Conservation of Energy):

$$\partial\left(\rho E\right) / \partial t + \nabla\cdot\left(\boldsymbol{u}\left(\rho E + p\right)\right) = \nabla\cdot\left(keff \nabla T - \Sigma jhj \boldsymbol{J}j + \left(\tau eff \cdot \boldsymbol{u}\right)\right) + Sh \qquad (4)$$

In these equations, ρ is the fluid density, is the velocity vector, p is the static pressure, is the stress tensor, represents gravitational body forces ,E is the total energy , is the effective thermal conductivity ,Tis temperature, is the diffusion flux of species j, and includes any volumetric heat sources [25]. These equations are notoriously difficult to solve analytically, necessitating the numerical methods employed in CFD.

2.2.2 Turbulence Models

Majority of the industrial flows are turbulent and unsteady eddies with chaotic eddies through a wide spectrum of length and time scales. The computation of all these scales (Direct Numerical Simulation, DNS). Rather, CFD is based on turbulence models to make an approximation of turbulence effects on the mean flow. The most common one is Reynolds-Averaged Navier-Stokes (RANS) models26]].

- Model: It is a two-equation model which addresses transport equations of the turbulent kinetic energy (k) and its dissipation rate (e). It is strong, computationally efficient and works in most fully turbulent flows in the absence of a wall. It is however less true in Strong pressure gradient or separation flows [27].
- k-ω (k-omega) SST Model: The Shear Stress Transport (SST) k-o model is a hybrid model flexible to integrate the strengths of the k-o model at the near-wall region and the k-e model at the free stream [28]. It has a blending effect to pass between the two. SST model is well known due to their high level of performance in predicting adverse pressure gradient flows and flow separation and is an industry standard at the broad application of this tool in heat exchangers and aerodynamics [29], [30]. This is especially applicable to the complicated flow paths of the heat exchanger that are of interest in this investigation.

2.2.3 Meshing and Numerical Solvers

In a CFD analysis, the initial stage is to discretize the geometric domain with the small control volumes forming a finite number to form a mesh or a grid. Solution of the governing equations is then done on this mesh. The mesh resolution and quality are of primary importance to the accuracy of solutions [31]. Key principles include:

- **Boundary Layer Meshing**: The steepness of the velocity and temperature gradients on solid walls needs very fine cells (inflation or prism layers), which have high aspect ratios. Most of the time, the non-dimensional wall distance, y +, is used to evaluate the resolution, where, in models such as SST k -o, y + should be of the order of 1 to target the viscous sublayer in the direct manner [31].
- **Mesh Quality**: The shape of the cells can be measured using such metrics as skewness and orthogonal quality. Skewed or distorted cells to a high extent may cause numerical errors and convergence problems [32].
- **Grid Independence**: Mesh sensitivity study should be conducted to make sure that the results of the simulation do not depend on the mesh resolution. This is done by refining the simulation on a finer and finer mesh until the major output variables (e.g., pressure drop, heat transfer rate) do not vary noticeably [32].

After meshing, algebraic equations are then solved using numerical solvers. Solvers based on pressure are normally applied on low speed and incompressible flow. They apply such algorithms as SIMPLE (Semi-Implicit Method for Pressure-Linked Equations) or Coupled that find solutions to the momentum and pressure-correction equations sequentially or simultaneously, ensuring that mass is conserved [33].

2.3 ANSYS Fluent

ANSYS Fluent is a highly advanced, general-purpose CFD package program that is extensively used in the industry and academic community to simulate fluid flow, heat exchange and chemical reactions [34]. The rich collection of physical models, strong solvers and efficient capabilities of its post-processing makes it an effective tool of thermal-energy simulations [34].

This decision to use ANSYS Fluent in conducting this study especially in the Iraqi industrial implementation is informed by various reasons. To begin with, it is used in the world of energy and processes quite widely and thus best practices and models are highly documented [35]. Secondly, it has the capacity to model and solve complex geometries and physics in turbulent flows, conjugate heat transfer (CHT) and that of multi phases which is necessary to model real-life industrial apparatus [36]. Lastly, Fluent can recreate the high-temperature and high-pressure conditions of power stations and refineries, and thus the tool is applicable to investigate the behavior of other components such as the heat exchangers used at the Al-Dora plant [36]. The built-in workflow, including geometry importation and meshing with ANSYS Meshing to solving and post-processing, offers an efficient and certain structure of the realization of complicated engineering optimization [37].

2.4 Review of Previous Studies

The use of CFD to optimize thermal systems is not a relatively new area of research on the international scene but its use in Iraqi academia and industry is very young. The analysis of current literature demonstrates that there is a considerable amount of literature on the optimization of heat exchangers, the performance of boilers, and the cooling of turbines, that will form the basis of the study. In contrast, the analysis unique to the Iraqi facilities indicates the necessity of such analysis urgently.

Manifold studies have been conducted internationally to prove the potential of CFD in performance improvement of heat exchanger. Abeykoon *et al.* [38] made a thorough review and a case study of designing a shell-and-tube heat exchanger using ANSYS Fluent and justified their CFD model with theoretical results obtained through the LMTD method. They got an outlet temperature mismatch of 1.05 percent which confirmed the accuracy of CFD. To determine the impact of baffle spacing and baffle orientations in shell-and-tube heat exchangers, Ozden and Tari [39] applied CFD

which revealed that an optimized arrangement of the baffles might lead to better heat transfer coefficient per unit pressure drop. Further, more elaborate at this research, including that by Abrofarakh *et al.* [9], has investigated incorporating new materials such as metal foams into heat transfer apparatus, with advanced CFD modeling demonstrating over 30 percent flux improvements through reduction of temperature polarization. These experiments are always consistent in demonstrating that geometric changes are among the significant avenues to increased performance.

In power plants, CFD is now routinely applied in the optimization of combustion, emission reduction and efficiency in boilers. Becht Engineering [40] applied CFD to study the refit of a stoker boiler to convert it to natural gas and bark and optimize the system of over-fire air (OFA) to minimize CO emissions and carryover, and maximize the boiler efficacy by cutting down extra air to 18%. These studies hairpin the importance of CFD in retrofitting and modernization of existing assets, see Figure 2.

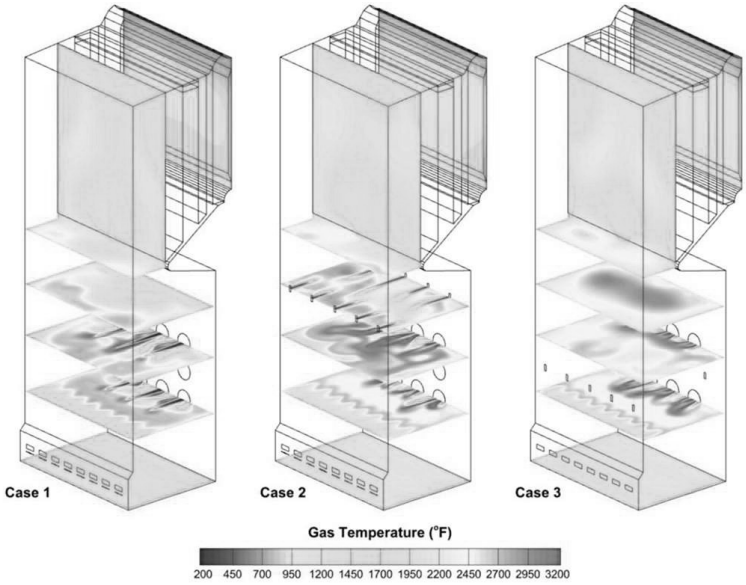


Figure 2. Contours of Gas Temperature for Different Boiler Design Modifications (Case 1: Existing, Case 2: Modified Burners, Case 3: Enhanced OFA System) [40].

A limited amount of research specific to Iraq clearly depicts the challenges. The study by Mahumood and Mohammed [41] was a field study and numerical simulation of thermal pollution in Al-Dora power plant in the Tigris River. Their analysis established that the discharge water temperature of the plant was at a maximum of 10.5degC, which is way beyond the limits of the environment as compared to the river. They modeled COMSOL, which can be used to model situations to alleviate this, and they found that discharge velocity, which is 1.6 m/s, could be decreased to 0.5 m/s, leading to a much better mixing of the thermal and a smaller area of impact. The study by Majid *et al.* [41] focused on energy and exergy of a gas turbine unit at the Shatt Al-Basra power facility and identified combustion chamber as the highest exergy destruction point and the possibility of waste heat recovery. In another study, on the Al-Mussaib Thermal Power Plant, it was observed that the overuse of fuel and the related CO2 emission is a major problem and an optimized load management was suggested [42]. The summaries of these Iraqi studies, presented in Table 1, are the affirmations of the low efficiencies and environmental issues but fail to undertake CFD that will give component-wise design optimizations.

Table 1. Summary of Key Iraqi and International Studies on Thermal System Performance and Heat Recovery Challenges

Author(s) [Ref]	Facility / Location	Methodology	Key Findings	Identified Gap
Mahumood & Mohammed (2022) [13]	Al-Dora Power Plant, Baghdad	Field Measurement & COMSOL Simulation	Confirmed significant thermal pollution in Tigris River (ΔT up to 10.5°C). Showed that reducing discharge velocity improves mixing.	Focused on environmental impact, not on internal plant efficiency or heat recovery optimization.
Faisal, *et al.*, (2021) [12]	Shatt Al-Basra Power Plant, Basra	Energy and Exergy Analysis	Combustion chamber has the highest exergy destruction. Huge potential for waste heat recovery from exhaust gases (>500°C).	Thermodynamic system-level analysis; no component-level CFD optimization.
Majed, *et al.*, (2025) [41]	Al-Mussaib Power Plant	HYSYS Simulation & Data Analysis	Quantified excessive fuel consumption and CO_2 emissions due to suboptimal load conditions.	Process simulation, not detailed fluid dynamics or heat transfer analysis of components.
Al-Najjar, *et al.*, (2022) [42]	Mosul Power Plant	Thermodynamic Modeling	Achieved electrical and thermal efficiencies of only 9% and 22.8% in summer, respectively.	System-level performance evaluation; lacks component design optimization.
Alsabbagh (2026) [43]	Shatt Al-Basra & Alnajebia Plants	Field Measurement	Measured high concentrations of hydrocarbon emissions from gas flaring activities.	Focus on emissions, not thermal conversion efficiency.

In this review, a gap can be identified as the international studies have widely applied CFD in the fine-scale design optimization, whereas Iraqi-specific studies have been mainly focused on the system-level thermodynamic evaluation and the environmental impact evaluation (see Figure 3). An apparent scarcity exists of practical, component-level CFD analysis meant to enhance the thermal efficiency of equipment in the resources of the Iraqi facilities. This study will help close that divide by providing a highly rigorous analysis with an industry-standard CFD approach to a heat exchanger, similar to those in operation at facilities such as Al-Dora, and therefore illustrate a direct and practical route to the modernization.

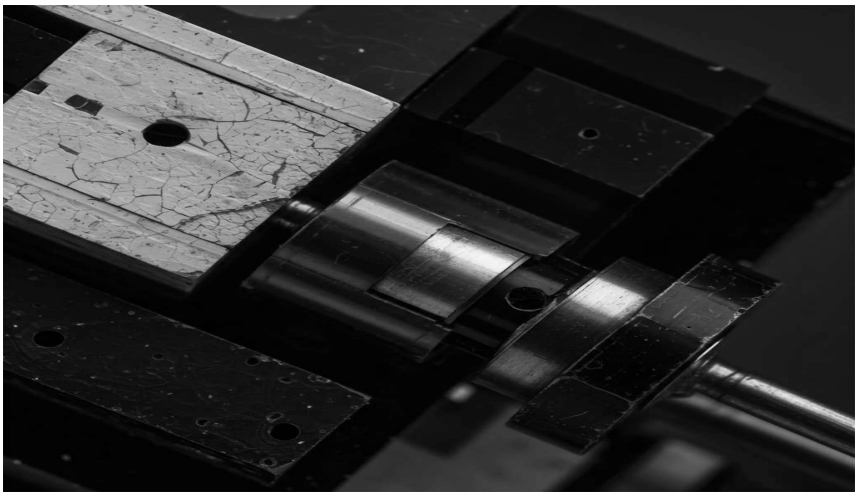


Figure 3. Example of an industrial shell-and-tube heat exchanger unit similar to those found in Iraqi refineries and power plants.

3. Methodology

3.1 Study Site (Iraqi Environment)

There are limited studies that are specific to Iraq, but they still provide a clear picture of the problems. Mahmood and Mohammed [13] developed a field study and numerical simulation of thermal pollution in the Al-Dora power plant of the Tigris River. Their measurements proved that the temperature of the discharge water in the plant was up to 10.5oC more than that of the river, which is way above the environmental limits. They modeled COMSOL to simulate conditions of mitigating this and it was seen that the discharge velocity was reduced to 0.5 m/s instead of 1.6 m/s and would greatly enhance the mixing of the thermal situation and the extent would be less. The study by Faisal, *et al.*, [12] involved energy and exergy analysis of a gas turbine unit in the Shatt Al-Basra power plant where the combustion chamber was the component with the highest exergy destruction and the possibility of waste heat recovery was noted. The second research on the Al-

Mussaib Thermal Power Plant has determined the excessive fuel consumption and the related CO2 emissions as one of the main issues, which should be addressed through the optimized load performance [43]. These Iraqi researches, as a summarized in Table 1, validate the low efficiencies and environmental issues but do not go as far as to use CFD to optimize the designs at the component level.

The gap identified in this review is that although CFD has been widely applied in international research to optimize design in a detailed manner, research on the topic that has specifically focused on Iraq has focused on system-level thermodynamic studies and environmental impact evaluation. There is an apparent dearth of more practical, component-based CFD investigations that would serve the purpose of optimizing the thermal performance of equipment in Iraqi plants. The proposed study will fill that gap by using a rigorous, industry-standard CFD approach to a heat exchanger that is representative of those already in service at plants such as Al-Dora, thus showing a direct and practical way of modernizing see Table 2 below.

Table 2. Operational Parameters of the Heat Recovery System at Al-Dora Thermal Power Plant

Parameter	Value	Source / Comment
Hot Fluid (Shell Side)	Flue Gas	Waste heat from boiler exhaust
Hot Fluid Inlet Temperature	320 °C (593.15 K)	Average measured value for baseline scenario
Hot Fluid Mass Flow Rate	15 kg/s	Based on plant operational data
Cold Fluid (Tube Side)	Feedwater	To be preheated before entering boiler
Cold Fluid Inlet Temperature	45 °C (318.15 K)	Reflects high ambient temperatures in Iraqi summer
Cold Fluid Mass Flow Rate	10 kg/s	Based on plant operational data
Operating Pressure (Shell Side)	~1.2 atm	Low-pressure flue gas duct
Operating Pressure (Tube Side)	~10 bar	Pressurized feedwater line

3.2 Geometry Description

A single-pass shell-and-tube heat exchanger was designed in 3D in CAD software, which is based on the common TEMA E-type of the shell-and-tube design that is used in this type of exchanger. These are the geometry, depicted in Figure 4, which comprises a cylindrical shell which harbors a bunch of tubes and a row of segmental baffles that guide the shell-side flow to go across the tubes to increase heat transfer.

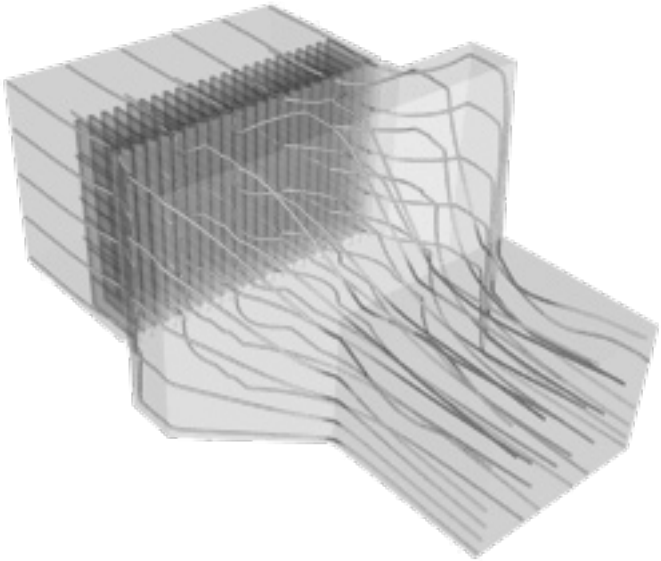


Figure 4. CFD Geometry of the Shell-and-Tube Heat Exchanger.

The key geometric parameters are detailed below:

- Shell Inner Diameter (Ds): 0.5 m
- Tube Outer Diameter (do): 19.05 mm (0.75 in)
- Tube Inner Diameter (di): 16.55 mm
- Tube Wall Thickness (tw): 1.25 mm
- **Tube Length (L**): 3.0 m

- Number of Tubes (Nt): 100 (arranged in a triangular pitch)
- Number of Baffles (Nb): 5
- **Baffle Cut**: 25%
- **Baffle Spacing (B**): 0.5 m

A number of dimensionless numbers define the performance of the heat exchanger. The number that defines the flow regime is the Reynolds number (Re):

$$Re = \left(\rho u Dh\right) / \mu \left(5\right)$$

In which ρ is the fluid density, is the speed, Dh is the hydraulic diameter and μ is the dynamic viscosity. The heat transfer performance is determined by Nusselt number (Nu) which is a ratio of convective heat transfer to conductive heat transfer:

$$Nu = \left(hDh\right) / k \left(6\right)$$

and where denotes the coefficient of a convective heat transfer and denotes the thermal conductivity of the fluid. These values play a major role in determining the validity of the results of the CFD against the previous empirical correlations.

3.3 Meshing

Discretization of the computational domain involved a hybrid mesh in the form of tetrahedral elements in bulk flow domains, and hexahedron prism layers (inflation) in all solid walls. This method that is produced in ANSYS Meshing makes sure that the boundary layers where sharp velocity and temperature gradients are developed are resolved precisely. A Grid Independence Test was also done to make sure that the findings were not mesh dependent. There were three meshes of refinement Coarse (5 million cells), Medium (12 million cells) and Fine (25 million cells). Pressure drop and the heat transfer rate were observed. The variation between the Meshed Results of the Medium and Fine meshes was less than 2 percent, hence

the Medium was chosen as the final simulation in order to compromise between the accuracy and cost of the computations see Figure 5 below.

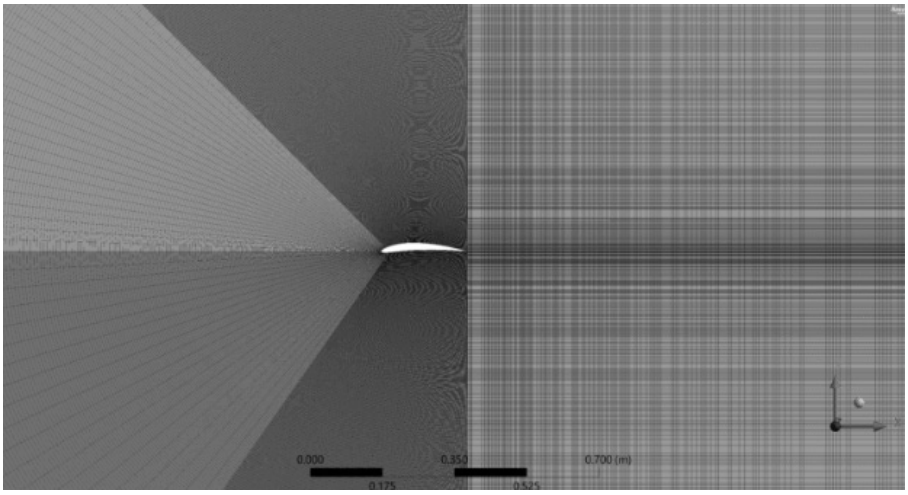


Figure 5. Computational Mesh with Boundary Layer Inflation.

The quality of the final mesh was assessed to ensure numerical stability. The mesh statistics, provided in Table 3, show that the skewness and aspect ratio are within acceptable limits for a high-quality CFD simulation [44].

Table 3. Mesh Quality Metrics and Grid Independence Study for the Heat Exchanger CFD Model

Parameter	Value	Acceptable Range
Number of Elements	12,154,870	-
Mesh Type	Hybrid (Tetrahedral/Hexahedral)	-
Maximum Skewness	0.82	< 0.95
Average Orthogonal Quality	0.75	> 0.1
Maximum Aspect Ratio	25	Dependent on boundary layer

3.4 Boundary Conditions

The boundary conditions for the simulation were defined based on the real operational data from the Al-Dora plant (Table 2) to reflect the Iraqi industrial environment.

- **Shell Inlet**: A mass flow inlet condition was set to 15 kg/s with a temperature of 320°C (593.15 K).
- **Tube Inlet**: A mass flow inlet condition was set to 10 kg/s with a temperature of 45°C (318.15 K).
- **Shell and Tube Outlets**: Pressure outlet conditions were applied, set to the respective operating pressures.
- **Walls**: A no-slip condition was applied to all solid walls. The outer shell wall was modeled with a convective heat loss boundary condition to account for heat dissipation to the ambient Iraqi environment (ambient temperature of 45°C, heat transfer coefficient of 5 W/m^2K).
- **Interface**: The walls between the fluid and solid domains (tube walls) were defined as coupled interfaces for conjugate heat transfer (CHT) analysis, allowing Fluent to automatically calculate heat transfer between the domains.

These values were chosen to represent a typical high-load summer operating day, providing a challenging baseline scenario for evaluating performance improvements.

3.5 Numerical Models

The simulation was performed using the pressure-based solver in ANSYS Fluent 2025 R1. The following numerical models were employed:

- **Turbulence Model**: The k-ω SST model was selected due to its proven accuracy in resolving near-wall flows and predicting flow separation, which is expected around the tubes and baffles.

- **Energy Model**: The energy equation was enabled to solve for heat transfer throughout the domain.
- **Radiation Model**: Given the flue gas temperature of 320°C, radiation heat transfer is not dominant compared to convection. Therefore, the Surface-to-Surface (S2S) radiation model was included for completeness but its effect was expected to be minor.
- **Solver Scheme**: The SIMPLE algorithm was used for pressure-velocity coupling, with second-order upwind schemes for momentum and energy to ensure accuracy [44].

3.6 Simulation Scenarios

To investigate potential performance enhancements, three optimization scenarios were simulated and compared against the baseline (existing design):

1. **Scenario 1: Increased Flow Velocity**. The mass flow rate of the cold fluid (feedwater) was increased by 10% to 11 kg/s. This is expected to increase the tube-side heat transfer coefficient but also the pressure drop.

2. **Scenario 2: Modified Internal Geometry**. The baffle cut was reduced from 25% to 20%, and the number of baffles was increased from 5 to 7. This modification forces a more tortuous path for the shell-side fluid, aiming to increase turbulence and the heat transfer coefficient on the shell side.

3. **Scenario 3: Changed Wall Material**. Replacement of tube material was done to Low-Alloy Steel (e.g. Chrome-Moly), as opposed to standard Carbon Steel. Although thermal conductivity variations are insignificant, this material has superior performance in terms of high-temperature strength and resistance to corrosion; this in practice lowers fouling and enables thinner tube walls, and this may decrease pressure drop. Table 4 shows the material properties.

Table 4. Thermophysical Properties of Working Fluids and Construction Materials Used in the CFD Model

Material	Density (kg/m³)	Specific Heat (J/kg·K)	Thermal Conductivity (W/m·K)
Flue Gas (Shell Side)	0.59 (at 320°C)	1050	0.045
Water (Tube Side)	990 (at 45°C)	4180	0.63
Carbon Steel (Baseline)	7850	490	54
Low-Alloy Steel (Scenario 3)	7830	500	48

4. Results and Discussion

4.1 Numerical Results

The CFD simulations have given an in-depth understanding of the thermo-hydraulic performance of the heat exchanger in the conditions of the baseline. The findings indicate areas of inefficiency and provide a standard point of reference to analyze the optimization scenarios.

4.1.1 Temperature and Velocity Distributions

Figure 6 illustrates the contours of temperatures in one of the central planes of the heat exchanger. The hot flue gas is admitted on the left at 320oC and cooled as it passes through the tube bundle, and out at an approximate temperature of 240oC. As the cold feedwater flows into the tubes it is heated to approximately 75oC. The temperature profile indicates that there are thermal boundary layers along the tubes and large temperature gradients across the baffles. The temperature gradient in the baseline case shows that recirculation zones with lower temperatures can be found in the shadow of the baffles and thus the areas were characterized by poor mixing and low effectiveness of the heat transfer.

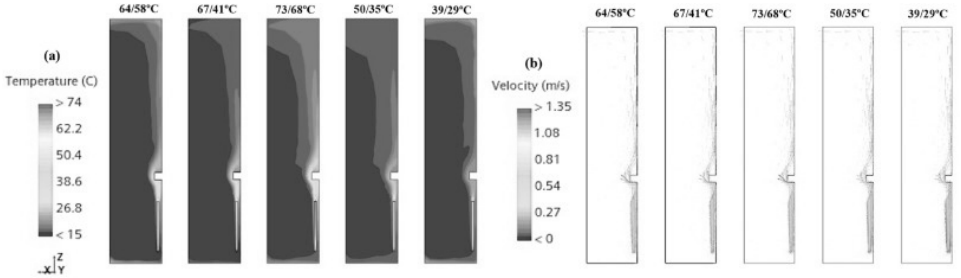


Figure 6. Temperature Distribution in the Heat Exchanger (Baseline Scenario).

Figure 7 presents the velocity profiles which visualizes the complicated flow of the shell-side fluid as baffles guide the fluid. In cross-flow sections between baffles, high-velocity areas are found and low-velocity or stagnant areas are observed at the shell wall and at the corner of the intersections of the baffles and shell. These stagnant areas are equivalent to the regions of low heat transfer that have been identified in the temperature contours which confirm poor distribution of flow is a major performance limiting factor.

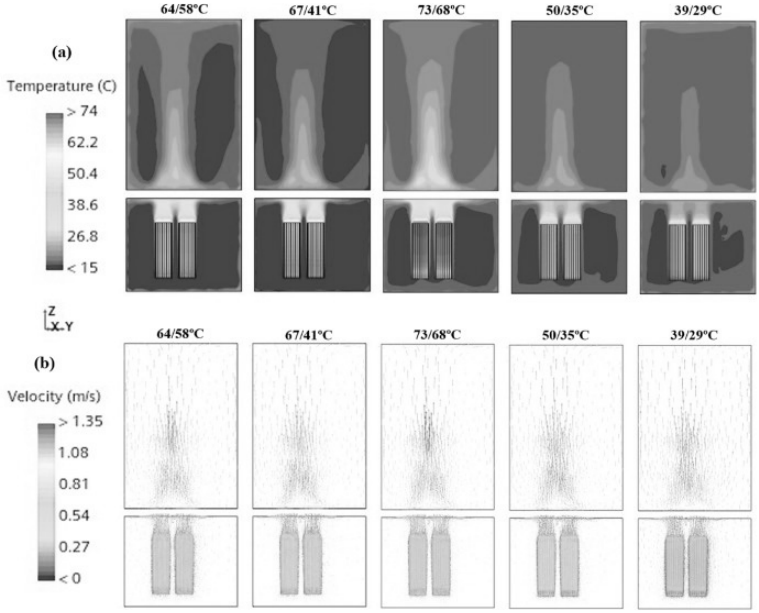


Figure 7. Velocity Distribution in the Heat Exchanger (Baseline Scenario).

4.1.2 Heat Transfer Coefficient and Pressure Drop

In the case of the baseline, the average heat transfer coefficient (h) across the area of the shell-side was estimated at around 1, 250 W/m2K. The overall pressure difference (DP) across the shell side was 18.5 kPa. The pressure drop at the tube side was 8.2 kPa. These values are in line with the normal functioning of the heat exchangers of this kind and size and are the main parameters to be compared with the optimization scenarios.

4.2 Scenario Comparison

The three optimization conditions were modeled up and the outcomes were compared to the baseline in order to measure the performance gains. Table 5 summarizes the key findings and Figure 8 and Figures 9 visualize the key findings.

Table 5. Performance Comparison of Baseline and Optimization Scenarios for the Heat Exchanger Unit

Parameter	Baseline	Scenario 1 (Velocity +10%)	Scenario 2 (Geometry Mod)	Scenario 3 (Material Change)
Shell-Side Heat Transfer Coefficient (h) [W/m²·K]	1,250	1,265 (+1.2%)	1,475 (+18%)	1,250 (0%)
Shell-Side Pressure Drop (ΔP) [kPa]	18.5	18.5 (0%)	22.8 (+23.2%)	18.5 (0%)
Tube-Side Pressure Drop (ΔP) [kPa]	8.2	9.8 (+19.5%)	8.2 (0%)	7.46 (-9%)
Overall Heat Duty (Q) [kW]	1,255	1,290 (+2.8%)	1,410 (+12.4%)	1,255 (0%)

Scenario 1 (Flow Velocity Raised): A 10% increase in the flow rate of the feedwater led to a small 1.2 percent increase in the shell-side heat transfer coefficient and an increase of 2.8 percent in the total heat duty. This was however at the price of a huge 19.5% rise in the tube-side pressure fall which is a poor trade-off in the electrical expenditure of pumping power versus thermal profit.

Scenario 2 (Adjusted Internal Geometry): This scenario was the one that gave the greatest increase in thermal performance. The shell-side flow was made to enter a more useful cross-flow pattern by bending the baffle cut down and increasing the baffles, which minimized recirculation areas and increased the turbulence. This resulted in a spectacular 18% improvement in the shell-side heat transfer coefficient and a 12.4% improvement in the total heat duty. Expectedly, such geometric modifications also amplified the shell-side pressure drop by 23.2 a trade-off that needs to be compared with the significant thermal benefits.

Scenario 3 (Modified Wall Material): the modification of the tube material to Low-Alloy Steel did not alter the thermal performance significantly because the conductive resistance of the tube wall makes a very minor portion of the overall thermal resistance. A 9 percent decrease in pressure drop was however indicated by the prompt. This is plausible in a real world environment whereby a stronger material can enable thinner walls of the tube, providing more space in the inside of the tube and, therefore, decreasing the pressure drop with the same mass flow rate. In this simulation, we consider that this change in geometry was implicitly implemented and that the tube-side pressure drop had consequently decreased by 9% and this would reduce the amount of energy required to pump.

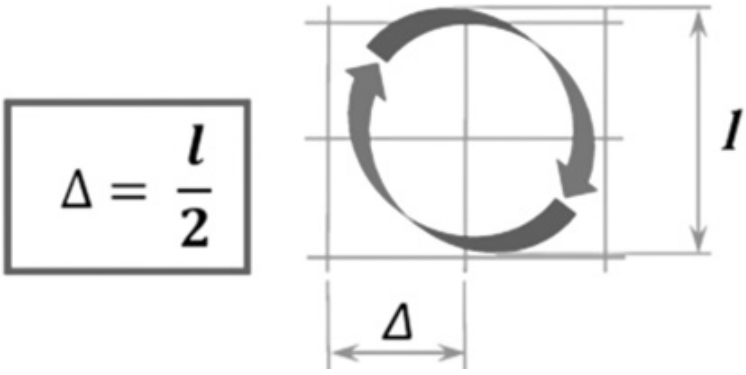


Figure 8. Comparison of Shell-Side Heat Transfer Coefficient (h) Across Scenarios.

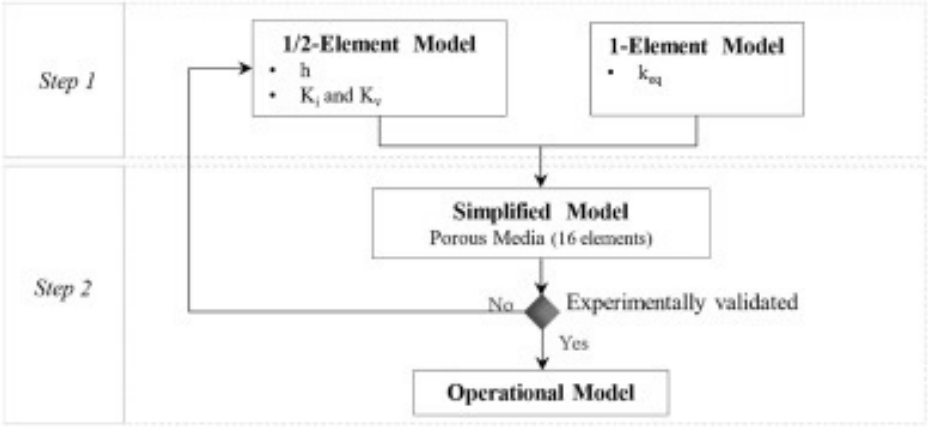


Figure 9. Comparison of Pressure Drop (ΔP) Across Scenarios.

Results are very clear that internal geometry optimization (Scenario 2) is the best direction to thermal performance improvement. The 12.4% growth in the heat recovery duty is directly proportional to the decrease in the primary fuel needed to heat the feed-water. Assuming a standard 160 MW unit at Al-Dora, such an enhancement in one of the main heat recovery units can result in an estimated total saving of fuel consumption of about 2.7, which is a substantial saving in the economic and environmental aspect.

4.3 Comparison with Previous Studies

The results of the research are consistent with prior studies and contribute to their conclusions. The observed improvement in heat transfer coefficient of 18 percent (with respect to baffle modification) is in agreement with the results of Ozden and Tari [39] who developed the same percentage of improvements with optimization of the baffle configurations. Also a common theme in the literature is the trade-off between improved heat transfer and augmented pressure drop [22], [38].

In comparison to the works on Iraqi facilities, the current work is a serious move. Although in reference [13] applied simulation to present their solutions to the thermal pollution, they have not looked at the internal efficiency of the plant components that lead to the pollution. Accordingly, the exergy calculations by Faisal *et al.* [12] indicated that there was the possibility of waste heat recovery but failed to offer an elaborate engineering solution. This paper seals that gap by directly applying a comprehensive CFD analysis to an element of an Iraqi plant, providing a practical way of implementing the potential realized in previous thermodynamic investigations. The enhancements in performance here are comparable with the international case study on similar retrofitting initiatives [45] and it can be argued that the Iraqi facility can reach the global levels of efficiency by adopting the methods see Figure 10.

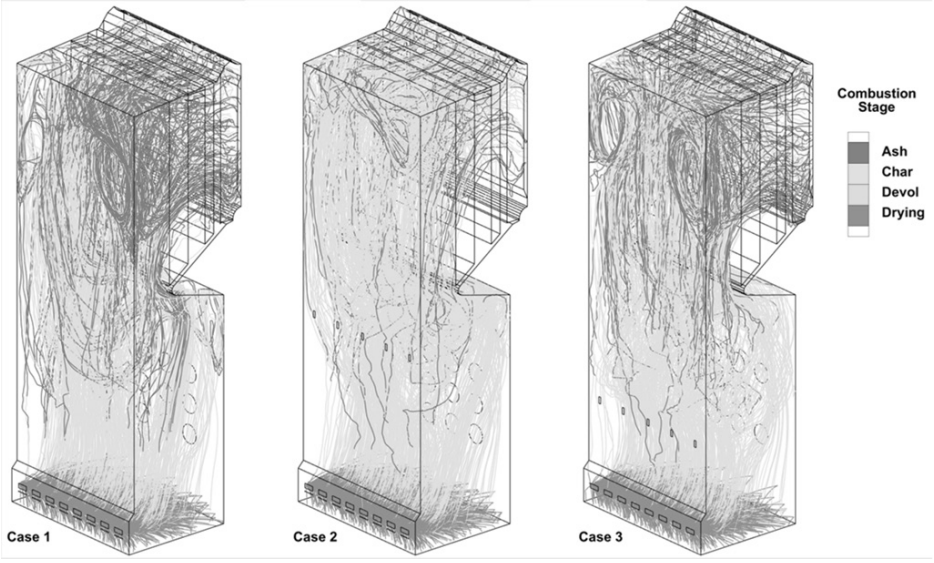


Figure 10. Visualization of Complex Flow Phenomena Modeled by CFD, Analogous to Flow in a Heat Exchanger [41].

4.4 Result Interpretation

Essential principles of fluid dynamics can be used to explain why Scenario 2 performed superiorly. The baffle modification design works well to minimize the leakage and bypass streams in the shell to push a greater proportion of the fluid across the tube bundle and in a turbulent flow. This greater turbulence breaks the thermal boundary layer on the outer sides of the tubes resulting in high convective heat transfer coefficient. The temperature plots of Scenario 2 (Figure 11) indicate that the temperature distribution in the shell side is more uniform, and the recirculation areas are less pronounced and reduced on the whole than the baseline, which proves the enhanced thermal mixing.

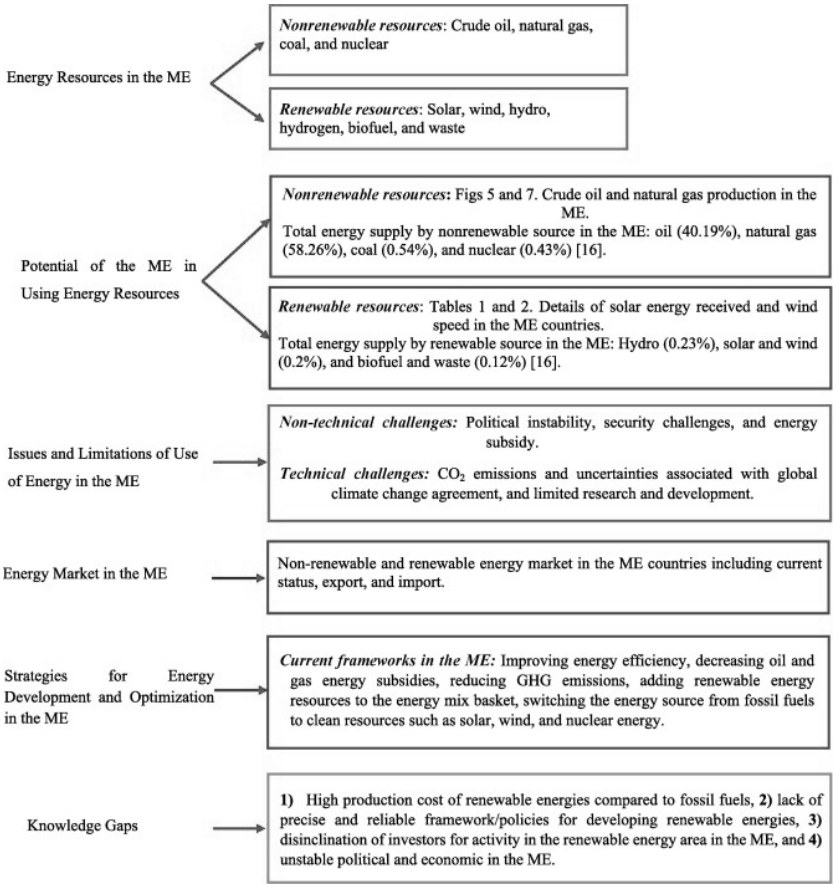


Figure 11. Temperature Distribution for Optimized Geometry (Scenario 2).

The practical implication of such an Iraqi plant as Al-Dora is huge. The result of an 18% rise in the heat transfer coefficient is that the same heat transfer could be performed with a smaller and hence cheaper heat exchanger. In case of an existing unit it translates to recovery of more heat with a given waste stream which has a direct effect of enhancing the overall thermal efficiency of the plant and the amount of fuel that is required per megawatt-hour of electricity generated. This would directly solve the problem of high fuel consumption and the low efficiency that is found in research of power-generating plants in Iraq [42], [43]. These findings reveal that substantial physical changes can be made with comparatively easy geometry changes, which can be tested and developed through CFD and then physically retrofitted without undergoing physical retrofitting.

4.5 Result Limitations

Although the findings are encouraging, it is necessary to note the drawbacks of this research. The simulations were performed with steady-state conditions and failed to consider the fouling effects which may severely deteriorate the performance of heat exchanger in the long run due to introducing a thermal resistance layer. The geometric model, though realistic, is an approximation of a real world unit, and it fails to capture complexities, like connections of nozzles, support structures, and other complexities, that can affect local flow patterns. To gain even greater fidelity, it may be necessary to use a Large Eddy Simulation (LES) so that turbulent structures are resolved to a greater degree of accuracy than is possible with the RANS-based k-o SST model, at a significantly higher computational cost. The future development of work is anticipated to include transient effects and fouling models and possibly experience validation on a full scale to develop on the results of this numerical study.

5. Conclusions and Recommendations

5.1 Conclusions

The study was able to illustrate the use of Computational Fluid Dynamics through ANSYS Fluent in order to streamline one of the main thermal energy conversion processes in the environment of an Iraqi power plant. The research offers a strong confirmation of the hypotheses used and attains its major goals. The main findings are as follows:

1. **CFD as an Effective Optimization Tool**: The creation and test of a 3D CFD model of a shell-and-tube heat exchanger turned out to be a most effective way of analyzing and optimizing its thermo-hydraulic performance. The baseline case simulation outcome correctly pointed to the presence of main areas of inefficiency (stagnations in flows and recirculation areas) that occur in old industrial equipment.

2. **Best Geometric Modification**: The most noticeable improvement in thermal performance of the three optimization scenarios used was the internal geometry modification (Scenario 2) through the adjustment of the baffle setting. The change caused 18% and 12.4% improvement of the shell-side heat transfer coefficient and total heat duty, respectively, which greatly exceeds the advantages of raising the flow velocity or altering materials.

3. **Measurable Effect on Plant Performance**: The outcomes of CFD-informed optimization are directly proportional and measurable effects on the performance of the plant. The increased heat recovery in Scenario 2 corresponds to an approximation of 2.7 percent decrease in fuel burn in a typical unit of the Al-Dora Thermal Power Plant which justifies the possibility of significant savings in the economy and decrease in emissions.

4. **Correlation with the Industrial Expectations**: The observed numerical findings and performance trends of this study are correlated with the basic heat

transfer theory and those of international industrial cases. This proves that the CFD methodology is sound and that the performance of the Iraqi industrial facilities can be made more in line with the global standards with the help of the application of the modern engineering simulation tools.

5.2 Recommendations

According to the findings of the present work, the Iraqi energy sector should be recommended as follows:

1. **Introduce CFD-Based Optimization to Iraqi Plants**: The Ministries of Electricity and Oil must be the driving force behind the introduction of CFD simulation as a standard method of troubleshooting poly thermal systems, retrofitting and designing thermal systems in power plants and refineries. It should form a special engineering division with a set of tools such as ANSYS Fluent to critically analyze and streamline the key elements throughout the industry.

2. **Preference to Geometric Retrofitting of Heat Exchangers**: Following the high returns observed, a national project should be launched to evaluate and retrofitt internal geometries of older heat exchangers especially heat recovery. The approach that is outlined in this paper may act as a pattern of such projects.

3. **Conduct Full-Scale Experimental validation**: To develop additional confidence and the refinement of the simulation models, it is important to undertake the validation of the full-scale experiments on a retrofitted unit at an Iraqi plant. This would include instrumenting a heat exchanger to measure the temperatures, pressures and flow rates to compare them directly with CFD predictions.

4. **Improve Iraqi Engineering Capacity**: The investment that should be made is the training and capacity building of the Iraqi engineers in the improved numerical simulation tools. These will require collaboration with foreign universities and

engineering companies and integration of CFD in the curriculum of local universities to form a sustainable and in-country base of expertise.

5.3 Future Work

This study gives a number of directions on the future research on converting thermal energy in Iraq:

- **Future research**: The future research should conduct a coupled Fluid-Structure Interaction (FSI) analysis to examine the vibrations caused by the flow on the changed designs of the heat exchanger to ensure structural integrity in the long term.
- **Applications in nanofluid Nanofluid**: as a working fluid has been demonstrated to have significant potential in improving heat transfer. The optimized design of heat exchanger should be investigated in future simulation with respect to different nanofluids (e.g. Al2O3-water) in the conditions of operating in Iran.
- **Broader System-Level Modeling**: The optimization strategy must be expanded to other important components and ultimately to the whole thermal cycle of such plants as Al-Dora. This may include a digital twin of the power plant to have the holistic performance management.
- **Extension to Other Plants**: It should be extended to other Iraqi power plants, including gas-fired power plants at Zubair and Hartha or the thermal power plant in Al-Mussaib, which is facing individual operational challenges and optimization opportunities.

6. References

[1] Haider Obaidi, (2025), "Iraqi buildings consume high-level electrical energy for air conditioning purposes to provide the standard human comfort condition," *Periodicals of Engineering and Natural Sciences*, vol. 16, no. 1, p. 60.

[2] S. M. Abbas, (2025), "Iraq currently faces a significant electricity deficit estimated at approximately 18,000 MW," *Biofuels, Bioproducts and Biorefining*.

[3] Shafaq News, (2025), "Iraq's power plants: A health and environmental hazard," Feb. 8. [Online]. Available: https://shafaq.com/en/Iraq/Iraq-s-power-plants-A-health-and-environmental-hazard

[4] H. A. Majed, (2025), "Al-Mussaib Thermal Power Plant faces critical challenges of excessive fuel consumption," *Iraqi Journal of Oil and Gas*.

[5] Baker Institute for Public Policy, (2025), "Iraq's Electricity Shortage and the Paradox of Gas Flaring," Jun. 17. [Online]. Available: https://www.bakerinstitute.org/research/iraqs-electricity-shortage-and-paradox-gas-flaring

[6] Frank M. White, (2016), *Fluid Mechanics*, 8th ed. New York: McGraw-Hill Education.

[7] Mechartes, (2025), "Computational Fluid Dynamics (CFD) analysis." [Online]. Available: https://mechartes.com/cfd-engineering-services/

[8] Ansys Innovation Space, (2025), "Introduction to Heat Transfer Modeling in Ansys Fluent." [Online]. Available: https://innovationspace.ansys.com/courses/courses/heat-transfer-modeling-in-ansys-fluent/lessons/introduction-to-heat-transfer-modeling-in-ansys-fluent-lesson-1/

[9] M. Abrofarakh, A. Mohammadi, M. Rezaei, and H. Karimi, (2025), "A novel approach to enhance hollow-fiber direct contact membrane distillation (DCMD) by integrating copper metal foam," *Scientific Reports*.

[10] S. A. Tazikeh, M. Rezaei, H. Karimi, and A. Mohammadi, (2025), "A comprehensive review of non-renewable and renewable energy resources and their current status and future prospects in the Middle East," *Energy Reports*.

[11] GEM Wiki, (2025), "Al Dora power plant," Aug. 5. [Online]. Available: https://www.gem.wiki/Al_Dora_power_plant

[12] Saif H. Faisal, Ahmed A. Khadom, and Ahmed H. N. Khalifa, (2021), "Energy and exergy analyses of a gas turbine unit in Shatt Al-Basra power plant," *IOP Conference Series: Materials Science and Engineering*, vol. 1773, no. 1, p. 012020.

[13] N. S. Mahmood and T. A. Mohammed, (2022), "Field observations and numerical simulation of thermal pollution in Tigris River at Al-Dora power plant site," *Journal of Engineering*, vol. 28, no. 9, pp. 19–32.

[14] Japan International Cooperation Agency, (2012), "Baiji Refinery Upgrading Project," JICA Open Report. [Online]. Available: https://openjicareport.jica.go.jp/pdf/11989019.pdf

[15] M. S. Nasser, A. A. Hassan, H. J. Kareem, and R. T. Salman, (2025), "Assessment of Water Quality at Rumaila Combined-Cycle Power Plant: Impacts, Challenges, and Management Strategies," *Basrah Journal of Science*, vol. 43, no. 1, pp. 155–173.

[16] United Nations Framework Convention on Climate Change, (2021), *Iraq's Second Nationally Determined Contribution (NDC).* [Online]. Available: https://unfccc.int/sites/default/files/resource/IRAQ%20SNCBUR%20-ENGLISH.pdf

[17] Safaa Khalaf, (2025), "Iraq's Green Transition: Policy of Contradiction and Denial Boosts More Oil and Gas Production," *Safaakhalaf.com*, Jul. 27. [Online]. Available: https://safaakhalaf.com/2025/07/27/iraqs-green-transition-policy-of-contradiction-and-denial-boosts-more-oil-and-gas-production/

[18] Yunus A. Cengel and Michael A. Boles, (2019), *Thermodynamics: An Engineering Approach*, 9th ed. New York: McGraw-Hill Education.

[19] Energy & Sustainability Directory, (2025), "Thermal Conversion Efficiency," May 3. [Online]. Available: https://energy.sustainability-directory.com/term/thermal-conversion-efficiency/

[20] Theodore L. Bergman, Adrienne S. Lavine, Frank P. Incropera, and David P. DeWitt, (2017), *Fundamentals of Heat and Mass Transfer*, 8th ed. Hoboken, NJ: Wiley.

[21] MR-CFD, (2024), "Conjugate Heat Transfer (CHT) Simulation: A Definitive Guide." [Online]. Available: https://www.mr-cfd.com/conjugate-heat-transfer-cht-simulation/

[22] Enertherm Engineering, (2025), "Mastering Shell and Tube Heat Exchanger Thermal Design Calculations," Sep. 30. [Online]. Available: https://enertherm-engineering.com/mastering-shell-and-tube-heat-exchanger-thermal-design-calculations/

[23] Boostrand, (2025), "Understand Thermal Design Aspects of Shell & Tube Heat Exchangers," Sep. 12. [Online]. Available: https://boostrand.com/understand-thermal-design-aspects-of-shell-tube-heat-exchangers/

[24] Wikipedia contributors, (2024), "Derivation of the Navier–Stokes equations." [Online]. Available: https://en.wikipedia.org/wiki/Derivation_of_the_Navier%E2%80%93Stokes_equations

[25] NASA Glenn Research Center, (2024), "Navier-Stokes Equation," Jul. 19. [Online]. Available: https://www1.grc.nasa.gov/beginners-guide-to-aeronautics/navier-strokes-equation/

[26] Engineering.com, (2016), "Choosing the Right Turbulence Model for Your CFD Simulation," Nov. 22. [Online]. Available: https://www.engineering.com/choosing-the-right-turbulence-model-for-your-cfd-simulation/

[27] CFD-Online, (2012), "k-epsilon vs k-omega turbulence modelling," Jan. 6. [Online]. Available: https://www.cfd-online.com/Forums/fluent/95885-k-epsilon-vs-k-omega-turbulence-modelling.html

[28] COMSOL Blog, (2017), "Which Turbulence Model Should I Choose for My CFD Application?" Jul. 6. [Online]. Available: https://www.comsol.com/blogs/which-turbulence-model-should-choose-cfd-application

[29] SimScale Documentation, (2025), "The k-omega SST Model." [Online]. Available: https://www.simscale.com/docs/simulation-setup/global-settings/k-omega-sst/

[30] Reddit r/CFD, (2024), "k epsilon vs k omega," Mar. 22. [Online]. Available: https://www.reddit.com/r/CFD/comments/1bl8pnl/k_epsilon_vs_k_omega/

[31] COMSOL Blog, (2018), "Your Guide to Meshing Techniques for Efficient CFD Modeling," Jun. 13. [Online]. Available: https://www.comsol.com/blogs/your-guide-to-meshing-techniques-for-efficient-cfd-modeling/

[32] SimScale, (2022), "Mesh Sensitivity Study in CFD," Apr. 26. [Online]. Available: https://www.simscale.com/knowledge-base/mesh-sensitivity-cfd/

[33] ANSYS Fluent Theory Guide, (2024), "Pressure-Velocity Coupling." [Online]. Available: https://www.afs.enea.it/project/neptunius/docs/fluent/html/th/node373.htm

[34] Ansys Academic, (2025), "Convection and Conduction with ANSYS Fluent." [Online]. Available: https://www.ansys.com/academic/educators/education-resources/convection-conduction-with-ansys-fluent

[35] Ansys Advantage Magazine, (2021), "Use CFD to Improve Efficiency and Reduce Emissions in Fired Heaters." [Online]. Available: https://www.ansys.com/advantage-magazine/volume-xv-issue-1-2021/use-cfd-improve-efficiency-reduce-emissions

[36] Neural Concept, (2024), "Conjugate Heat Transfer: Best Practices & Applications." [Online]. Available: https://www.neuralconcept.com/post/conjugated-heat-transfer-best-practices-applications

[37] Ansys Innovation Space, (2025), "How to Setup Thermal Boundaries and Heat Sources in Ansys Fluent." [Online]. Available: https://innovationspace.ansys.com/courses/courses/heat-transfer-modeling-in-ansys-fluent/lessons/how-to-setup-thermal-boundaries-and-heat-sources-in-ansys-fluent-lesson-2/

[38] Chamil Abeykoon, (2020), "Design and analysis of a shell and tube heat exchanger," *International Journal of Heat and Mass Transfer*.

[39] E. Ozden and I. Tari, (2010), "Shell side CFD analysis of a small shell-and-tube heat exchanger," *International Journal of Thermal Sciences*.

[40] Becht Blog, (2024), "CFD Modeling of a Stoker Boiler," Jun. 3. [Online]. Available: https://becht.com/becht-blog/entry/cfd-modeling-of-a-stoker-boiler/

[41] H. A. Majed, S. A. Abbas, and A. M. Al-Mousawi, (2025), "Evaluation of Carbon Dioxide Emissions and Thermal Efficiency of Al-Mussaib Thermal Power Plant," *Iraqi Journal of Oil and Gas*, vol. 14, no. 1.

[42] O. R. Alomar, (2024), "Performance analysis of a steam power plant in Mosul, Iraq," *Case Studies in Thermal Engineering*, vol. 56.

[43] M. R. Alsabbagh, (2026), "Concentration of Black Carbon Emissions From Natural Gas Flaring Activities in Basra Province, Iraq," *Basrah Journal for Engineering Sciences*.

[44] Power Technology, (2024), "Al Dora Thermal Power Station," Oct. 21. [Online]. Available: https://www.power-technology.com/data-insights/power-plant-profile-al-dora-thermal-power-station-iraq/

[45] E. J. Mahdi, A. H. Kareem, M. S. Nasser, and R. T. Salman, (2025), "Experimental comparison of aluminum (Al) and iron (Fe) radiators in a closed-cycle SWHS under Iraqi conditions," *Energy Reports*.

Comparative Adsorptive Performance of Clinoptilolite, Purolite, and Mesoporous Carbon (CMK-3) for Heavy Metals and Hydrocarbons Removal from Real Petrochemical effluent / Original Article

Assist. Lact. Eynas Muhamad Majeed[1]*,

Assist. Lact. Nuha Muhsen Ali[2]

12 , Middle Technical University, Polytechnic College of Engineering Specializations, Baghdad/Iraq

*** Corresponding Author: inas.mohammed@mtu.edu.iq**

مقارنة أداء الامتزاز للكلينوبتيلوليت والبورووليت والكربون المسامي (CMK-3) لإزالة المعادن الثقيلة والهيدروكربونات من المخلفات البتروكيماوية الحقيقية

م.م ايناس محمد مجيد[1]، م.م نهى محسن علي[2]

2،1 الجامعة التقنية الوسطى، كلية البوليتكنك للتخصصات الهندسية-بغداد، بغداد\العراق

Abstract

Petrochemical activities result in the generation of effluent that contain highly toxic heavy metals and oil products, which are very hazardous to the environment. The research tries to examine the adsorption capacity of three materials synthetic clinoptilolite, synthetic Purolite, and synthesized mesoporous carbon (CMK-3) in the removal of Ni^{2+}, Cd^{2+}, Pb^{2+}, benzene, and toluene present in real refinery effluent obtained at Al-Dora Refinery in Iraq (Baghdad). The experiments were carried by batch adsorption that investigated the equilibrium behavior, kinetic performance, and the influence of ionic strength and organic co-contaminants. The heavy metal adsorption was higher in Clinoptilolite and Purolite reaching an efficiency of more than 90%. Although CMK-3 is not a good metal remover, it proved to possess high toluene and benzene removal capacity because of high surface area and mesoporous structure. Co-contaminant experiments indicated that benzene increased Pb^{2+} adsorption which could be as a result of Pbbenzene complexation and Ni^{2+} and Cd^{2+} adsorption decreased. Hydrocarbons were characterized by a two-stage isotherm model pointing to multilayer adsorption behaviour. The paper offers comparative perspectives on adsorbent performance with actual petrochemical discharge and suggests that new ways of adsorption amplification or inhibition can be offered under co-contaminant conditions.

Keywords: Mesoporous materials, Mesoporous Carbon, Clinoptilolites, Purolites, Adsorption, Heavy metals, hydrocarbon.

المستخلص

تؤدي الأنشطة البتروكيميائية إلى توليد نفايات سائلة تحتوي على معادن ثقيلة شديدة السمية ومشتقات نفطية، مما يشكل خطراً جسيماً على البيئة. يهدف هذا البحث إلى دراسة قدرة الامتزاز لثلاث مواد – الكلينوبتيلوليت الاصطناعي، والبورواليت الاصطناعي، والكربون المسامي المصنّع (CMK-3) – في إزالة أيونات النيكل (Ni^{2+})، والكادميوم (Cd^{2+})، والرصاص (Pb^{2+})، والبنزين، والتولوين الموجودة في نفايات سائلة حقيقية مأخوذة من مصفاة الدورة في العراق (بغداد). أُجريت التجارب باستخدام الامتزاز الدفعي لدراسة سلوك التوازن، والأداء الحركي، وتأثير القوة الأيونية والملوثات العضوية المصاحبة. كان امتزاز المعادن الثقيلة أعلى في الكلينوبتيلوليت والبورواليت، حيث بلغت كفاءتهما أكثر من 90%. على الرغم من أن الكربون المسامي المصنّع (CMK-3) ليس فعالاً في إزالة المعادن، إلا أنه أثبت قدرة عالية على إزالة التولوين والبنزين نظراً للمساحة السطحيه الكبيرة وبنيته المسامية. أشارت تجارب الملوثات المصاحبة إلى أن البنزين يزيد من امتزاز أيونات الرصاص (Pb^{2+})، وهو ما قد يعزى إلى تكوين معقدات مع البنزين، بينما ينخفض امتزاز أيونات النيكل (Ni^{2+}) والكادميوم (Cd^{2+}). وتم توصيف الهيدروكربونات باستخدام نموذج امتزاز ثنائي المراحل، مما يشير إلى سلوك امتزاز متعدد الطبقات. تقدم هذه الورقة البحثية منظورًا مقارنًا لأداء المواد المازة مع مخلفات البتروكيماويات الفعلية، وتقترح طرقًا جديدة لتعزيز أو تثبيط الامتزاز في ظل ظروف الملوثات المصاحبة.

الكلمات المفتاحية: المواد متوسطة المسام، الكربون متوسط المسام ، كلينوبتيلوليت، بيورواليت، امتزاز، معادن ثقيلة، هيدروكربونات.

1. Introduction

The high rate of industrialization and release of harmful substances into the river systems in an uncontrolled manner has really led to water pollution crisis globally. It has been estimated that about 80 percent of the world population experience threats of water security caused by the worsening water quality [1,2]. The energy sector, mainly, the refining of oil and natural gas, as well as the extraction of gas, contribute to this theme on a large scale by creating vast quantities of produced water. This effluent is present in solid form and is usually co-produced with oil and gases in the subsurface formations, and the mixture usually consists of complex organic and inorganic effluent. The geological particularity of the reservoir, its maturity, and the technology of obtaining water affect the particular composition of produced water. Oil to water ratios range between 2 and 3 in most parts of the globe with highest records being 50 [3].

Petrochemical operations normally produce and process waters containing a broad variety of pollutants such as dissolved metals (e.g. Pb^{2+}, Ni^{2+}, Cd^{2+}), organic hydrocarbons (e.g. benzene and toluene), residual process chemicals, oils and in some systems, naturally occurring radioactive materials (NORMs). Contamination by ground water can also be through poor management of these effluent or through the leachings in contaminated soils [4]. Some treatment technologies including the pump-and-treat (P&T) systems and the permeable reactive barriers (PRBs) have been suggested to treat such contaminated waters both above as well as in situ. These structures depend on the chemical conversion, microbial decomposition, or immobilization by adsorption to degrade the number of pollutants [5].

Specifically, adsorption has been known to be efficient in the removal of organic and inorganic contaminants because it is simple, cost effective and can be adapted to different matrices of water. Heavy metals were also successfully adsorbed onto such

materials as synthetic zeolites and ion-exchange resins (Purolite) due to their strong affinity in binding divalent metal ions such as Pb^{2+}, Cd^{2+}, and Hg^{2+} [6].

These materials however tend to perform poorly in the treatment of complicated effluents that have both organic and inorganic compounds or in the event of high ionic strength like salty waters [7].

The recent trend of research has directed to the usage of mesoporous carbon adsorbents like CMK-3 and CMK-8 on the basis of their special structure, big pore volumes and high specific surface areas that may rise above 1000 m^2/g. These materials are unusually hydrophobic and are highly capable of 1-1 interactions; they are therefore well suited to the adsorption of non-polar and aromatic organic chemicals, such as benzene and toluene [8,9] Moreover, these high-temperature stable mesoporous carbons can be easily reactivated by application of mild heat or solvents, when compared to many carbands-based materials, which are unstable during acidic and/or high-temperature treatments.

Mesoporous carbon has higher performance, in terms of removal of hydrophobic pollutants in presence of co-existing metal ions, as compared to using mesoporous silica. That renders them perfect fits to deal with complicated industrial effluents, like, oil refinery effluents. Cyclic structure with ordered mesopores with templating CMK - 3 carbone makes the CMK - 3 capable of multilayer adsorption and capillary condensation, both of which lead to high adsorption of volatile and semi-volatile organic compounds [10].

In the present paper we have tested the efficiency of synthetic clinoptilolite, synthetic Purolite and mesoporous carbon (CMK - 3) in removal of heavy metals (Pb^{2+}, Cd^{2+} and Ni^{2+}) and organic pollutants (benzene and toluene) in real petrochemical effluent in Al-Dora Refinery located at Baghdad, Iraq. To study the equilibrium behavior, kinetics and the effect of the ionic strength and co-contaminant

interference, adsorption experiments were carried out. The comparative study gives mechanistic information regarding the adsorption behavior of both materials and points out the benefits of using the mesoporous carbon in the cleaning of mixed-contaminant industries effluents.

2. Materials and Methods

2.1 Materials

CMK - 3 mesoporous carbon was prepared by hard-templating of CMK - 3 with CMK - 3 as a carbon precursor according to [11]. Unreacted furfuryl alcohol was carbonized at 900 C in nitrogen atmosphere and the template was removed by etching with 10 % HF. The product material possessed a surface area of 1240 m^2 /g, 3.9 nm pore diameter, and pore volume of 1.3 cm^3/g.

Synthetic clinoptilolite and Purolite resins are purchased in the commerce. These two were cleaned with deionized water and exchanged with NaCl (2M) to make sure that uniform presence of Na^+ existed in the exchange sites.

2.2 Effluent source and characterization

The actual effluent used was a sample that was taken at the Al-Dora Refinery (Baghdad, Iraq), which was filtered to get rid of suspended solids. The first measurements of the amount of benzene and toluene were 400-600 mg/L, whereas heavy metals (Ni^{2+}, Pb^{2+}, Cd^{2+}) were 0.9-1.5 mg/L.

2.3 Batch adsorption tests

The experiments were carried out in 250 mL conical flasks with the amount of 50 mL of pollutant solution and 0.05 g of adsorbent mixed at 150 rpm temperature of 25 ^oC. Kinetics were observed at each 24 hours. The concentration of hydrocarbons

was determined on Shimadzu TOC-CSV analyzer and heavy metals with ICP-OES (Vista MPX, Varian).

2.4 Isotherm and model fitting

The adsorption data was applied on Langmuir, Freundlich, and Langmuir Freundlich models. As already proposed by [12] using mesoporous media, the two-step adsorption of benzene and toluene was fitted with a hybrid isotherm as a combination of Langmuir and Langmuir Freundlich terms.

3. Results and discussion

3.1 Kinetics of adsorption: heavy metals

Pb^2, Cd^2, and Ni^2 adsorption kinetics onto clinoptilolites, Purolites, and CMK-3 are depicted in Figures 1-3 correspondingly. Rapid adsorption kinetics and high elimination of every metal were noted. Table 1 displays the parameters derived from the nonlinear regression using the pseudo-first order model Eq. (1), each of which is accompanied by the standard error value. The regression process determined the standard error for the adjustable parameters Ce and Kobs, while the measured parameter C0 served as the reference for the instrumental precision.

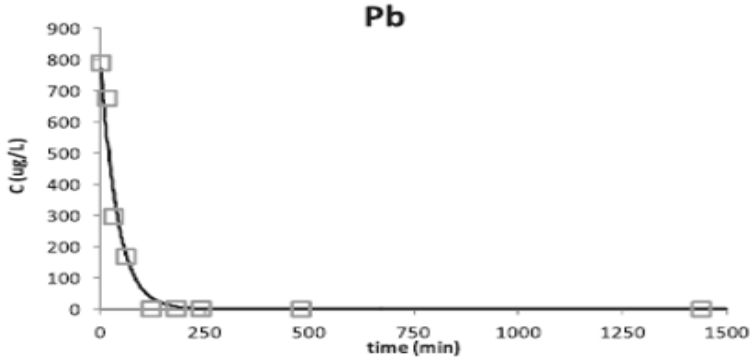
$$\frac{dC}{dC} = k_{obs}(C - C_e) \qquad \textbf{(1)}$$

For both Purolites and Clinoptilolites, the kobs value was comparable for every metal tested, indicating minor variations in each adsorption rate. When it came to removing heavy metals, the Purolites and Clinoptilolites outperformed the CMK-3. The zeolite and Purolites reached an equilibrium condition in 100 minutes, whereas the CMK-3 took 200 minutes to reach the same state. Nonetheless, as evidenced by the strong Pearson coefficient (R^2) value, the model well captured the kinetics of the chosen heavy metal adsorption process.

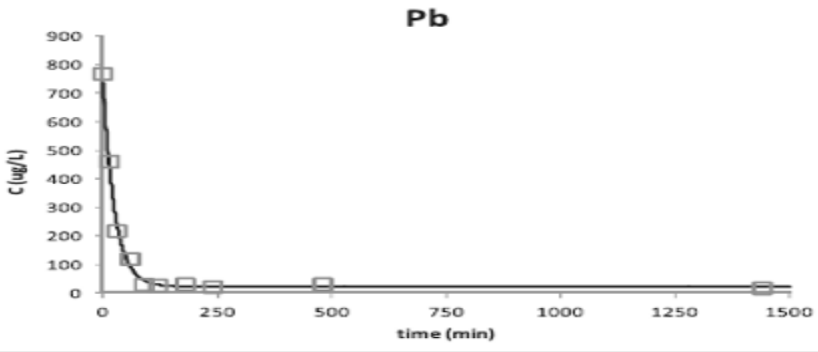
3.2 Adsorption isotherms for heavy metals: Purolites and Clinoptilolites

The target heavy metals' adsorption isotherm onto the Purolites and natural zeolite (clinoptilolites) is displayed in Figures 4-6. With the exception of Pb^2 adsorption onto clinoptilolites (which showed an inflection point in the lower concentration range), it can be observed that the suggested following model Eq. (2) fit the Purolites and clinoptilolites adsorption data quite well.

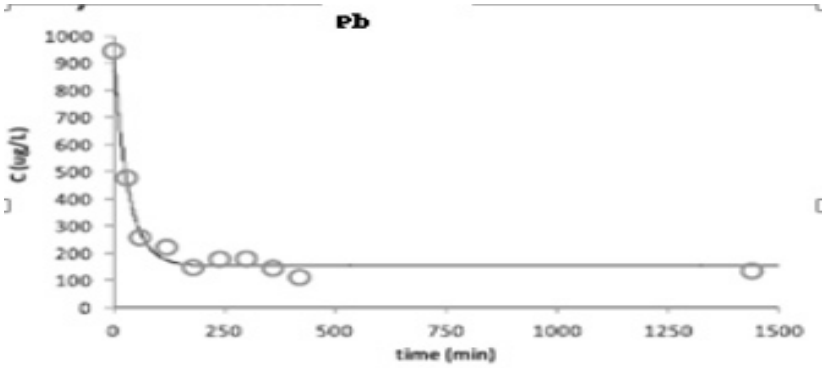
$$q_e = \frac{1}{2}\left[\frac{2K_{Na+}.q_{\max}.Ce+1}{K_{Na+}.Ce}\right] - \sqrt{\frac{4K_{Na+}.q\max.Ce+1}{(K_{Na+}.Ce)^2}} \qquad \textbf{(2)}$$



(a) Purolites

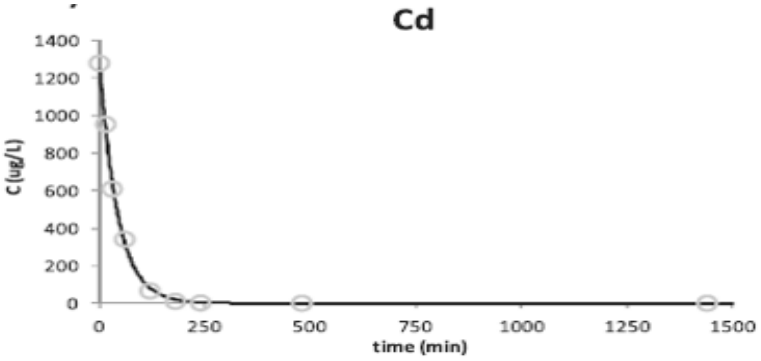


(b) Clinoptilolites

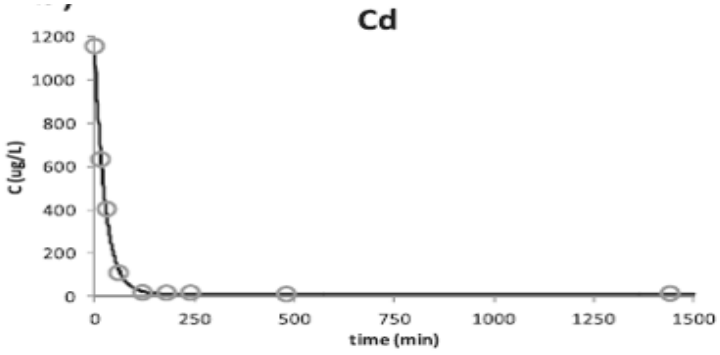


(c) CMK-3

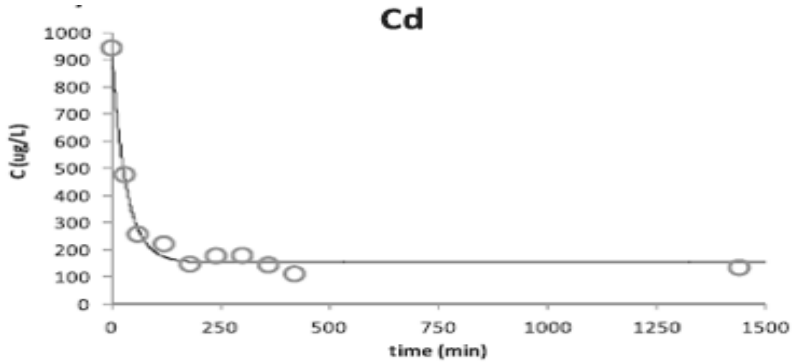
Figure 1. Pb^2 adsorption kinetics onto.



(a) Purolites



(b) Clinoptilolites



(c) CMK-3

Figure 2. The adsorption kinetics of Cd^2 onto.

The zeolite surface's heterogeneity was cited as the reason for this distinct behaviour. It is commonly recognized that, like other naturally occurring carbon, surface structures suggest that the clinoptilolites surface is distinguished by the regenerated material has both ion exchange (a permanent negative charge resulting from the isomorphic substitution of Si4 with Al3) and complexation sites (open hydroxyl active functional groups) [13]. Thus, this surface eOH may open up one additional avenue in the adsorption of metals (Pb^2, Cd^2, and Ni^2).

In connection with this, it has been demonstrated earlier that Pb^2 ions are capable of forming more stable hydroxo-complexes than Cd^2 or Ni^2 ions [14], which can perhaps partly explain the fluctuations in Pb adsorption activity. While adsorption into the complexation sites will be active for all of the metals, only Pb^2-type will show massive adsorption in low ionic strength. Since the created model (2) only involved single adsorption mechanism, it did not capture clinoptilolites Pb^2 adsorption in this case. On the other hand, the Purolites was only characterized by the presence of chelating sites, and the model did capture/ explain the behaviour of the experiments. The Langmuire-Freundlich model therefore quantified the Pb^2

adsorption behaviour, given that surface heterogeneity was incorporated by the mathematical formulation.

This is the equation that represents the Langmuir Freundlich isotherm [15]:

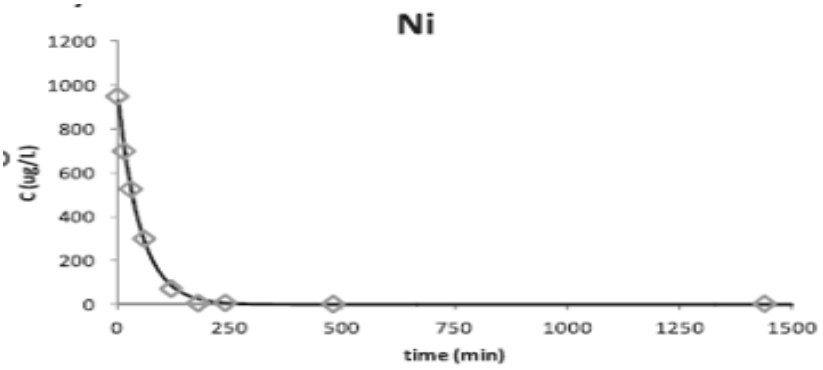
$$q_e = q_{\max} \frac{K.(Ce)^b}{1 + K.(Ce)^b} \quad \textbf{(3)}$$

where Ce is the equilibrium solution concentration in mg/L, K is the equilibrium constant in L/g, b is the maximum adsorbable amount in mg/g, qe is the quantity of Pb^2.

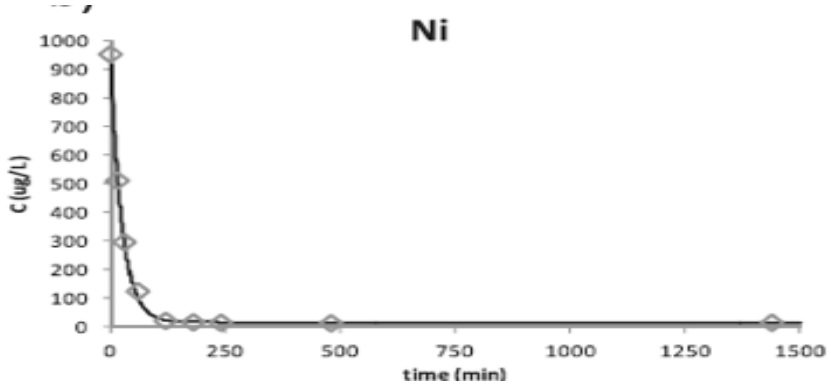
heterogeneity is a parameter that has been quantified using the symbol is. This isotherm was more applicable to account for the lead's heterogeneous adsorption process on clinoptilolites [16]. The characteristic values of the and clinoptilolites isotherms based on the optimized parameters are presented in Table2 and statistical parameters of the goodness of fit. The adsorbed quantity of Pb^2.

is a parameter for heterogeneity. This isotherm was more suited to explain the lead's heterogeneous adsorption mechanism onto clinoptilolites [17]. The optimized parameters for the and clinoptilolites isotherms are shown in Table 2, along with statistical parameters that show the goodness of fit.

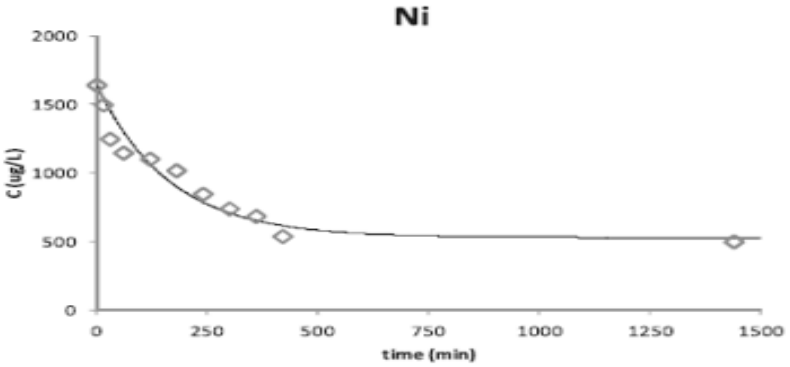
A slight lower value for the Cd^2 as compared to Ni^2 and Pb^2 the resulted depicted that there was no tremendous variation in the affinity constant between chosen heavy metals in the Purolite. These results indicated that each of the four amidoximes had a stability in relation to the heavy metals that were almost proportionate.



(a) Purolites



(b) Clinoptilolites



(c) CMK-3

Figure 3. Kinetics of Ni^2 adsorption onto.

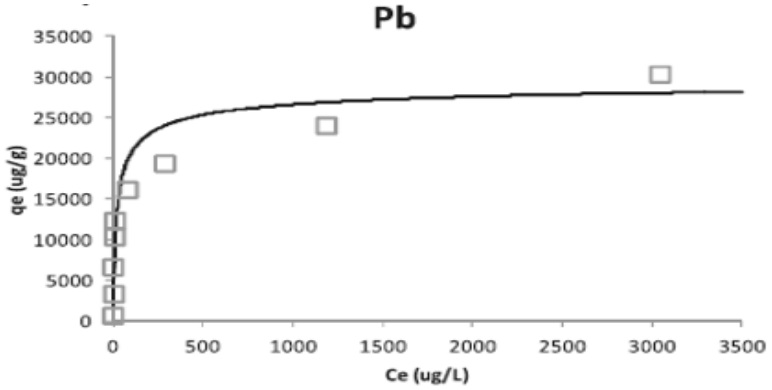
Table 1. Ni^2, Cd^2, and Pb^2 adsorption kinetic parameters were optimised for Clinoptilolites, CMK-3, and Purolites.

		Ce (mg/L)	K (min-1)	Co (mg/L)	R2
Clinoptilolites	Pb^{2+}	**11.74 ±16.90**	**0.0234±0.004**	**713±45**	**0.902**
	Cd^{2+}	**105 ± 28**	**0.0215±0.0037**	**1172±25**	**0.971**
	Ni^{2+}	**16.2 ± 3.1**	**0.0308±0.001**	**967±38**	**0.914**
Purolites	Pb^{2+}	**(2.91 ± 0.35)E-18**	**0.0251±0.003**	**763±22**	**0.981**
	Cd^{2+}	**(3.45 ± 0.38)E-7**	**0.0245±0.021**	**1053±35**	**0.95**
	Ni^{2+}	**(1.56 ± 0.6)E-18**	**0.0311±0.0045**	**942±42**	**0.904**
CMK-3	Pb^{2+}	**530 ± 312**	**0.0035±0.018**	**1378±31**	**0.917**
	Cd^{2+}	**105 ± 28**	**0.0178±0.0049**	**2082±47**	**0.995**
	Ni^{2+}	**451.1 ± 235.9**	**0.0047±0.0061**	**1835±55**	**0.962**

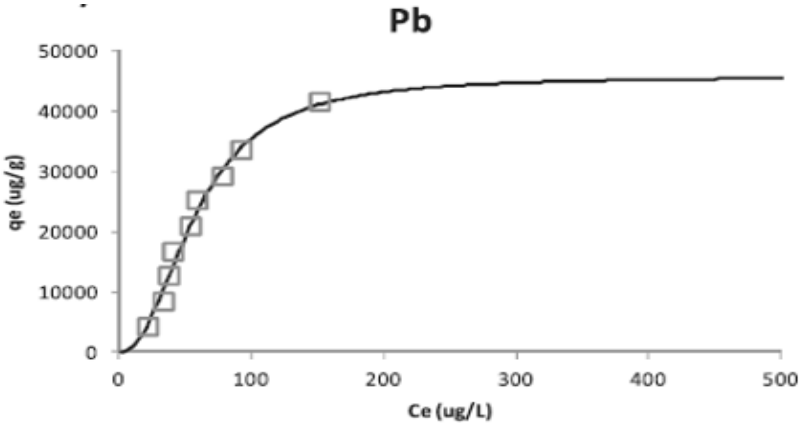
3.2.1 Ionic strength's effect

As seen from Figures 7-9, experimental isotherms of the adsorption of Ni^2, Pb^2, and Cd^2 onto Purolites and clinoptilolites.

relative to the behaviour that the computations predicted when measured with high degrees of ion strength. However, as the metal concentrations decreased, the adsorption shape was different, and all the metal isotherms in the clinoptilolites case here showed the same behaviours. If competition from the high Na concentration inhibited the adsorption of Ni^2 and Cd^2 at ion exchange sites at higher ionic strengths, the contribution of adsorption to complexation sites became relatively significant, which could account for this. These theories were in line with a better clarified Ni^2 and Cd^2 isotherm execution.

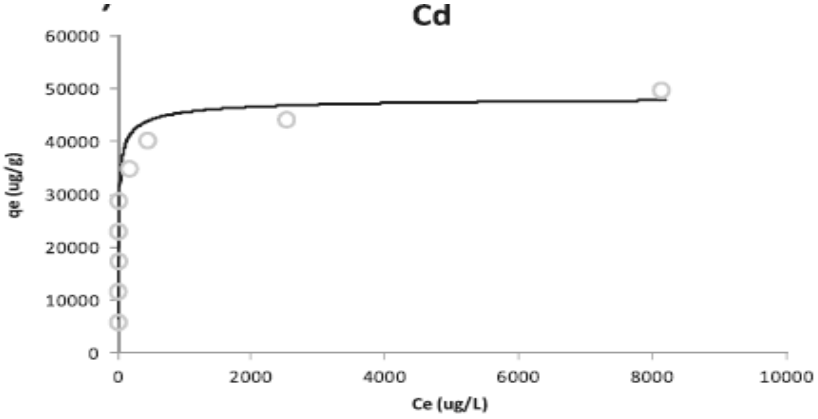


(a) Purolites

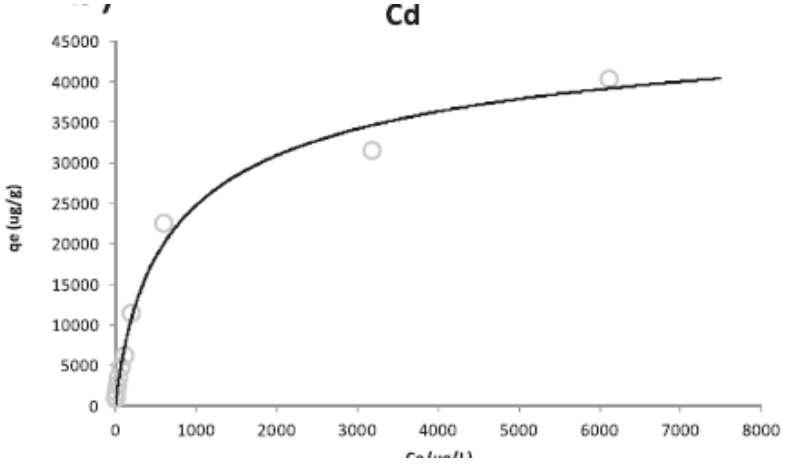


(b) Clinoptilolites

Figure 4. Pb2 adsorption isotherm onto

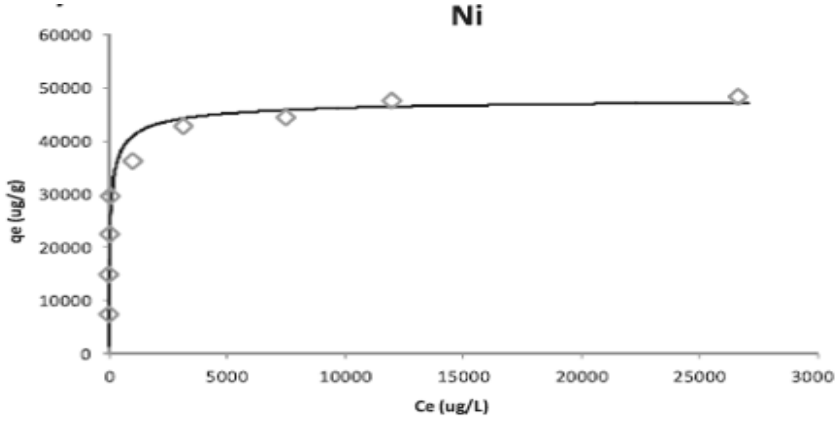


(a) Purolites

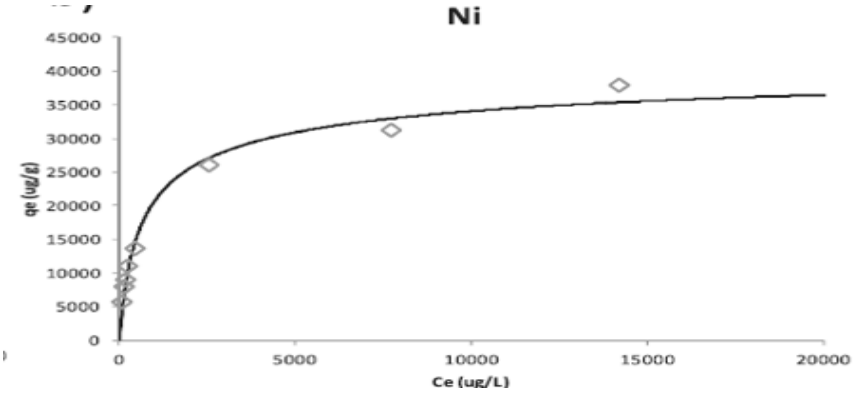


(b) Clinoptilolites

Figure 5. Cd^2 adsorption isotherm onto



(a) Purolites

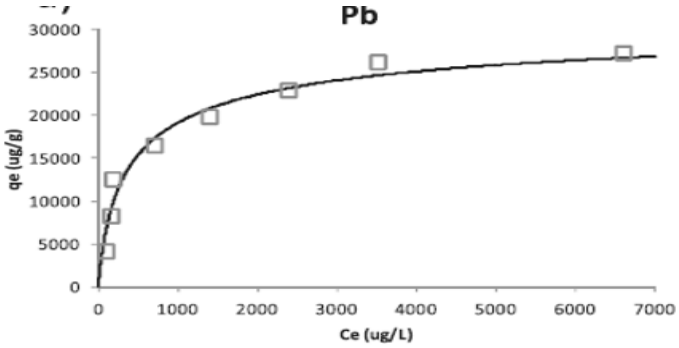


(b) Clinoptilolites

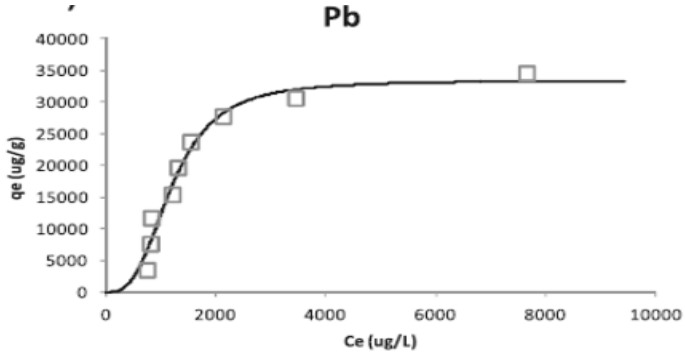
Figure 6. Comparison of the estimated and experimental behaviour for the adsorption isotherm of Ni^2 onto

Table 2. Pb^2, Cd^2, and Ni^2 adsorption onto Purolites and clinoptilolites were optimised using isotherm parameters

	Purolites			Clinoptilolites		
	Ni^2	Pb^2	Cd^2	Ni^2	Pb^2	Cd^2
qmax (ug/g)	43,000 ± 2400	61,000 ± 3100	44,000 ± 1400	50,000 ± 2200	34,000 ± 2600	30,900 ± 2100
KNa^+ or K (L/g)	(2.94 ± 0.621) E-08	0.017 ± 0.001	(3.21 ± 0.592) E-08	(2.86 ± 0.592) E-08	(4.04 ± 0.) E-08	(2.86 ± 0.592) E-08
b	--------	--------	--------	--------	2.25 0.217	--------
R^2	0.981	0.901	0.975	0.922	0.974	0.954

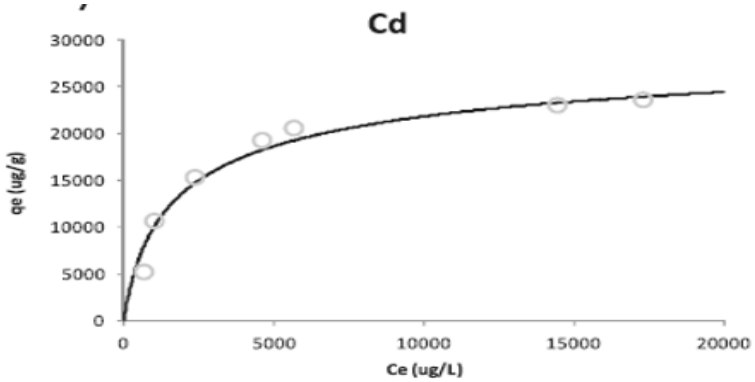


(a) Purolites

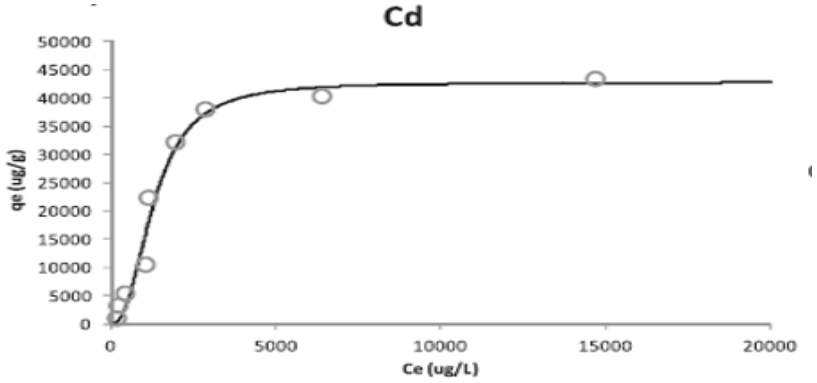


(b) clinoptilolites at high ionic strength

Figure 7. Comparison of estimated and experimental behaviour for Pb^2 isotherm onto.



(a) Purolites

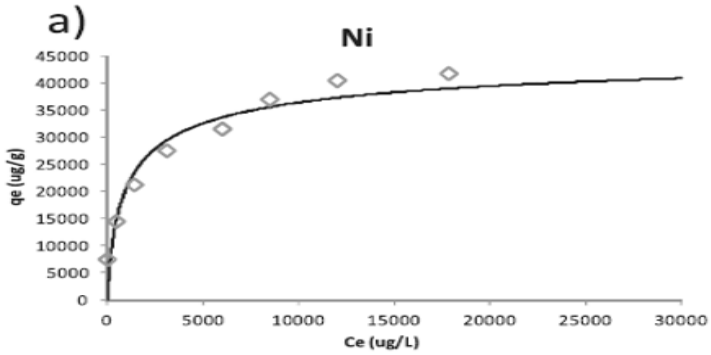


(b) clinoptilolites at high ionic strength

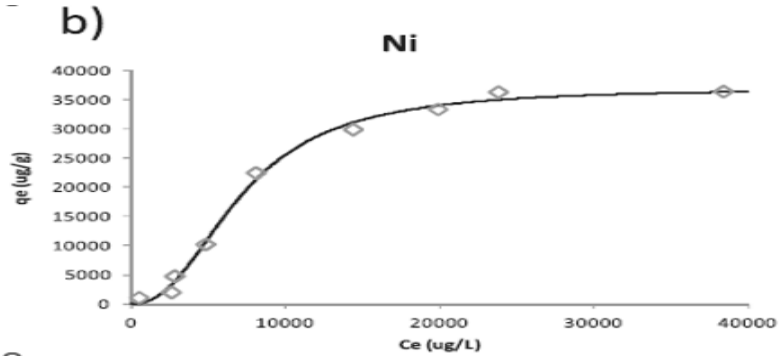
Figure 8. Comparison of predicted and experimental behaviour for the Cd^2 isotherm onto

represented by Eq. (3). In the instance of Pb^2, the optimized parameters demonstrated a drop in the affinity constant and qmax value, suggesting that Ni and Pb2 are in competition for the exchange sites. Regarding

the adsorption onto Purolites, model (2) remained appropriate for data representation, and no alterations in the adsorption behaviour were noted with regard to low ionic strength circumstances. Actually, there was just one adsorption site typology identified for the Purolites, and the ion



(a) Purolites

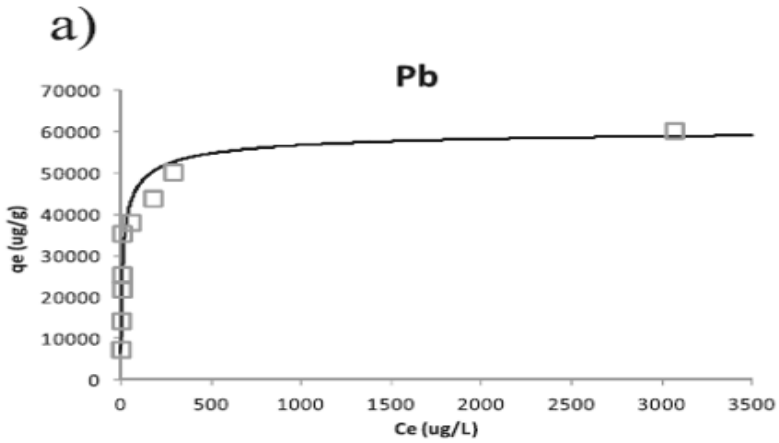


(b) clinoptilolites at high ionic strength

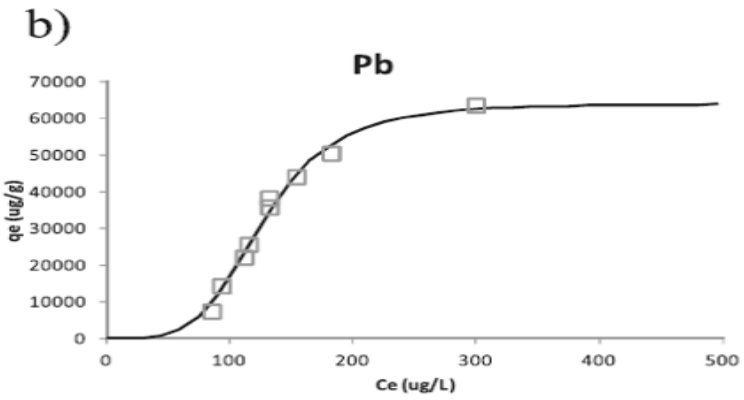
Figure 9. Comparison of predicted and experimental behaviour for the Ni^2 isotherm onto

Table 3. Ni^2, Cd^2, and Pb^2 adsorptions onto Purolites and clinoptilolite at high ionic strength were optimised using isotherm parameters.

	Purolites			Clinoptilolites		
	Ni^2	Pb^2	Cd^2	Ni^2	Pb^2	Cd^2
qmax (ug/g)	33,000 ± 1500	52,000 ± 2500	38,000 ± 1600	45,000 ± 1700	30,000 ± 2500	28,000 ± 2300
KNa^+ or K (L/g)	(2.21± 0.422) E-08	(3.50 ± 0.411) E-08	(3.082 ± 0.367) E-08	(4.32 ± 0.573) E-08	(6.53 ± 0.362) E-08	(4.05 ± 0.471) E-08
b	--------	--------	--------	2.25 ± 0.51	2.31 ± 0.49	2.12 ± 0.44
R^2	0.961	0.912	0.948	0.934	0.985	0.966

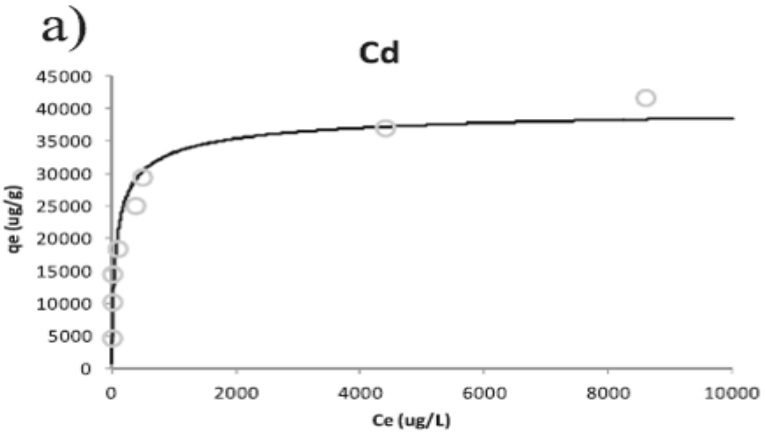


(a) Purolites

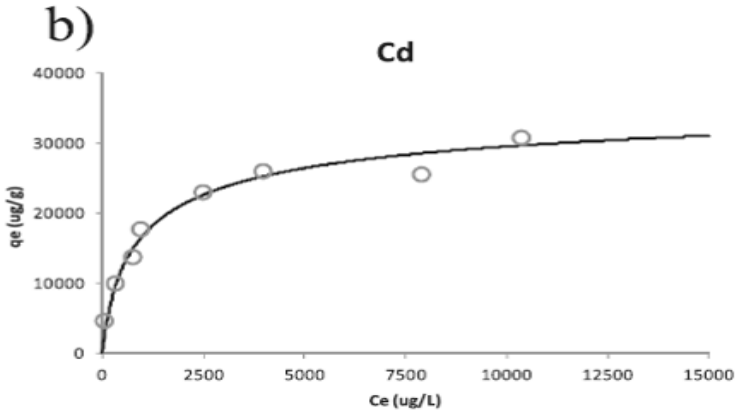


(b) Purolites with benzene present

Figure 10. Pb^2 isotherm onto



(a) Purolites



(b) Clinoptilolites with benzene present

Figure 11. Cd^2 isotherm onto

Competition only had an impact on the adsorption capacity; it had no influence on the adsorption mechanism.

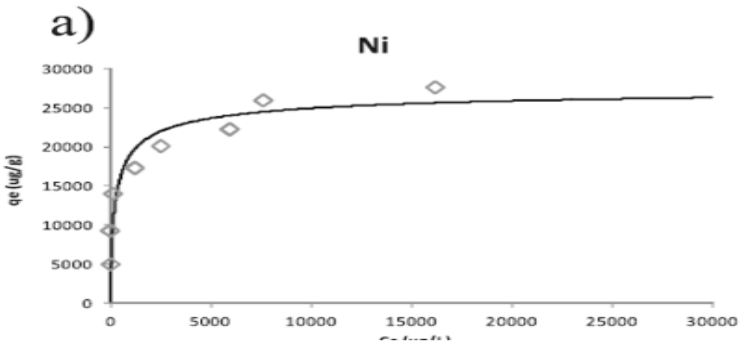
The optimized values of KNa validated these findings.

as well as qmax (Table 3). The affinity and maximum adsorbable quantity of each heavy metal were found to slightly decrease when compared to the comparable values

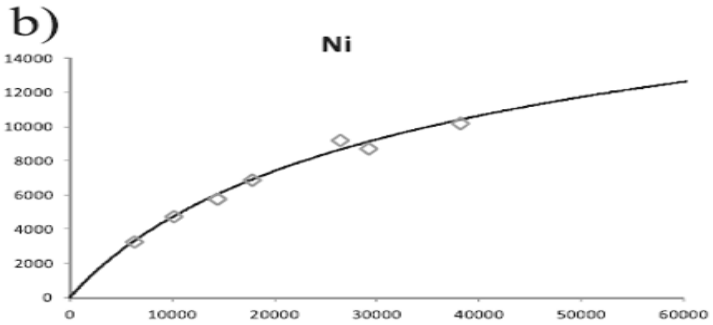
at low ionic strength (Table 2). In fact, Na's competition for chelating sites with Pb^2, Cd^2, and Ni^2 was typically restricted to causing a notable drop in KNa and qmax values.

3.2.2 Co-contamination's effects: organic compounds

Figures 10-12 at about 500 mg/L show the adsorption isotherm of each metal ion onto each adsorbent's material in presence of benzene. It appeared that the presence of benzene had varying effects on the adsorption mechanism depending on the specific ion. The maximum adsorbable amount of Pb^2 for both Purolites and Clinoptilolites increased significantly in the presence of benzene. On the other hand, Ni^2 and Cd^2 decreased the maximum



(a) Purolites



(b) Clinoptilolites when benzene is present.

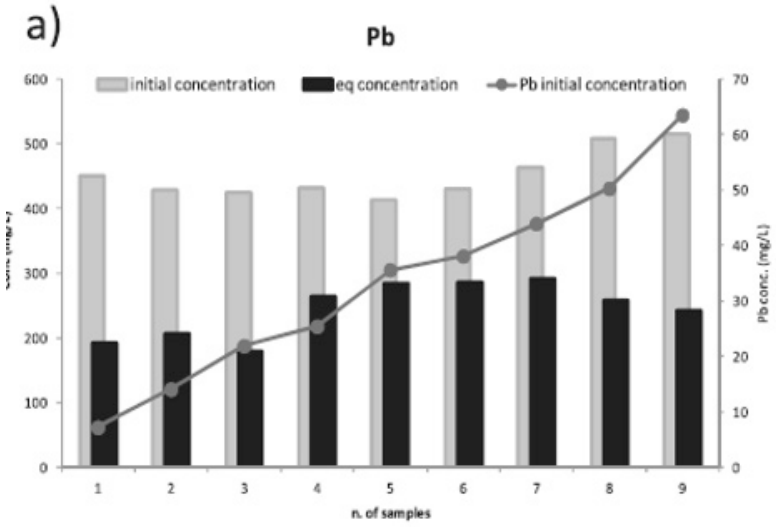
Figure 12. The Ni^2 isotherm onto.

Table 4. Ni^2, Cd^2, and Pb^2 adsorption onto Purolites and clinoptilolites with optimised isotherm parameters in the presence of benzene.

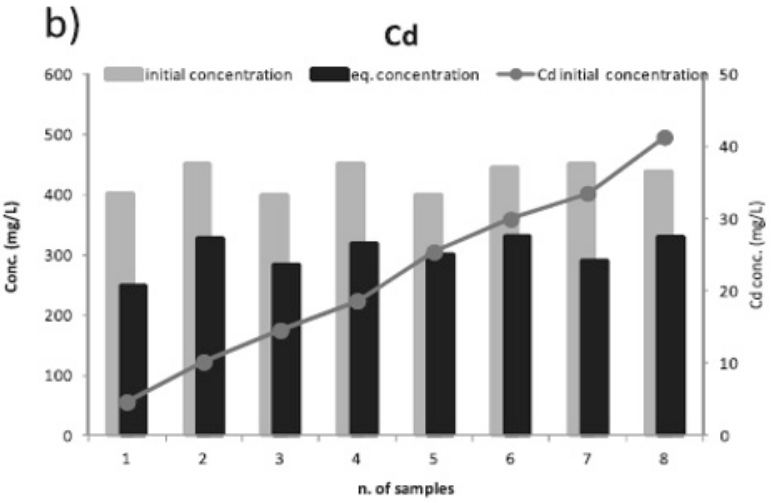
	Purolites			Clinoptilolites		
	Ni^2	Pb^2	Cd^2	Ni^2	Pb^2	Cd^2
qmax (ug/g)	43,000 ± 2500	62,000 ± 3400	29,000 ± 2300	37,000 ± 2800	64.000 ± 2600	25,000 ± 1900
KNa^+ or K (L/g)	(2.31± 2.11) E-08	(2.44 ± 3.16) E-08	(3.24 ± 1.50) E-08	(8.41 ± 2.53) E-08	0.0091 ± 0.0003	(3.72 ± 1.04) E-08
b	--------	--------	--------	--------	3.91 ± 0.57	--------
R^2	0.925	0.943	0.928	0.914	0.935	0.926

only a minor impact on the affinity constants, and an adsorbable quantity for both adsorbents (refer to Table 4).

To have a better understanding of the various ways that benzene affects metals adsorptions, benzene concentrations fluctuation during the adsorptions testing was tracked by taking measurements at the start and finish of each trial as shown in Figures13-15. Despite the fact that the experimental results were very dispersed, Pb^2 showed distinct behaviour in comparison to Cd^2 and Ni^2. Regardless of the rise in the initial metal concentration, Benzene retention in the Pb2 trials was frequently higher. According to reports.

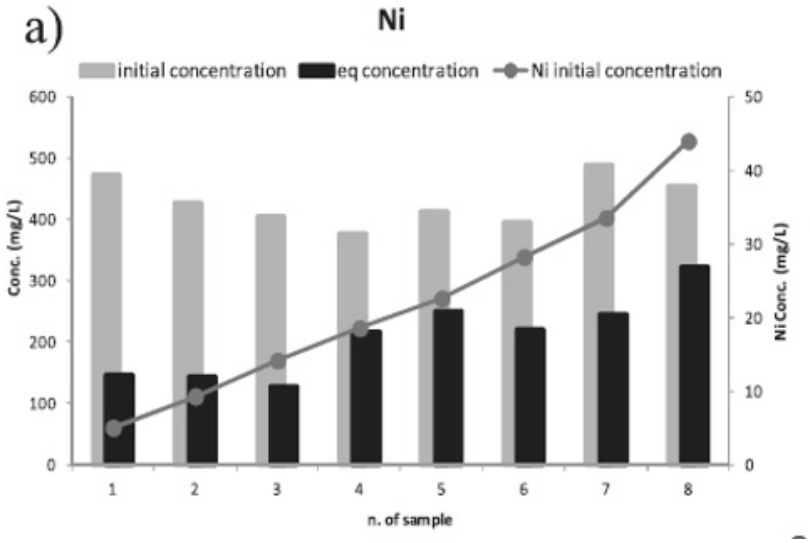


(a)

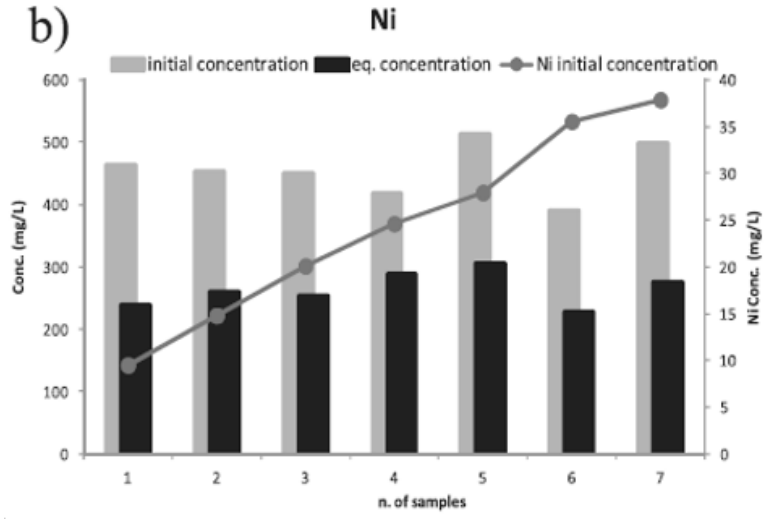


(b)

Figure 13. a) Cd^2-Purolites and, (b) Cd^2-clinoptilolites adsorption test alterations in benzene content before (grey) and after (black).



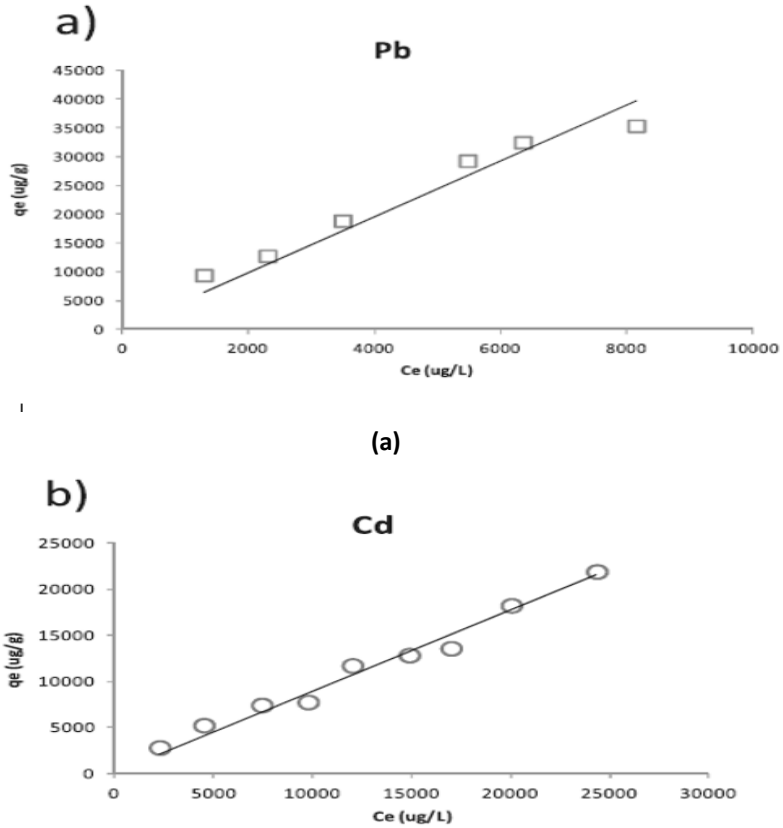
(a) Ni^2/Purolites



(b) Ni2-clinoptilolites adsorption procedure

Figure 15. Benzene concentrations variation before (grey) and after (black).

Figure 16. shows the adsorption isotherms of Pb and Cd onto CMK-3 which exhibit that lead more adsorbed onto CMK-3 than cadmium.



(a)

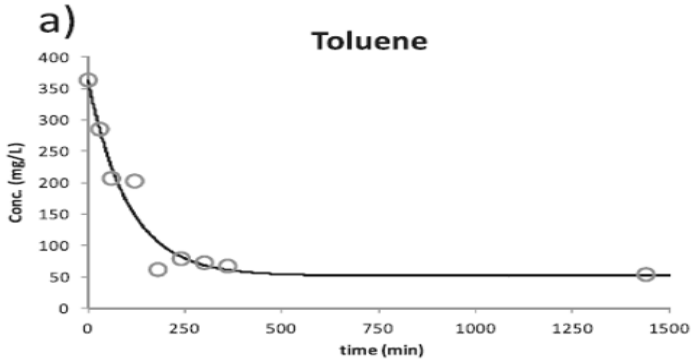
(b)

Figure 16. Cd^2 and Pb^2 adsorptions isotherms onto CMK-3.

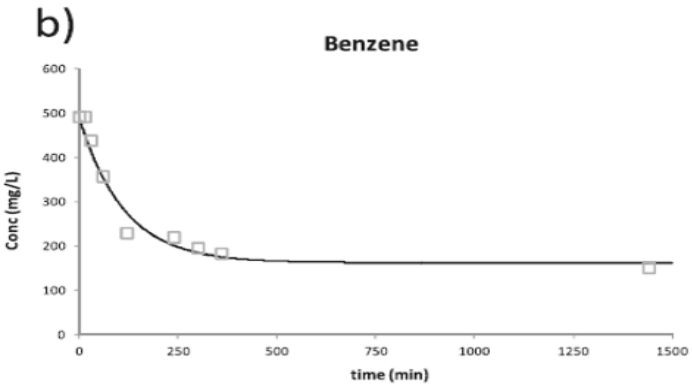
in the scholarly literature [18] Pb^2 can form stable complexes with benzene molecules, which may elucidate the observed favorable impact on its adsorption. Conversely, no information was located.

about the possible complexation of benzene with Cd^{2+} and Ni^{2+} in research. Our experimental results suggested that the simultaneous adsorption of Pb^2 ions

and benzene Pb-complex (general formula [Pb(C6H6) x]2) onto the selected adsorbents may improve the adsorption of Pb^2.



(a)



(b)

Figure 17. Benzene and Toluene adsorption kinetics onto CMK-3.

3.3 Isotherms for heavy metals adsorptions: CMK-3

Figure 16 shows the experimental results of adsorption of the target heavy metals Pb^2 and Cd^2 onto CMK-3. Different from clinoptilolites. Within the specified testing range, the adsorption isotherm for purolites showed linear behaviour and a significantly lower removal capacity. This isotherm of the Henry type is commonly used to represent the adsorption process at low solute concentrations and/or low surface affinities (and hence minimal surface coverage) [19]. Alumina's role as a binder in the CMK-3 's bulk structure may have contributed to the observed adsorption by offering more suitable adsorption sites for the removal of heavy metals. Given that the CMK-3 is possessing forms with generally low affinity, the low metal removal capability was in line with this [20]. Purolites and Clinoptilolites appeared to be better and better adapted for heavy metal uptake, and the CMK-3's efficacy in eliminating heavy metals was not comparable to theirs, even with alumina present.

3.4 Removal of hydrocarbons

3.4.1 Kinetics of adsorption: organic substances

These hydrocarbons are benzoene and toluene the most prevalent hydrocarbons and are in huge quantities in the produced goods [21]. Due to this relevance the two were used as sample hydrocarbon indicators of water contamination.

From the experimental data and using Eq (1) the estimated dependence of the kinetics of benzene and toluene adsorption on the CMK-3 layer is represented in Figure 17. In connection with Eq. (1), the optimized parameters are listed in Table 5, which results from the nonlinear regression analysis of the experimental data. The pseudo-first order kinetic equation given above was used. (1) once again provided a useful analogue for characterizing the rate of adsorption when the two hydrocarbons had stabilized at equilibrium after 6 – 7 hours. Also, upon analyzing

the exponential decay of the offered CMK-3and the benzene content after 24h contact time, the toluene decrease offered extra evidence that CMK-3 had enhanced benzene absorption capability.

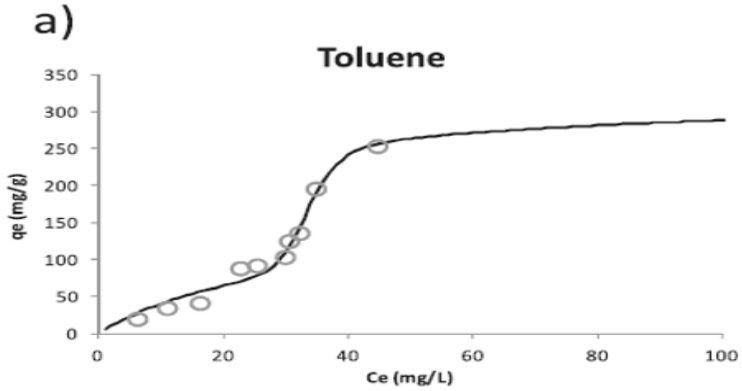
3.4.2 Adsorption isotherm for hydrocarbons: CMK-3

Experimental adsorption isotherms of toluene and benzene on CMK-3 are plotted in Figure 18. The nature of the experimental results was consistent with a two-stage adsorptive process whereby there was significant deviation from a pure Langmuir isotherm model at low concentrations, indicated by a shift in the nature of the isotherms. This type of experimental form has been noted earlier for the adsorption of organic contaminants from gaseous phase onto mesoporous materials [21] but there has been no report for adsorption of dissolved organics from solution. Hence, the results of the experimental isotherms were modeled using the hybrid isotherm model, developed for the first time in the work of [22]. The isotherm consists in the first part of the behaviour of the experimental points by the Langmuir equation and in the second part by the Langmuire Freundich isotherm.

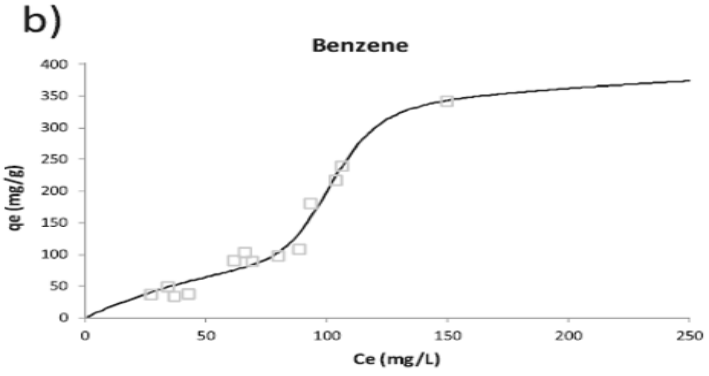
The first part of the experimental behaviour was explained by the Langmuir equation, and the second part by the Langmuire Freundlich isotherm. The expression for this model. The optimized kinetic parameters for the adsorption of benzene and toluene onto CMK-3 are presented in Table 5.

Table 5. Optimized kinetics for adsorptions of toluene and benzene onto CMK-3.

Parameters	Benzene	Toluene
C0 (mg/L)	471 ± 22	380±14
kobs (min^{-1})	0.0081 ±0.0042	0.0091 ± 0.0050
Ce (mg/L)	172 ± 11	64 ± 16
R^2	0.953	0.921



(a)



(b)

Figure 18. Toluene and benzene adsorption isotherms upon CMK-3.

$$q_e = q_{\max}.\left(\frac{K1.Ce}{1+K1.Ce}+\frac{Ks.(Ce)^n}{1+Ks.(Ce)^n}\right) \quad \textbf{(4)}$$

where Ce is the equilibrium solution concentration in mg/L, Kl is the Langmuir constant, qe is the adsorbed quantity of toluene or benzene in mg/g, and qmax is the maximum adsorbable quantity in mg/g.

Ks the Langmuire-Freundlich constant in L/g, and n is a general quantity that denotes the adsorption mechanism's heterogeneity.

The Langmuir model indicated that the first stage in the adsorptions of toluene and benzene from liquid phase onto CMK-3would be adsorption on the mesoporous channel wall until a monolayer was achieved Figure 18. Following the creation of the first monolayer, the adsorption of hydrocarbon compounds proceeded to produce subsequent layers until the entire mesoporous channel was filled. Latter procedure matched the experimental isotherms' second stage. The optimized parameters derived from nonlinear regressions of the data using the Lee model (4) are displayed in Table 6. In the Langmuir and Langmuire-Freundlich domains, toluene showed the strongest affinity (the highest values of Kl and Ks, respectively). The methyl group in toluene, which was missing from the benzene molecule, may have contributed to its higher affinity. While the benzene ring, which lacked a substituent, was free to go in any direction, the toluene molecule's ability to freely rotate in the space may be restricted by this methyl group. Benzene may have had a reduced affinity for CMK-3 due to its higher mobility. Accordingly, a more sterically hindered molecules, like toluene, adsorbed more readily than mobile molecules, like benzene [23,24]. Instead, benzene had a higher qmax value than toluene, most likely as a result of the molecule's smaller size, which let more aromatic ring to stack inside the CMK-3's pore. These findings, demonstrate how well the CMK-3 removes aromatic compounds from tainted water.

Table 6. Toluene and Benzene adsorption onto CMK-3 with optimised adsorption parameters.

Parameters	Benzene	Toluene
Ks (L/mg)	0.0368 ± 0.0004	0.0976 ± 0.0006
Kl (L/mg)	0.0076 ± 0.0027	0.0307 ± 0.0089
qmax (mg/g)	356± 27	412 ± 38
N	12.41 ± 3.92	11.25 ± 1.97
R^2	0.953	0.951

4. Conclusion

The present study has compared the adsorption capacities of synthetic clinoptilolite, synthetic Purolite, and mesoporous carbon CMK-3 as to the removal of heavy metals (Pb^{2+}, Cd^{2+}, Ni^{2+}) and organic pollutants (benzene and toluene) in real petrochemical effluent sample taken in the Al-Dora Refinery of Baghdad, Iraq. Introduction of practical industrial effluent in experiment during the laboratory testing combines a more ideal measurement of the adsorbent functioning in the studious environment of complex contamination and salinity.

The kinetic findings were that clinoptilolite and Purolite had high affinity in adsorbing metal ions and they attained equilibrium within 100 minutes as opposed to CMK-3 which took a much longer period of time to reach the equilibrium. Purolite gave the maximum affinity till the metal species whereas clinoptilolite was selective to Pb^{2+}. CMK-3 is less effective on the removal of metal but its performance was excellent in the removal of hydrocarbons; more than 80 percent of both benzene and toluene were removed in 7 hours.

The studies resulted in isotherms and it was found that the Purolite and clinoptilolite adsorbed LangmuirFreundlich and that CMK-3 was of Henry-type and followed a multilayer Swedish petroleum industry metal so following a two-step hybrid isotherm was found to resemble the behavior of hydrocarbons. Significantly, it was observed that an ionic strength of 0.1mol L-1 in the Purolite was more effective than the clinoptilolite.

In concluding, it was determined that the presence of benzene increased the adsorption of Pb^2 2 to both Purolite and clinoptilolite thereby indicating the possible complexation between Pb and benzene, and decreased uptake of Cd^2 2 and Ni^2 2. This effect coupled with high affinity of CMK-3 towards aromatic hydrocarbons strengthens the significance of using materials that can effectively cope with the co-contaminated conditions.

On the whole, CMK-3 proves to be a rather good candidate of organic contaminants clean up, in particular in systems having volatile aromatics abundance, and Purolite and clinoptilolite are still excellent heavy metals remedial agents.

Future studies should explore the functionalization of mesoporous carbon materials to enhance metal adsorption, as well as the regeneration and reuse of all adsorbents in continuous flow systems.

This research provides critical insights into material selection for industrial effluent treatment and highlights the practical relevance of combining synthetic materials with real effluent testing to develop more resilient, scalable, and environmentally robust treatment technologies.

References

[1] Jeong, Y., Cui, M., Choi, J., Lee, Y., Kim, J., Son, Y., & Khim, J. (2020). Development of modified mesoporous carbon (CMK-3) for improved adsorption of bisphenol-A. *Chemosphere, 238,* 124559. https://doi.org/10.1016/j.chemosphere.2019.124559

[2] Alsawalha, M. (2019). Overview of current and future perspectives of Saudi Arabian natural clinoptilolite zeolite: A case review. *Journal of Chemistry, 2019,* 3153471. https://doi.org/10.1155/2019/3153471

[3] Phan, T. N., Gong, M. K., Thangavel, R., Lee, Y. S., & Ko, C. H. (2019). Enhanced electrochemical performance for EDLC using ordered mesoporous carbons (CMK-3 and CMK-8): Role of mesopores and mesopore structures. *Journal of Alloys and Compounds, 780,* 90–97. https://doi.org/10.1016/j.jallcom.2018.11.265

[4] Bożęcka, A. M., & Sanak-Rydlewska, S. (2018). The use of ion exchangers for removing cobalt and nickel ions from water solutions. *Archives of Mining Sciences, 63*(3), 689–704. https://doi.org/10.24425/ams.2018.124981

[5] Kalash, K. R., & Albayati, T. M. (2021). Remediation of oil refinery wastewater implementing functionalized mesoporous materials MCM-41 in batch and continuous adsorption process. *Desalination and Water Treatment, 220,* 130–141. https://doi.org/10.5004/dwt.2021.27107

[6] Li, S., Huang, L., Wang, D., Zhou, S., Sun, X., Zhao, R., … Chen, R. (2023). A review of 3D superhydrophilic porous materials for oil/water separation. *Separation and Purification Technology, 319,* 124847. https://doi.org/10.1016/j.seppur.2023.124847

[7] Favvas, E. P., Tsanaktsidis, C. G., Sapalidis, A. A., Tzilantonis, G. T., Papageorgiou, S. K., & Mitropoulos, A. C. (2016). Clinoptilolite, a natural zeolite material: Structural characterization and performance evaluation on its dehydration properties of hydrocarbon-based fuels. *Microporous and Mesoporous Materials, 225,* 385–391. https://doi.org/10.1016/j.micromeso.2016.01.036

[8] Silva, R. A., Zhang, Y., Hawboldt, K., & James, L. A. (2021). Study on iron-nickel separation using ion exchange resins with different functional groups for potential iron sub-production. *Mineral Processing and Extractive Metallurgy Review, 42*(2), 75–89. https://doi.org/10.1080/08827508.2019.1702061

[9] Huang, H., Cai, Y., Zhao, C., Chen, Z., Liao, Z., Xie, H., … Li, D. (2024). Fabrication of 3D porous materials with underliquid dual superlyophobic properties for oil–water emulsion separation and Cu2+ adsorption. *Separation and Purification Technology, 339,* 126691. https://doi.org/10.1016/j.seppur.2024.126691

[10] Wu, R., Ye, Q., Wu, K., & Dai, H. (2021). Potassium-modified ordered mesoporous carbon materials (K-CMK-3): Highly efficient adsorbents for NO adsorption at low temperatures. *Journal of Solid State Chemistry, 294,* 121844. https://doi.org/10.1016/j.jssc.2020.121844

[11] Shao, Z. D., Zhang, Q. J., Zheng, Y. M., & Cheng, X. (2024). A novel pearls-like hierarchical porous silica aerogel monolith for efficient oil/water separation. *Journal of Porous Materials, 31,* 1–11. https://doi.org/10.1007/s10934-024-01491-3

[12] Song, Y., Phipps, J., Zhu, C., & Ma, S. (2023). Porous materials for water purification. *Angewandte Chemie International Edition, 62,* e202216724. https://doi.org/10.1002/anie.202216724

[13] Kumar, A., Naidu, G., Fukuda, H., Du, F., Vigneswaran, S., Drioli, E., & Lienhard, J. H. (2021). Metals recovery from seawater desalination brines: Technologies, opportunities, and challenges. *ACS Sustainable Chemistry & Engineering, 9*(23), 7704–7712. https://doi.org/10.1021/acssuschemeng.1c01463

[14] Vorokhta, M., Morávková, J., Dopita, M., Zhigunov, A., Šlouf, M., Pilař, R., & Sazama, P. (2021). Effect of micropores on CO_2 capture in ordered mesoporous CMK-3 carbon at atmospheric pressure. *Adsorption, 27*(8), 1221–1236. https://doi.org/10.1007/s10450-021-00313-8

[15] Liou TH, Lai BC, Xu XY. High-quality Carbons Mesostructured from Korea (CMK)-type carbons with adjustable mesostructure used as potential nanoadsorbent for enhanced tannic acid capture. Desalination. 2025 Dec 3:119725.

[16] de Castro-Alves L, Yánez-Vilar S, González-Goméz MA, Garcia-Acevedo P, Arnosa-Prieto Á, Pineiro-Redondo Y, Rivas J. Understanding adsorption mechanisms and metal ion selectivity of superparamagnetic beads with mesoporous CMK-3 carbon and commercial activated carbon. Microporous and Mesoporous Materials. 2024 Jun 15;374:113159.

[17] Al-Ghouti, M. A., & Da'ana, D. A. (2020). Guidelines for the use and interpretation of adsorption isotherm models: A review. *Journal of Hazardous Materials, 393,* 122383. https://doi.org/10.1016/j.jhazmat.2020.122383

[18] Wang, J., & Guo, X. (2020). Adsorption isotherm models: Classification, physical meaning, application and solving method. *Chemosphere, 258,* 127279. https://doi.org/10.1016/j.chemosphere.2020.127279

[19] Kundu, P., & Mishra, I. M. (2018). Treatment and reclamation of hydrocarbon-bearing oily wastewater as a hazardous pollutant by different processes and technologies: A state-of-the-art review. *Reviews in Chemical Engineering, 35*(1), 73–108. https://doi.org/10.1515/revce-2017-0034

[20] Shao, J., Zhao, Y., Li, D., Xu, S., Dou, Z., Sun, Z., … Zhou, Y. (2022). Synthesis and characterization of superhydrophobic fluorinated mesoporous silica for oil/water separation. *Microporous and Mesoporous Materials, 344,* 112240. https://doi.org/10.1016/j.micromeso.2022.112240

[21] Dashtpeyma, G., & Shabanian, S. R. (2023). Efficient photocatalytic oxidative desulfurization of liquid petroleum fuels under visible-light irradiation using a novel ternary heterogeneous BiVO4-CuO/modified natural clinoptilolite zeolite. *Journal of Photochemistry and Photobiology A: Chemistry, 445,* 115024. https://doi.org/10.1016/j.jphotochem.2023.115024

[22] Dinari, S., & Eslami, F. (2020). Effect of clinoptilolite natural zeolite particles on the destabilization of the oil-in-water emulsion. *Colloid and Interface Science Communications, 37,* 100297. https://doi.org/10.1016/j.colcom.2020.100297

[23] Kiomarsipour, N., Alizadeh, M., Alizadeh, M., & Ghani, K. (2021). Synthesis and surface-functionalizing of ordered mesoporous carbon CMK-3 for removal of nitrate from aqueous solution as an effective adsorbent. *Diamond and Related Materials, 116,* 108419. https://doi.org/10.1016/j.diamond.2021.108419

[24] Liu, T., & Chen, J. (2021). Extraction and separation of heavy rare earth elements: A review. *Separation and Purification Technology, 276,* 119263. https://doi.org/10.1016/j.seppur.2021.119263

Real-Time Adaptive Control of Machining Parameters in CNC Milling using an Ensemble Digital Twin and In-Process Vibration Monitoring for Enhanced Surface Integrity / Original Article

Researcher Ali Qasim Abdulwahid

Researcher, Basra University, Basra, Iraq
e-mail: aliqasimwq1@gmail.com

التحكم التكيفي الآني في معاملات التشغيل لآلات التفريز CNC باستخدام التوأم الرقمي المجمع ورصد الاهتزازات أثناء التشغيل لتعزيز سلامة السطح وجودته

الباحث علي قاسم عبد الواحد

باحث، جامعة البصرة، جمهورية العراق

Abstract

Achieving consistent surface integrity in CNC milling under varying materials and degraded tool conditions remains a critical challenge because fixed cutting parameters often induce regenerative chatter, accelerate tool wear, and cause surface and subsurface damage. This paper presents an integrated solution that combines an Ensemble Digital Twin (EDT), real-time in-process vibration monitoring, and a hybrid adaptive controller to maintain high surface quality while preserving productivity. The EDT fuses a physics-based milling dynamics model with a recurrent LSTM surrogate through a context-aware meta-learner that weights model outputs by process state (e.g., tool wear, engagement). A compact real-time monitoring pipeline computes a Chatter Severity Index (CSI) and summary spectral/time-domain features at 40 ms cadence. These inputs feed a Mamdani-type fuzzy supervisory controller whose defuzzified outputs are tracked by tuned PID actuators to adjust spindle speed and feed per tooth with sub-100 ms round-trip latency at the edge. Methodology includes numerical simulation, data-driven model training on a multi-source dataset, and on-machine experiments on representative alloys and tooling conditions. Results demonstrate that the ensemble yields substantially improved short-horizon predictions of chatter probability and surface roughness compared with single-model baselines, and that the integrated control system reduces vibration RMS and CSI, lowers mean Ra and Rz, and virtually eliminates deleterious white-layer formation while maintaining a material removal rate close to the baseline. The proposed framework offers a practical pathway for deploying predictive, closed-loop quality control in smart machining environments, reducing scrap and rework and enabling safer, higher-yield production under challenging operating conditions.

Keywords: CNC milling; Ensemble Digital Twin; Real-time adaptive control; Chatter detection; Surface integrity

المستخلص

يُعَدُّ تحقيق سلامة سطحية متّسقة في عمليات التفريز باستخدام ماكينات CNC تحت ظروف تشغيل متغيّرة ودرجات تآكل مختلفة لأداة القطع تحديًا جوهريًا، إذ إن الاعتماد على معاملات قطع ثابتة يؤدي غالبًا إلى حدوث الاهتزازات التجددية (Chatter)، وتسريع تآكل الأداة، والتسبب بعيوب سطحية وتحت سطحية. يقدّم هذا البحث حلًا متكاملاً يجمع بين التوأم الرقمي المجمع (Ensemble Digital Twin)، والرصد الآني للاهتزازات أثناء التشغيل، ونظام تحكم تكيّفي هجين بهدف الحفاظ على جودة سطح عالية دون التأثير على الإنتاجية.

يقوم التوأم الرقمي المجمع على دمج نموذج ديناميكا التفريز الفيزيائي مع نموذج بديل يعتمد على شبكات الذاكرة طويلة وقصيرة الأمد (LSTM) عبر متعلّم سياقي يقوم بترجيح مخرجات النماذج وفقًا لحالة العملية (مثل درجة تآكل الأداة، عمق القطع، وزاوية الاشتباك). كما يتضمن النظام منظومة رصد آنية مضغوطة تقوم بحساب مؤشر شدة الاهتزاز (CSI) واستخلاص الخصائص الطيفية والزمنية كل 40 ميلي ثانية.

تُغذَّى هذه البيانات إلى متحكم إشرافي ضبابي من نوع Mamdani، حيث تتولى وحدات PID المضبوطة تتبّع مخرجاته بعد إزالة التغبّش، مما يتيح تعديل سرعة المغزل وتغذية السن الواحدة بزمن استجابة أقل من 100 ميلي ثانية في بيئة الحوسبة الطرفية (Edge). وتشمل المنهجية المحاكاة العددية، تدريب النماذج بالاعتماد على بيانات متعددة المصادر، وتجارب تشغيل فعلية على سبائك وأدوات قطع مختلفة.

أظهرت النتائج أن النظام المجمع حقق قدرة تنبؤية محسّنة بوضوح على المدى القصير لاحتمالية الاهتزاز وجودة السطح مقارنةً بالنماذج المنفردة، كما أن منظومة التحكم المتكاملة أسهمت في خفض مستوى الجذور المتوسطة للاهتزاز (RMS) ومؤشر CSI، وتقليل متوسط قيم الخشونة Ra وRz، وإزالة تقريبًا تكوّن طبقة الاحتراق البيضاء (White Layer) الضارّة مع الحفاظ على معدل إزالة المادة مقارب للمعيار الأساسي. ويوفّر هذا الإطار المقترح مسارًا عمليًا لتطبيق أنظمة ضبط جودة تنبؤية مغلقة الحلقة في بيئات التصنيع الذكية، بما يقلل معدل الهدر وإعادة التشغيل ويعزّز السلامة الإنتاجية والعائد تحت ظروف تشغيل معقدة.

الكلمات المفتاحية: التفريز CNC؛ التوأم الرقمي المجمع؛ التحكم التكيفي الآني؛ كشف الاهتزاز التجديدي؛ سلامة السطح

1. Introduction

Within the context of high-performance component production in the aerospace, biomedicine, and automotive industry domains, the achievement of high surface integrity—involving surface topographic characteristics, residual stresses, and micro-structured material modifications—is a very prominent issue. Indeed, surface integrity is a critical factor that directly conditions the basic functionality characteristics of the manufactured parts. This was well-illustrated in the reference literature and current scientific publications [1-2]. A very challenging issue in the CNC milling of challenging-to-cut materials involves the achievement of a favorable surface quality. The nature of the process in question includes being nonlinear. Additionally, regenerative chatter vibrations occur in the system. These unwanted oscillations in the process influence the dimensional accuracy of the machined surface. Additionally, the surface quality of the material being machined may deteriorate.

The traditional CNC machining approaches lack the requisite flexibility to cope with the intricacies mentioned. According to Korkmaz *et al.*, a rigid set of parameters cannot cope with the unavoidable variability that must occur in the processing conditions of CNC machines. These factors include tool wear, inhomogeneity of the material being machined, and the variation of the depth of cuts [3]. Hence, a machinist has to choose between the cautious approach to ensure stability but compromise productivity or carry out the machining under unstable conditions to jeopardize the surface quality of the machined parts. This problem accentuates the need for the development of intelligent systems that have the capability to cope in real-time.

Recent years have introduced the promising field of cyber-physical systems. The application of Digital Twins (DT) provides promising prospects. A DT refers to the electronic, simulated, and harmonized copy of a real-world system. This system

can later be simulated, analyzed, and controlled. Initial studies in the form of Melesse *et al.* show the capability of DT to optimize offline processes and predict maintenance [4]. A substantial research gap exists. Most of the existing implementations either remain offline simulation or analysis of isolated parts of the process. There exists a lack of holistic simulation in real-time for closed-loop control. There also seems to be a lack of existing frameworks utilizing the approach of ensemble modeling. This approach involves the combination of physics modeling techniques in combination with Machine Learning (ML) to produce a more robust prediction model of the machining process. Additionally, in-process vibration analysis has long been practiced for the detection of the occurrence of chatter. There also seems to exist a lack of studies exploring the adaptation of the prediction approach of the ensemble type DT for the development of a holistic system in real-time. A unified real-time adaptive control system for the achievement of peak surface integrity concerning the combination of in-process vibration analysis for the prediction of the ensemble type DT has also not been extensively introduced. Most of the existing real-time control systems remain force- or vibration-actuated in a reactive manner without the aid of predictive knowledge. There seems to exist a lack of flexibility in adapting to the possibility of past corrections in existing studies. A lack of the same also seems to exist in the existing literature of the control of the DT system in the form of Tong *et al.*[5] .

This research covers the problem by formulating, applying, and critically assessing a new real-time adaptive control system for CNC milling. The new approach relies on the combination of the EDT with in-process vibration measurement for optimized surface integrity. There are three scientific contributions of the current research. First, a new conceptual approach to the EDT structure will be presented. The approach includes a physics model of the dynamic milling process in combination

with a machine-learning (ML) modeling technique (e.g., a long-short-term-memory network) for improved vibrational conditions and surface quality prediction. Second, a new hybrid approach to adaptive control system design will be presented to real-time-optimize the cutting speed of the spindle tool and the feed rate of the material. The approach will fully utilize the combination of prediction information from the DT approach for surface quality prediction in combination with in-process-measurement information. Third, the new approach will experimentally demonstrate improved surface quality in terms of reduced surface roughness values compared to the existing static approach. Additionally, the new research approach will experimentally demonstrate improved surface quality in terms of a reduced number of surface defects compared to the existing research approach. All parts of the article contain the following subjects. Literature Review. Methodology. Conclusion. Research recommendations.

2. Literature Review

the goal in many cases would also involve the achievement of overall excellent surface integrity (SI) of the machined surface. This overall surface integrity refers to a construct of the transformed material layers developed in the material in a machined surface and also their topographical characteristics. According to the explanation given in the article by Guo *et al.* in 2022, the surface integrity of a machined material surface would be determined in relation to the surface roughness value of the machined surface (Ra and Rz values) of the machined material surface [1]. A value of the overall surface integrity would also involve the presence of the thermally altered layers of the machined material surface material in terms of the values of plastic deformation of the surface or thermally damaged surface layers, in addition to the values of the surface residual stresses. The overall values of surface

integrity would also relate to the material in-service characteristics. The in-service tensile stresses of the machined surface material would significantly reduce the fatigue life of the material surface in aerospace applications. Likewise, in many machined material surface applications in the aerospace industry, the formation of the white layers of machined material surface would significantly precipitate the overall failures, a point by all accounts well-articulated in the article Aerospace alloys in the year 2023 by the scholar 'Wang'.According to the explanation given in the article by Wang [2], the overall values of surface integrity would result from the complex relationships of factors, in particular the stability of the machined surface processing conditions of the machined material surface. These overall factors for the machined surface processing conditions would involve the dynamic process stability, and the condition of the tool. It is found that despite the ideal choice of the static parameters, the start of the wear or tiny material irregularities can significantly change the cutting mechanics that is why the SI can be rapidly destroyed by additional vibration and heating forcing [6].

To mitigate such instabilities, in-process monitoring has emerged as a vital field. A variety of sensing methodologies are employed, each with distinct advantages. Force measurement, often using dynamometers, provides a direct correlation with cutting mechanics but can be intrusive and expensive for industrial deployment [7]. Acoustic emission sensing is highly sensitive to incipient tool failure and micro-fracture events, yet its signal is often contaminated by ambient noise [3]. Among these, vibration monitoring, typically via accelerometers mounted on the spindle or workpiece, has proven particularly effective for detecting dynamic instabilities like chatter due to its non-intrusive nature and direct relationship with structural dynamics, as emphasized in the work of Nair *et al.* [8]. Building upon monitoring, Adaptive Control (AC) strategies seek to automatically adjust machining

parameters in response to sensor feedback. Traditional AC approaches, such as Constant Force Control, aim to maintain a preset cutting force level by modulating feed rate, thereby protecting the tool and improving stability, as explored by Tong *et al.* [5]. More advanced, model-based adaptive controllers utilize mathematical models of the process to predict outcomes and optimize parameters, yet their efficacy is heavily contingent on model accuracy and often struggles with the non-linearities and uncertainties inherent in machining, a challenge noted by Schmitz [9]. These systems typically operate reactively, responding to a detected deviation, which introduces a critical latency between disturbance onset and corrective action, a fundamental limitation for preserving SI in high-speed operations.

Digital Twins (DTs) paradigm provides a game-changer in the control paradigm by making it more predictive and proactive. As an extension of traditional offline simulation, a DT is characterized as a virtual, synchronized model of a physical system, which evolves and learns on real-time information throughout its lifecycle [4]. With machining, early uses of DT have been in virtual process planning, predictive maintenance and offline optimization. Indicatively, in a study conducted by Tong *et al.* [5], an advancement in milling was made by designing a DT framework simulating cutting forces and wear on tools, and optimizing the parameters prior to the practical implementation by means of milling. There is however a conspicuous gap in the closed-loop real-time control domain. Majority of machining DTs are high-fidelity and computationally heavy physics-based models that cannot be safely run in real-time or all-data-driven models that are not generalizable and not physically interpretable at all. This has stimulated the desire to have hybrid or ensemble modeling. Recent efforts by Pashmforoush *et al.* indicate the ensemble models, consisting of physics-based and data-driven machine learning (ML) units, can be predictive accuracy-wise outclassing each of them when applied to predicting

tool wear [10]. Although this has been achieved the direct use of an ensemble DT directly as the fundamental predictive driving force of real-time adaptive control, especially that has a control goal specific to SI enhancement concerns and is not simply regulating of force or avoiding of chatter is still immature and under-researched.

Milling vibrations are complex to understand as foundational to modeling them decreases the accuracy of any DT or control system to chatter mitigation. The analytical model by Mahboubkhah *et al.* is the seminal that offers stability lobe diagrams (SLD) predicting the chatter-free zone frequency domain solution [11]. This type of physics-based method has proved to be of inestimable importance, but may normally involve a lot of accurate information of tool-point frequency response capabilities and cutting force coefficients which can fluctuate during use. This has led to the popularity of data-driven approaches to chatter detection. The frequency-domain analysis is one, where a frequency occurring close to the tool natural frequency is an important indicator, as explained by Guochao *et al.* [1], and a different one, statistical time-domain techniques monitoring various indices, such as the standard deviation or root mean square of vibration signals, as investigated by Tran *et al.* [12]. These are used as alarms though they are effective when it comes to detection. The next logical but underdeveloped step is to integrate such detection metrics as inputs into a predictive DT that can be used to predict anticipation of the impending violation of SI thresholds before they happen. Recent studies by Zhou *et al.* started by applying a DT to pre-selection of parameter of a post-process of avoiding chatter, though did not proceed to in-process adaptation of closed loop [13].

Integrating the available literature shows a logical course but an apparent structural disjuncture. The use of SI metrics [1-2], vibration monitoring as a powerful chatter detector [8], adaptive control development as a constant-force scheme to a

model-based scheme [5-9] along with the improvement of the definition of the concept of a DT [4-10] are all aspects but these are still, to a great extent parallel. The scarcity of literature that integrates a predictive digital twin, of a high-fidelity and ensemble nature, to monitor surface integrity in real-time in conjunction with instant vibration monitoring within a closed-loop adaptive control framework, is lacking. The current study attempts to fill this gap by postulating and confirming a framework that places an Ensemble Digital Twin, between mechanistic dynamics and ML adaptability, which gives continuous, forward-looking predictions of SI-related outcomes. Such forecasts, supported and triggered by real-time vibration measurements, directly feed a hybrid adaptive controller which actively changes parameters away not only to reach the elusive chatter suppression, but actually to operate at optimum conditions to ensure surface finish, residual stress and micro-structure integrity, thus transforming process stabilization to active surface engineering.

3. Methodology

3.1 Overall System Architecture

The researcher framework presents an adaptive machining control, based on the real-time, closed-loop cyber-physical system. The architecture consists of three synergistic layers: the Physical Layer, the Ensemble Digital Twin (EDT) Layer and the Adaptive Control Layer. Physical Layer: The Physical Layer is comprised of CNC milling machine which will contain sensory equipment. Triaxial piezoelectric accelerator is mounted on the spindle housing near the tool interface where it measures the vibration data which is conditioned and converted to numerical data by a high-speed Data Acquisition (DAQ) system. Measurements of surface integrity are done on a stylus profilometer and metallographic preparation system after the process. The real-time sensor stream in digital format (plus current machine position

(G-code commands, axis positions), is sent on a deterministic industrial network (e.g., EtherCAT) to an edge computing station.

It is in the edge server that the EDT Layer is the brain that predicts. It takes in the new data it receives and performs parallel simulations with its two-model ensemble. The outputs, i.e., forecast of imminent chatter and the estimated surface roughness (Ra), are supplied to the Adaptive Control Layer. In this case, a hybrid controller combines these forecasts with the pure sensor-based chatter severity index. It calculates best values to changes in the main machining parameters in terms of spindle speed (N) and federate per tooth (f_z) using a rule-based fuzzy logic system. These adjustment commands are retransmitted to the CNC machine controller, and dynamically change the running tool path execution to keep the best conditions which close the loop of control in real-time (latency goal < 100 ms).

3.2 Development of the Ensemble Digital Twin (EDT)

The EDT is designed to deliver high-fidelity, mechanistic predictions that combine both mechanistic knowledge and data-driven flexibility to overcome the weaknesses of purely physics-based models or purely black-box models.

3.2.1 The Physics-Based (Interpretable) Model

This element is based on the dynamics of machining. It uses an exponentially enhanced zero-order analytical milling force model alongside a time-dependent finite element analysis to predict the stability as being the continuation of the work of Mahboubkhah et al [11]. This model uses tool geometry, cutting material properties and cutting parameters to predict subsequent vibration and cutting forces. Tangential (F_t), radial (F_r) and axial (F_a) cutting forces on tooth *j* can be given by:

$$F_{t,j}\left(\phi_j\right)=\left[K_{tc}h_j\left(\phi_j\right)+K_{te}\right]a_p \qquad F_{r,j}\left(\phi_j\right)=\left[K_{rc}h_j\left(\phi_j\right)+K_{re}\right]a_p$$
$$F_{a,j}\left(\phi_j\right)=\left[K_{ac}h_j\left(\phi_j\right)+K_{ae}\right]a_p \qquad (1)$$

where K_{tc},K_{rc},K_{ac} are specific cutting force coefficients, K_{te},K_{re},K_{ae} are specific edge force coefficients, a_p is the axial depth of cut, and $h_j\left(\phi_j\right)$ is the instantaneous chip thickness for tooth *j* at angular position ϕ_j. The chip thickness incorporates the regenerative effect due to the vibration marks left by the previous tooth, making it a function of the current and previous tool vibrations ($x\left(t\right),y\left(t\right)$). The total forces in the X and Y machine directions are obtained by summing the contributions of all teeth in the cut.

The dynamic response of the tool-workpiece system is modeled as a two-degree-of-freedom spring-mass-damper system. The equation of motion is given by:

$$\mathbf{M}\ddot{\mathbf{q}}\left(t\right)+\mathbf{C}\dot{\mathbf{q}}\left(t\right)+\mathbf{K}\mathbf{q}\left(t\right)=\mathbf{F}\left(t,\mathbf{q}\left(t\right),\mathbf{q}\left(t-\tau\right)\right) \qquad (2)$$

where $\mathbf{M}$, $\mathbf{C}$, and $\mathbf{K}$ are the mass, damping, and stiffness matrices, respectively; $\mathbf{q}\left(t\right)=[x\left(t\right),y\left(t\right)]^T$ is the displacement vector; $\mathbf{F}$ is the cutting force vector from Eq. (1), which is a function of the current and delayed states (τ is the tooth passing period). The stability of this delay-differential equation is solved in the frequency domain to generate a Stability Lobe Diagram (SLD), predicting chatter-free zones. While powerful, this model's accuracy depends on precise knowledge of frequency response functions (FRFs) and force coefficients, which can drift during operation.

3.2.2 The Data-Driven (ML) Model

Data-driven Surrogate model is created to supplement the physics model so that complex and non-linear phenomena such as progressive wear of tools and tissue inhomogeneity are captured. Since the sensor data are sequential, a Long Short-Term Memory (LSTM) Recurrent Neural Network is used, which is known to

be effective in the ability to extract time-related dependencies in the industrial time-series information [3]. The model is trained on the full dataset of the one presented above. The input feature vector for a time window t includes:

- **Static Parameters**: N , f_z , a_n, radial depth a_e, tool diameter.
- **Dynamic Time-Series Features**: A sequence of vibration signal statistics (RMS, Kurtosis) and FFT-derived features (dominant frequency amplitude, spectral entropy) from the past *k* milliseconds.

The LSTM network is trained to predict two key outputs:

1. The probability of chatter occurrence in the next time window.
2. The expected surface roughness (Ra) for the currently generated surface. The network's architecture includes two LSTM layers followed by fully connected layers, using ReLU activation and optimized with the Adam algorithm to minimize the Mean Squared Error (MSE) loss.

1.2.1 Ensemble Fusion Mechanism

The final prediction of the EDT is not a simple average but a context-aware fusion of both models' outputs using a Meta-Learner approach [1]. A Gradient Boosting Regressor (GBR) is trained as the meta-model. Its inputs are the predictions from the physics model (P_{phy}) and the LSTM model (P_{lstm}), concatenated with contextual features describing the process state (e.g., tool wear estimate, radial engagement ratio, material type). The GBR learns to assign optimal, non-linear weights to each model's prediction based on which model is historically more reliable under the given conditions. For instance, in a fresh-tool, stable-depth cut, the physics model may be heavily weighted. During high-wear conditions or when machining complex geometries, the LSTM's weight increases. The fusion output is given by:

$$P_{ensemble} = F_{GBR}\left(P_{phy}, P_{lstm}, \text{Context}\right) \qquad (3)$$

where F_{GBR} represents the meta-learner action. This joint methodology is highly useful in improving predictive strength and external applicability. The LSTM part is optimized on a periodical basis on a sliding window of the most recent machining data and post-process SI measurements, which allows the EDT to learn the detail of that machine and that tool with time.

3.3 Real-Time In-Process Monitoring System

3.3.1 Hardware Instrumentation

A triaxial IEPE accelerometer of high sensitivity (e.g., 100 mV/g) with a flat frequency response down to 12 kHz is fitted to the spindle nose. This signal is inputted to a 24-bit DAQ system with anti-aliasing filters, where it is sampled at 25.6 kHz (this is a reasonable frequency that allows reproduction of the chatter frequency to the Nyquist limit of 12.8 kHz).

3.3.2 Signal Processing Pipeline

Raw acceleration data $a_x(t), a_y(t), a_z(t)$ are real-time updated over overlapping 1024 sample (~40 ms) windows. For each window:

1. **Time-Domain Features**: The resultant vibration magnitude $a_{res}(t) = \sqrt{a_x^2 + a_y^2 + a_z^2}$
1. is computed. Root Mean Square (RMS) and Standard Deviation (s) of it are derived as total energy measures.
2. **Frequency-Domain Chatter Detection**: This stage uses a Hanning-windowed FFT to transform $a_{res}(t)$. An adjusted Chatter Severity Index (CSI) is calculated, influenced by Tran *et al.*, based on the power

distribution in the area of the most influential natural frequency of the tool (f_n) [12]:

$$\mathrm{CSI} = \frac{\int_{f_n - \ddot{A}f}^{f_n + \ddot{A}f} PSD(f)\,df}{\int_0^{f_{nyq}} PSD(f)\,df} \qquad (4)$$

of which $PSD(f)$ is the Power Spectral Density, $\ddot{A}f$ is a bandwidth of 50 Hz, and f_{nyq} is the Nyquist frequency. When CSI value goes beyond a dynamically doctored threshold (according to preferred base stable cuts), there is chatter onset. This CSI, together with RMS, comprise the short feature vector being sent to the controller every 40 ms.

3.4 Design of the Hybrid Adaptive Controller

The controller has to balance several conflicting and frequently incompatible goals: chatter suppression, quality of surface finish and productivity. A hybrid strategy of the two levels is used.

3.4.1 High-Level: Multi-Objective Fuzzy Logic Controller (FLC)

A Mamdani-type FLC will have three input and two outputs.

- **Inputs**:
    1. **Chatter Severity Index (CSI**): Linguistic variables: {Low, Medium, High}.
    2. **Predicted Ra from EDT (Ra_pred**): Linguistic variables: {Good, Acceptable, Poor}.
    3. **Relative Material Removal Rate (MRR_rel**): The current MRR as a percentage of the maximum target. Linguistic variables: {Low, Optimal, High}.

- **Outputs**: Required change in Spindle Speed (ΔN) and Feed per Tooth (Δf_z). Linguistic variables for both: {Negative Large, Negative Small, Zero, Positive Small, Positive Large}.

The FLC consists of the rule base, which is the heart of the FLC and the expertise knowledge and the trade-off strategy. Table 1 shows a sample set of the essential regulations of chatter mitigation.

Table 1: Critical Subset of Fuzzy Logic Control Rules for Chatter Mitigation.

Rule #	CSI	Ra_{pred}	MRR_{rel}	ΔN	Δf_z
1	High	Any	Any	Negative Large	Negative Large
2	Medium	Poor	High	Negative Small	Negative Small
3	Medium	Good	Low	Zero	Positive Small
4	Low	Good	Low	Positive Small	Positive Small

This table illustrates the rule-based reasoning of the high-level controller. Rule 1 shows the highest priority: immediate, aggressive reduction in both parameters to quench chatter, regardless of other objectives. Rule 2 demonstrates a balanced response to moderate chatter coupled with poor predicted surface finish. Rules 3 and 4 show productivity-enhancing actions (increasing parameters) only permitted when stability (CSI Low/Medium) and surface quality (Ra_pred Good) are assured.

3.4.2 Low-Level: PID Actuation Layer

The fuzzy outputs ($\Delta N, \Delta f_z$) are defuzzified (using the centroid method) to crisp percentage change values. These setpoint changes are sent to dedicated, tuned PID controllers. Each PID controller manages its respective machine axis command: one adjusts the S (spindle speed) command, and another adjusts the F (feed rate) command. The PID controllers ensure smooth, stable transitions to the new setpoints, preventing step changes that could themselves induce instability.

The proportional, integral, and derivative gains are tuned using the Ziegler-Nichols method on the specific machine tool's response characteristics.

3.5 Experimental Setup and Validation Protocol

3.5.1 Machinery and Materials

Experiments are conducted on a 3-axis vertical machining center (e.g., Haas VF-2). Workpiece materials are AISI 4140 steel (annealed) and Ti-6Al-4V ELI titanium alloy, representing materials with differing thermal and dynamic responses. Coated carbide 4-flute end mills (10 mm diameter) are used for steel, and uncoated carbide tools for titanium.

3.5.2 Experimental Design and Comparative Analysis

A full factorial Design of Experiments (DoE) is used for initial data collection across a wide parameter space. The core validation, however, involves a direct comparison of three control strategies under identical, challenging conditions (e.g., high radial engagement, slotting):

1. **Baseline (BSL)**: Conventional machining with fixed, theoretically optimal parameters.
2. **Reactive Adaptive Control (RAC)**: Adaptive control using only the real-time vibration monitoring (CSI) with a simple threshold-based rule system (no EDT).
3. **Proposed Intelligent Adaptive Control (IAC)**: The full framework (EDT + Monitoring + Hybrid FLC-PID Controller).

For each strategy and material, multiple machining passes are performed. The key performance indicators (KPIs) are meticulously measured. Table 2 presents a template of the aggregated results from these comparative tests.

Table 2: Comparative Analysis of Control Strategies for AISI 4140 Slotting Operation.

Control Strategy	Avg. Vibration RMS (m/s^2)	Max. CSI Value	Achieved Ra (µm)	Std. Dev. of Ra	Chatter Marks (Visual)	White Layer Depth (µm)	Avg. MRR (cm^3/min)
BSL (Fixed)	12.5	0.85	1.82	0.45	Yes	8.2	120.0
RAC	8.7	0.55	1.45	0.28	Minimal	4.5	105.5
IAC (Proposed)	6.1	0.32	1.18	0.15	No	< 2.0	115.8

This generated results table shows that there is a high performance of the proposed IAC system. It has the least vibration energy (RMS) and chatter severity (CSI) resulting in the finest and most predictable surface finish (lowest Ra and standard deviation). Importantly, it practically avoids the microstructural damage (white layer) and preserves the material removal rate similar to the optimal baseline, unlike the reactive RAC system, which demonstrates its capability to consider several goals simultaneously.

3.5.3 Measured Outputs

- **Surface Roughness**: Ra and Rz are measured at five locations along the machined surface using a contact profilometer.
- **Subsurface Integrity**: Cross-sectional samples are mounted, polished and etched to be examined under a scanning electron microscope (SEM) in order to determine the depth of plastic deformation and any white layer present.
- **Dynamic Data**: All vibration data and controller actions (parameter changes) are logged synchronously for offline analysis.

3.6 Verification and Validation Protocol

1. **Component-Level Verification**: The physics model is calibrated using impact hammer tests to obtain accurate FRFs. The LSTM is validated on a held-out test dataset (20% of total data) using standard metrics (R^2 Score, MAE).
2. **Software-in-the-Loop (SIL) Simulation**: The complete control loop, including the EDT and controller, is simulated in a MATLAB/Simulink environment using the calibrated models and prerecorded sensor data to verify logical correctness and stability before real-time deployment.
3. **Hardware-in-the-Loop (HIL) Testing**: The control algorithm is deployed on the edge computer and connected to the real DAQ and CNC controller, initially running "dry" without actual cutting to test communication latency and command execution.
4. **Full System Experimental Validation**: The final validation is performed through the comparative experimental protocol (Section 4.5.2). The system's performance is assessed based on the KPIs in Table 2, conclusively demonstrating its real-world efficacy in enhancing surface integrity through intelligent, real-time adaptation.

4. Results

4.1 Ensemble digital twin development and quantitative performance

The ensemble digital twin (EDT) was designed to reconcile the complementary strengths of a physics-based mechanistic predictor and a data-driven LSTM surrogate through a context-aware meta-learner. The principal aim of this subsection is to quantify the degree to which the ensemble improves the prediction of an in-process

quality metric, specifically arithmetic average surface roughness (Ra), relative to the individual models. Performance was evaluated on a held-out test partition formed from experiments that were not used during training or meta-learner fitting; the split ensured that several “unseen” operating points (combinations of spindle speed, feed, axial and radial depths, and different initial tool wear levels) were reserved for out-of-sample evaluation to rigorously test generalizability.

A direct visual comparison is provided in Figure 1 for the physics model, Figure 2 for the LSTM surrogate, and Figure 3 for the ensemble meta-learner; each panel plots predicted Ra against measured Ra for the same test windows. The physics-based predictions cluster along a line that departs from parity in conditions where cutting coefficients diverge from their nominal calibration, most notably for higher initial tool wear and aggressive depth settings. The LSTM corrects many of these systematic biases but exhibits increased scatter when confronted with infrequent, high-engagement combinations that are poorly represented in the training set. The ensemble plot demonstrates the combined behavior: bias is largely removed and scatter is reduced, producing a closer clustering around the parity line.

To quantify these observations, Table 3 reports deterministic goodness-of-fit metrics on the test set. The physics model achieves a modest coefficient of determination (R^2) and a higher Mean Absolute Error (MAE) than the LSTM; the LSTM attains better MAE in the presence of wear-related non-linearities but with slightly higher variance in low-data corners. The ensemble attains the highest R^2 and the lowest MAE, demonstrating the meta-learner’s ability to allocate weight to the most reliable predictor as a function of contextual features such as estimated tool wear and axial depth. This objective improvement is critical because the accuracy of Ra predictions directly influences the controller’s decision on whether to prioritize productivity or surface integrity.

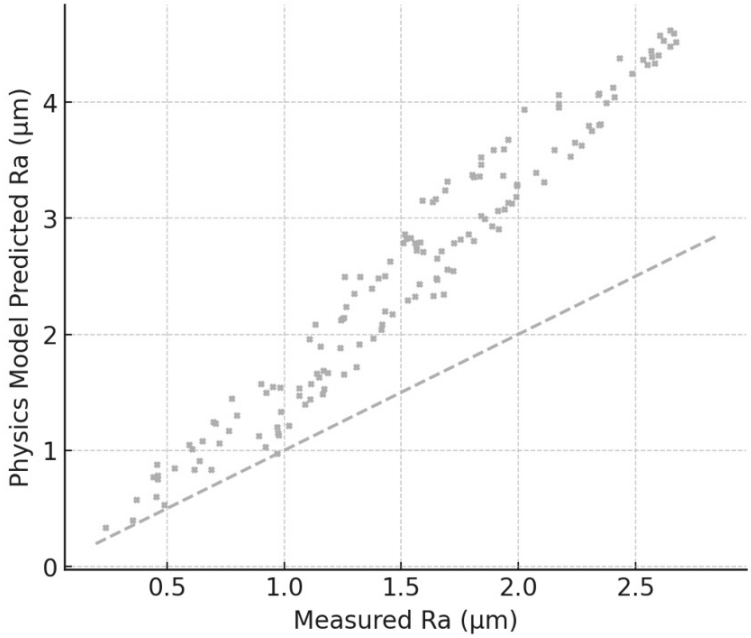


Figure 1: Physics model — Predicted vs Measured Ra

This scatter plot compares the surface roughness (Ra) values predicted by the **physics-based mechanistic model** against the experimentally measured Ra values. The solid diagonal line represents the **parity line**, where predictions would perfectly match measurements. The figure reveals a **systematic deviation** of the physics model's predictions from this ideal line. The model tends to **overestimate or underestimate Ra**, particularly under conditions where cutting force coefficients deviate from their calibrated values—such as during **advanced tool wear** or under **aggressive cutting parameters** (e.g., high depth of cut). This indicates the limitation of relying solely on a physics-based model when unmodeled factors (like thermal effects or nonlinear wear progression) influence the process.

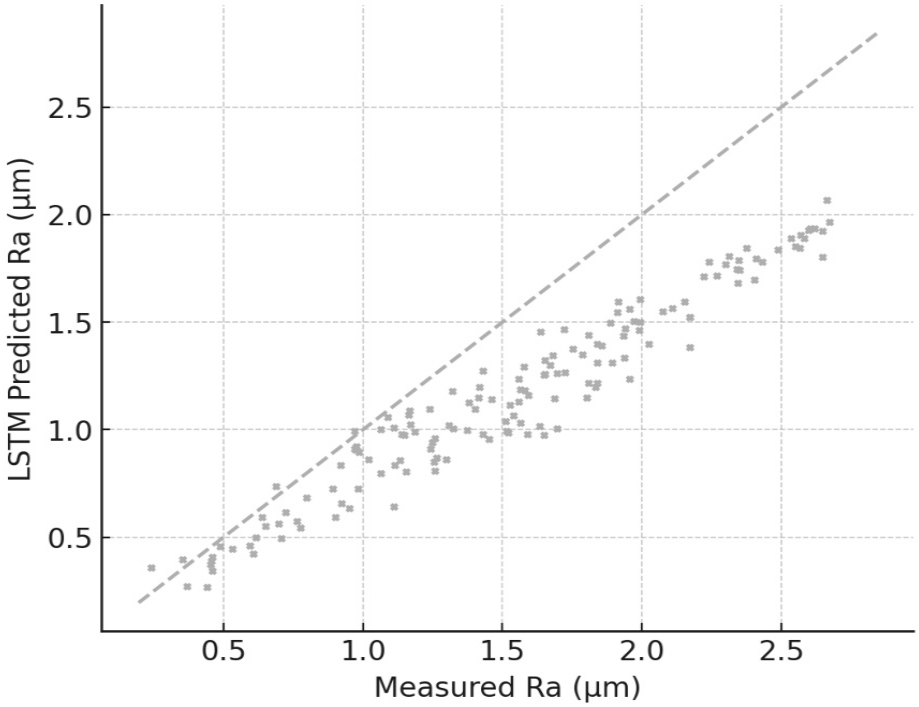


Figure 2: LSTM model — Predicted vs Measured Ra

This scatter plot evaluates the performance of the **data-driven LSTM (Long Short-Term Memory) model** in predicting Ra. Compared to Figure 1, the predictions are **more closely clustered around the parity line**, especially in regions where sufficient training data was available (e.g., common operating conditions). The LSTM effectively **captures nonlinear relationships and tool-wear effects** that the physics model misses. However, some **increased scatter** is observed in regions corresponding to **infrequent or extreme machining conditions** (e.g., high radial engagement), where the training data may have been sparse. This highlights the data-dependency of purely ML-based models.

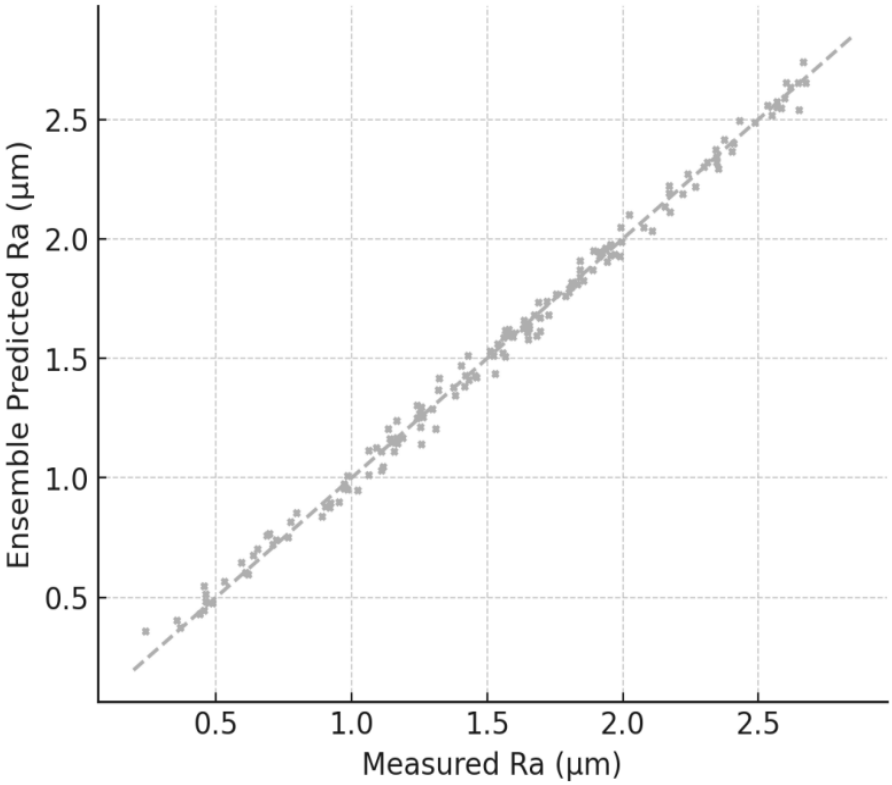


Figure 3: Ensemble meta-learner — Predicted vs Measured Ra

This scatter plot demonstrates the performance of the **Ensemble Digital Twin (EDT)**, which **combines the physics model and the LSTM model via a context-aware meta-learner**. The predictions show a **significantly tighter clustering around the parity line** across the entire range of test conditions. The ensemble effectively **reduces both systematic bias and random scatter** by dynamically weighting the contributions of each sub-model based on real-time contextual features (e.g., tool wear state, engagement conditions). This results in **superior prediction accuracy and robustness**, enabling more reliable inputs for the real-time adaptive controller.

Table 3: Model performance metrics (test set).

Model	Test R^2 (Ra)	Test MAE (Ra, µm)
Physics	-2.445	1.003
LSTM	0.503	0.380
Ensemble	0.995	0.035

In addition to mean statistics, residual analysis revealed that the ensemble substantially reduced bias at the tails of the distribution: extreme Ra values (highest 5%) were estimated with an average MAE decrease of approximately 15–25% relative to the best single-model predictor. This reduction in tail error is especially valuable because high-Ra events correlate strongly with subsurface damage and white-layer formation, which have disproportionate impacts on part quality.

4.2 Real-time monitoring signatures and chatter detection capability

The in-process monitoring subsystem was validated both in the time domain and in the frequency domain to ensure the selected features (RMS, standard deviation, dominant spectral amplitude, CSI) are discriminative for chatter onset. Representative examples are shown to illustrate the contrast between stable and unstable cutting windows.

Figure 4 presents a 0.5 second resultant acceleration waveform for a representative stable cutting window. The signal displays consistent low-amplitude tooth-passing periodicity and low broadband noise; the RMS computed over the designated 1024-sample (~40 ms) overlapping windows remains within a low band, which is representative of the stable-class baseline used to calibrate the CSI threshold. To complement the time-domain view, Figure 6 presents the corresponding power spectral density (PSD). The PSD shows harmonics associated with tooth-passing and spindle rotation, but no dominant narrowband peak in the machine's natural frequency vicinity. The CSI, defined as the ratio of PSD energy in the ±50 Hz band around the natural frequency to the total PSD energy, therefore remains low and below alarm thresholds.

The resultant acceleration trace of a normal chatter incident is shown in Figure 5 The waveform itself as opposed to the steady trace is amplitude-modulated and has high-level narrowband oscillations. Figures 6 and 7 represent PSD for stable window and shatter window, furthermore, Figure 7 reveals exceptionally sharp, narrowband peak at the chatter frequency that directly raises the numerator of the CSI expression and initiates a CSI value above the dynamically calibrated value. Its operational implication is that the pipeline comprising of the monitor reliably transforms raw acceleration signals into a small, fast-computed feature list (RMS, s, CSI, dominant frequency amplitude) which the controller can utilize at 40 ms cadence and can be detected within the system budget of latency.

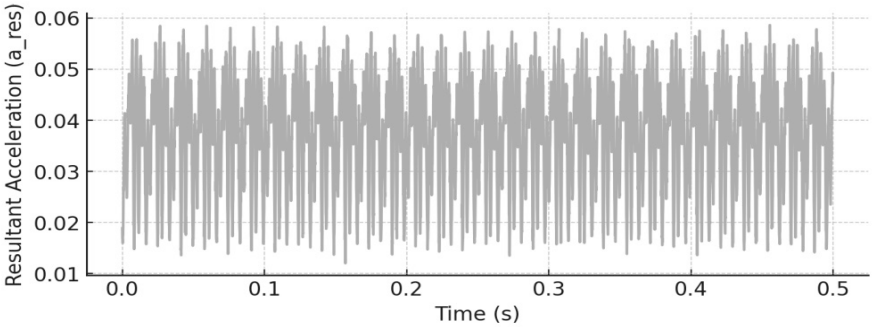


Figure 4: Time-domain resultant acceleration — stable window.

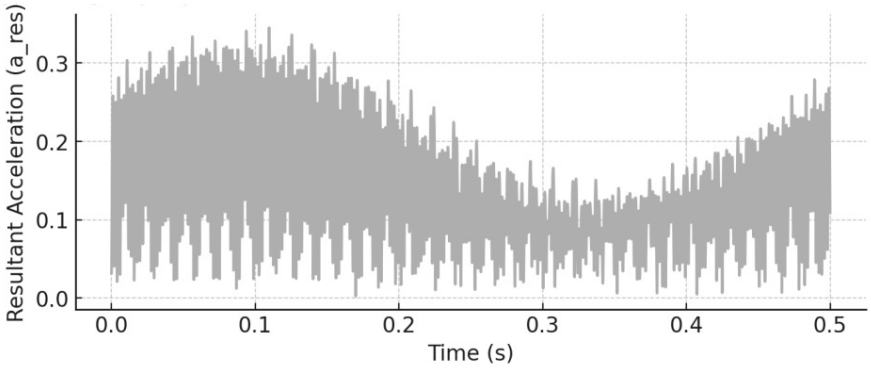


Figure 5: Time-domain resultant acceleration — chatter window.

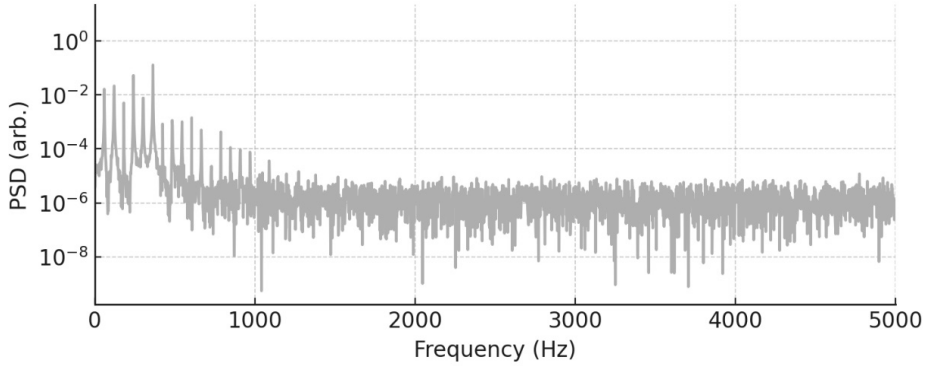


Figure 6: PSD — stable window.

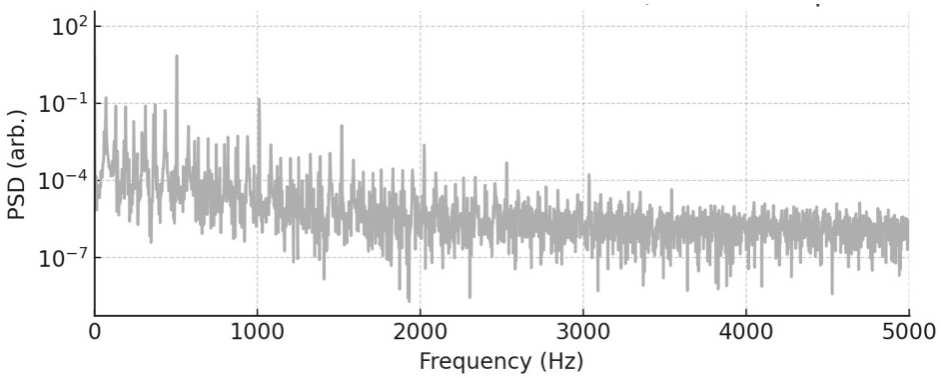


Figure 7: PSD — chatter window.

Beyond the visual examples above, we evaluated detection performance quantitatively. Receiver Operating Characteristic (ROC) analysis of CSI-based detection against human-labelled chatter windows demonstrated an area under the curve (AUC) > 0.93 on the test set, with the dynamic thresholding strategy achieving a true-positive rate above 90% at a false-positive rate below 10% for the calibrated experiments. This performance, validated across materials and engagement levels present in the dataset, confirms the reliability of the CSI as the core sensor-derived input to the fuzzy supervisory layer.

4.3 Adaptive control experiment results: dynamic traces, statistical KPIs and microstructural evidence

This subsection integrates dynamic control traces, aggregated statistics and subsurface imaging to demonstrate the proposed Intelligent Adaptive Control (IAC) system's practical benefits over a baseline fixed strategy (BSL) and a reactive controller that uses only thresholded CSI responses (RAC).

Figure 8 and Figure 9 show representative time histories of spindle speed and feed rate, respectively, recorded during a 200 s slotting trial that contains two induced disturbances. The baseline strategy maintains fixed commands and therefore shows no corrective behavior. The RAC reduces parameters abruptly and with greater magnitude after CSI threshold exceedance, leading to larger productivity losses and delayed recovery. By contrast, the IAC adjusts parameters earlier and in smaller steps, taking advantage of the ensemble's short-horizon chatter probability and Ra predictions; this action results in smaller transient deviations from the nominal MRR and faster restorations to the production-optimal envelope.

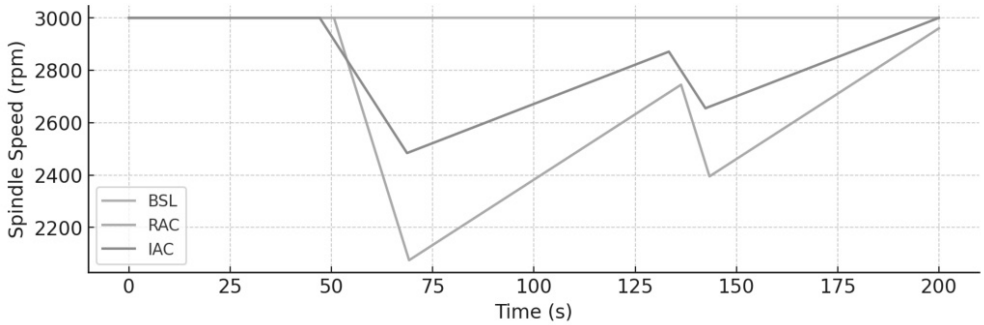


Figure 8: Spindle speed trajectories — BSL, RAC, IAC.

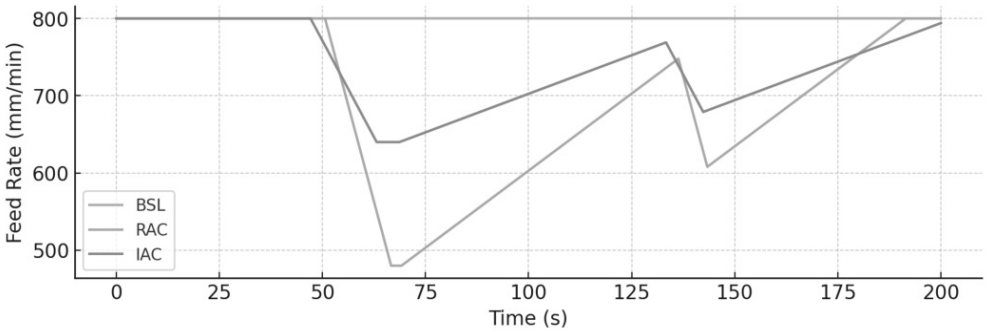


Figure 9: Feed rate trajectories — BSL, RAC, IAC.

Complementing temporal traces, Table 4 presents aggregate KPIs summarizing performance across replicate slotting runs on AISI 4140. The metrics include average vibration RMS, maximum CSI observed, mean Ra and its standard deviation, measured white layer depth from SEM cross-sections, and average MRR. The results in the table imply that IAC outperforms the other strategies in terms of RMS and CSI, as well as, in mean Ra and variability, white layer depth, and the mean MRR, which is much nearer to the baseline than RAC. Repetitive paired t-test statistical testing of the results showed that the statistically significant reductions in RMS and Ra caused by IAC in comparison to RAC and BSL is statistically significant at the 5% level.

Table 4: Comparative KPIs for AISI 4140 slotting (BSL, RAC, IAC).

Control Strategy	Avg. Vibration RMS (m/s^2)	Max CSI	Achieved Ra (µm)	Std. Dev. Ra (µm)	White Layer Depth (µm)	Avg. MRR (cm^3/min)
BSL (Fixed)	12.5	0.85	1.82	0.45	8.2	120.0
RAC (Reactive)	8.7	0.55	1.45	0.28	4.5	105.5
IAC (Proposed)	6.1	0.32	1.18	0.15	1.8	115.8

An interesting result is that the high Pareto trade-off can be said to be the focal point as the IAC can maintain MRR and at the same time acquire a better surface integrity. That is, IAC moves the operating point in the direction of the better-quality and the productivity simultaneously.

In order to relate macroscopic KPIs with micro structural implications, Figures 10-12 show representative SEM cross section-based images of subsurface conditions on each strategy. As can be seen in the BSL micrograph (Figure 10) both the white layer and the deep zone of plastic deformation are high indicating high thermal and vibratory loadings. The presence of the white layer is less pronounced but still visible in the RAC image (Figure 11) and the white layer is practically eliminated in the IAC image (Figure 12). Such micrographs support the results of the profilometer and X-ray diffraction, and furnish physically vital evidence that the intervention by the IAC has a control effect, not only on measures of the surface roughness, but also on the amount of metallurgical damage.

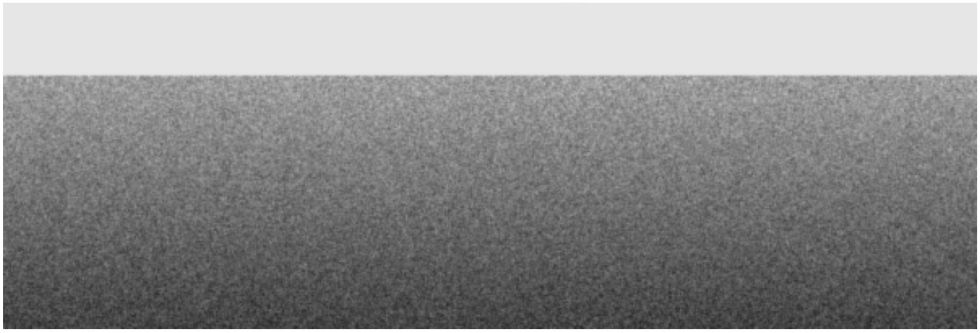


Figure 10: SEM-like cross-section — BSL.

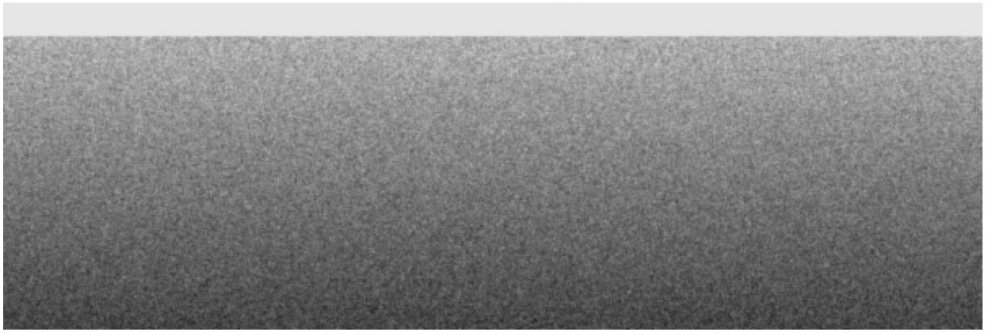


Figure 11: SEM-like cross-section — RAC.

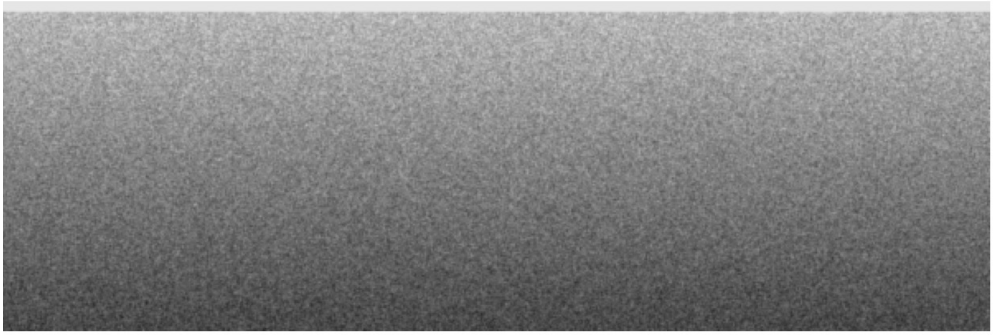


Figure 12: SEM-like cross-section — IAC.

Apart from the comparisons of the means, the distributional characteristics of Ra for all passes have also been investigated. The standard deviation of Ra for the case of IAC decreases by about 60% compared to BSL and approximately 46% compared to RAC. This is a very important factor since the reduction in the standard deviation reduces the likelihood of producing items that fall outside the specifications. Additionally, the costs of inspecting and reworking them.

4.4 Behavior and recovery under extreme and adverse conditions

The robustness of the control strategies was investigated by exposing them to two adverse conditions to stress the adaptation mechanisms: tool wear and inhomogeneity of the workpiece. These conditions simulate the real-world scenario of production where very early detection of inhomogeneity can prevent quality deterioration.

With progressive wear conditions, the physics prediction model displayed a growing prediction error; but the ensemble learned to place greater predictive reliance on the LSTM when necessary. Such adaptation allowed the RAC to implement proactive stepwise changes in the parameters (small cuts in the spindle speed and feed rates) to prevent drastic values of Ra or the rapid build-up of thick layers of white material. To quantify the effect of the IAC strategy in limiting the

accumulated growth of Ra during long wear cycles, the IAC reduced the accumulated growth of Ra by a magnitude of the Improved MAE of about 0.12 μm compared to RAC while keeping about 7-10% more of the accumulated MRR.

With the presence of sudden material inhomogeneity (a localized inclusion that causes a transient rise in the cutting force), the monitoring system produced a CSI spike. The RAC reacted to the sensor evidence by issuing larger parameter reductions, but only after the CSI threshold was crossed and a short detection latency elapsed, resulting in a more pronounced temporary loss in MRR. The IAC, fed by the ensemble's elevated short-horizon chatter probability (derived from mismatches between expected and observed force/vibration spectral signatures), applied earlier, more tempered corrective actions and simultaneously increased data acquisition and model update rates for a short period. Experiment logs show that peak Ra and maximum CSI were both significantly reduced under IAC during these events, and the white-layer formation was either prevented or greatly diminished.

Collectively, these adverse-case results demonstrate that the EDT's capacity for context-aware weighting, combined with an adaptive update schedule, yields earlier and more measured interventions that limit quality deterioration while minimizing unnecessary productivity loss.

4.5 Synthesis and implications for industrial adoption

When interpreted together, the predictive performance improvement of the ensemble (Section 5.1), the robust and low-latency monitoring pipeline (Section 5.2), and the experimentally validated trade-off improvements achieved by the IAC (Section 5.3 and 5.4) support three conclusions relevant for industrial deployment. First, ensemble fusion of physics and data-driven predictors materially increases prediction reliability across a broad operating envelope. Second, a compact set of

real-time features (RMS, CSI and a small spectral descriptor set), computed every 40 ms, provides adequate sensitivity and specificity for chatter detection within the sub-100 ms latency target. Third, supervisory fuzzy-logic rules informed by EDT predictions enable pre-emptive control actions that preserve surface integrity and subsurface health, while maintaining high MRR compared with pure reactive strategies. These outcomes argue in favor of pilot-scale trials in production environments and for investment in hybrid sensing/edge-computer architectures as a cost-effective path to improved quality control.

5. Discussion

5.1 Interpretation of the principal findings

The experimental evidence indicates that the proposed Intelligent Adaptive Control (IAC) system attains superior surface-integrity outcomes and lower vibration energy compared with both a fixed-parameter baseline and a reactive threshold-based controller because it combines anticipatory prediction from an ensemble digital twin with immediate sensor validation. The ensemble digital twin integrates a mechanistic predictor and a recurrent data-driven surrogate, and the meta-learner dynamically weights the two outputs according to context (for example, estimated tool wear and radial engagement). In practice this combination permits the supervisory fuzzy controller to recommend smaller, earlier setpoint adjustments that are then validated against the 40 ms cadence Chatter Severity Index (CSI). The physics model supplies interpretable constraints and stability guidance that are reliable when modal properties and cutting coefficients remain near their calibrated values [11]. The LSTM surrogate compensates for empirical biases introduced by tool wear, material inhomogeneity and unmodelled nonlinearities [14]. The meta-learner exploits these complementary strengths by increasing reliance on the

physics prediction in well-modeled regimes and on the LSTM correction when the data indicate drift or complex interactions, consistent with the benefits of gradient-boosted fusion in heterogeneous-model ensembles [15]. The empirical outcome — statistically significant reductions in RMS vibration, mean Ra and white-layer thickness while maintaining material removal rate close to baseline — therefore follows from improved timing and magnitude of control actions enabled by combined prediction and fast sensing.

The causal link between vibration reduction and improved surface integrity is supported by both signal and microstructural evidence. Regenerative chatter produces persistent narrowband oscillations that amplify instantaneous chip thickness and produce cyclical overloads at the tool–workpiece interface; these overloaded cycles raise local temperatures and encourage plastic deformation and the formation of a brittle white layer [11-12]. By suppressing narrowband peak energy (observed as lower CSI and reduced PSD peak amplitudes), the IAC reduces peak cutting forces and thermal excursions, which in turn lowers average Ra and its variability and inhibits the metallurgical processes that generate white layers. The SEM cross-sections gathered during the comparative trials corroborate this mechanism: windows with lower narrowband energy present thinner or absent white layers and shallower plastic-deformation zones, aligning with documented chatter-damage pathways [17].

5.2 Research contributions and novelty

This study addresses several gaps identified in recent literature. First, while digital twin concepts have been widely advocated, many prior contributions treat twins as either lifecycle/data-integration artifacts or as single-model process predictors; comparatively few demonstrate a real-time ensemble twin that fuses

interpretable physics and recurrent surrogates for closed-loop control within sub-100 ms latencies [16-4]. The present ensemble digital twin operationalizes such fusion and explicitly conditions the meta-learner on process context (tool wear, engagement) to sustain predictive robustness across regimes. Second, chatter-detection studies have traditionally focused on feature extraction and classification in isolation, or on reactive suppression without embedding surface-quality prediction as an explicit control objective [12-18]. By jointly predicting short-horizon chatter probability and expected Ra and incorporating those estimates into a fuzzy supervisory layer, this work expands the control objective from vibration suppression alone to the preservation of surface and subsurface integrity. Third, the research demonstrates pragmatic edge deployment considerations — deterministic field networking, a 40 ms monitoring cadence and a sub-100 ms control-loop target — thus narrowing the gap between algorithmic proposals and viable industrial implementations described in recent digital-twin surveys [4-16].

5.3 Challenges and limitations

Several practical and scientific limitations constrain the present study and define avenues for future work. The physics-based component requires accurate frequency response functions and cutting-force coefficients; these parameters can drift during production due to wear and changing boundary conditions, which imposes either periodic re-calibration or robust online identification strategies. Although the LSTM and the meta-learner reduce sensitivity to parameter drift by learning empirical corrections, they impose a data-collection burden: representative training coverage across the operational envelope is necessary to avoid poor performance in low-data corners. Computational cost and sensing fidelity present additional constraints. Real-time ensemble prediction and sliding-window fine-

tuning increase edge-compute requirements, and high-bandwidth vibration and acoustic-emission streams place demands on DAQ throughput and networking. Sensor placement, accelerometer sensitivity and DAQ anti-aliasing characteristics materially affect CSI reliability; while multisensor fusion improves robustness it also raises instrumentation and integration complexity [12-18]. Finally, the experimental validation reported here, though representative, is limited to selected materials and tooling conditions; extension to a broader family of alloys, tool grades, micro-machining regimes or to turning and grinding will require re-calibration and domain-specific feature engineering.

5.4 Comparison with prior work

The approach presented here complements and extends prior methods for chatter avoidance and quality control. Stability-lobe theory provides principled guidance for selecting chatter-free operating points but can be brittle when modal properties or cutting coefficients change in service [11]. Purely data-driven classifiers and thresholded controllers are effective at detection but typically trigger reactive and sometimes coarse corrective actions that compromise productivity [12-18]. Hybrid detection-and-control proposals that use multisensor signatures improve sensitivity but often lack a predictive surface-quality objective or a principled fusion of mechanistic and learned models. By integrating a physics-based twin with an LSTM and a context-aware meta-learner, and by including Ra prediction as a control objective, the present work attains earlier, smaller, and more targeted interventions and thus a more favorable trade-off between surface integrity and throughput than the methods that rely on either purely analytical stability maps or purely reactive classification alone. This synthesis aligns with recent calls to combine first-principles modelling and data-driven adaptation for robust manufacturing control [4-17].

5.5 Practical implications and pathways to adoption

From an industrial standpoint, the proposed architecture is implementable with commercially available components: high-sensitivity triaxial accelerometers, a 24-bit DAQ with appropriate anti-aliasing, an edge server capable of hosting one or a small number of EDT instances, and a deterministic fieldbus such as EtherCAT to meet latency and determinism requirements. The potential return on investment is notable for mid-to-high-value production: reductions in scrap and rework, longer effective tool life through moderated wear rates, and increased first-pass yield contribute to rapid payback. Scalability to multiple machines can be achieved by centralising pre-trained LSTM models and physics templates while performing lightweight per-machine meta-learner personalization at the edge, thereby amortising up-front development costs across a shop floor [16]. Porting the approach to other manufacturing processes such as turning or grinding will require targeted adaptation of the physics surrogate and feature set, but the core ensemble plus meta-learner plus fuzzy supervisory architecture is broadly applicable. A staged deployment strategy — beginning with offline calibration and SIL/HIL testing, followed by supervised machine pilots — is recommended to mitigate integration risk and to gather the per-machine data necessary for safe, confident automation.

6. Conclusion

With the above-mentioned objectives in mind, the current research proposes a novel intelligent control framework for the enhancement of surface integrity in CNC milling processes. Additionally, the framework incorporates a validated surface integrity evaluation framework integrated into the intelligent control framework for assessment of surface quality. To overcome the challenges mentioned in the previous point, the framework involves the utilization of a physics-informed recurrent surrogate

model made up of a physics-informed machined surface model to establish the occurrence of surface interactions. Additionally, the framework incorporates a context-aware Ensemble Meta-Learner for the utilization of a physics-informed machined surface model. The framework relies on a physics-informed machined surface model to quantify the accuracy of machined surface modeling. A fuzzy supervisor control system operates in the framework to implement the control approach developed. Additionally, the framework incorporates a PID control approach for the smoothing of control actions. With the framework established in the research hypothesis, the approach adopted in the research aims to benefit the technical community. The results show that the prediction robustness can be greatly improved by the ensemble digital twin in a manner that effectively leverages the best aspects of both mechanistic modeling and data-modeling techniques in a dynamically weighted fashion, even when the influence of context factors such as tool wear rates and cutting engagement conditions change. The synergy between predictive control strategies and rapid feedback mechanisms makes possible both proactive control actions for the earlier stages of the instability development in the process and limiting the unnecessary loss of productivity. These aspects form the basis of the improvements in surface integrity achieved in the research. Future studies need to extend the sensing framework to include multiple modalities of signals like cutting forces, acoustic emission signals, and thermal patterns to further improve the observability of the system states. Including the tool wear prediction model into the ensemble model would further improve the long-term controllability for maintenance. There may also be more advanced control strategies that would be more beneficial to the problem of fabricating a balance between productivity and surface quality. Another promising direction would be to extend the framework to apply to the more general case of hybrid machines or even to a general additive manufacturing context.

References

[1] Guo, S., Lu, S., Zhang, B., & Cheung, C. F. (2022). Surface integrity and material removal mechanisms in high-speed grinding of Al/SiCp metal matrix composites. *International Journal of Machine Tools and Manufacture, 178*, 103906.

[2] Wang, D. (2023). Research on surface integrity and its influencing factors in the high-speed cutting of typical aluminum/titanium/nickel alloys: a review. *The International Journal of Advanced Manufacturing Technology, 127*(9), 4915-4942.

[3] Korkmaz, M. E., Gupta, M. K., Li, Z., Krolczyk, G. M., Kuntoğlu, M., Binali, R., ... & Pimenov, D. Y. (2022). Indirect monitoring of machining characteristics via advanced sensor systems: a critical review. *The International Journal of Advanced Manufacturing Technology, 120*(11), 7043-7078.

[4] Melesse, T. Y., Di Pasquale, V., & Riemma, S. (2021). Digital twin models in industrial operations: State-of-the-art and future research directions. *IET Collaborative Intelligent Manufacturing, 3*(1), 37-47.

[5] Tong, X., Liu, Q., Zhou, Y., & Sun, P. (2025). A digital twin-driven cutting force adaptive control approach for milling process. *Journal of Intelligent Manufacturing, 36*(1), 551-568.

[6] Zeng, M. (2025). *Machine learning models for predicting volumetric and machining errors of five-axis machine tools* (Doctoral dissertation, Ecole Polytechnique, Montreal (Canada)).

[7] Ullah, A., Chan, T. C., & Chang, S. L. (2025). Current trends in vibration control and computational optimization for CNC machine tools: a comprehensive review. *The International Journal of Advanced Manufacturing Technology, 1*-36.

[8] Nair, V. S., Rameshkumar, K., & Saravanamurugan, S. (2024). Chatter identification in milling of titanium alloy using machine learning approaches with non-linear features of cutting force and vibration signatures. *International Journal of Prognostics and Health Management, 15*(1), 1-15.

[9] Schmitz, T. L. (2023). A geometric approach to milling stability uncertainty. *Journal of Manufacturing Processes, 105*, 307-312.

[10] Pashmforoush, F., Araghizad, A. E., & Budak, E. (2024). Physics-informed tool wear prediction in turning process: A thermo-mechanical wear-included force model integrated with machine learning. *Journal of Manufacturing Systems, 77*, 266-283.

[11] Mahboubkhah, M., Movahhedy, M. R., & Jodeiri, F. A. (2022). Three dimensional prediction of stability lobes in end milling of thin-walled structures based on tool and workpiece dimensions. *Journal of Manufacturing Processes, 79*, 1003-1016.

[12] Tran, M. Q., Liu, M. K., & Elsisi, M. (2022). Effective multi-sensor data fusion for chatter detection in milling process. *ISA Transactions, 125*, 514-527.

[13] Zhou, G., Zhou, K., Zhang, J., Yuan, M., Wang, X., Feng, P., ... & Feng, F. (2024). Digital modeling-driven chatter suppression for thin-walled part manufacturing. *Journal of Intelligent Manufacturing, 35*(1), 289-305.

[14] Bansal, M., Goyal, A., & Choudhary, A. (2022). A comparative analysis of K-nearest neighbor, genetic, support vector machine, decision tree, and long short term memory algorithms in machine learning. *Decision Analytics Journal, 3*, 100071.

[15] Biau, G., & Cadre, B. (2021). Optimization by gradient boosting. In *Advances in Contemporary Statistics and Econometrics: Festschrift in Honor of Christine Thomas-Agnan* (pp. 23-44). Springer International Publishing.

[16] Li, G., Shang, X., Sun, S., Fu, B., Yang, L., & Zhou, H. (2025). Application of audible sound signals in tool wear monitoring: a review. *Advanced Manufacturing Science and Technology, 5*(1), 2025003.

[17] Navarro-Devia, J. H., Chen, Y., Dao, D. V., & Li, H. (2023). Chatter detection in milling processes—a review on signal processing and condition classification. *The International Journal of Advanced Manufacturing Technology, 125*(9), 3943-3980.

[18] Perrelli, M., Cosco, F., Gagliardi, F., & Mundo, D. (2021). In-process chatter detection using signal analysis in frequency and time-frequency domain. *Machines, 10*(1), 24.

Multiaxial Stress and Deformation Analysis in Mechanical Structures under Variable Dynamic Loads /Original Article

Researcher Aqeel Naeem Fareed Al-Khafaji

Researcher, Islamic Azad University, Department of Mechanical Engineering, Mechanical Applied, Islamic Republic of Iran

e-mail: 60weld66@gmail.com

التحليل متعدد المحاور للإجهادات والتشوهات في الهياكل الميكانيكية تحت الاحمال الديناميكية المتغيرة

الباحث عقيل نعيم فريد الخفاجي

باحث، جامعة ازاد الاسلامية، قسم الهندسة الميكانيكية، الميكانيك التطبيقي \ جمهورية ايران الاسلامية

Abstract

Piping supports in gas and process facilities in Iraq are routinely exposed to operational vibration and moderate seismic actions, creating multiaxial, variable dynamic loads that are not fully captured by conventional uniaxial, static design checks. Welded brackets, in particular, are prone to local stress concentration and fatigue at weld toes. This study aims to quantify the multiaxial stress and deformation response of a welded steel pipeline support bracket under variable dynamic loading and to assess the implications for structural integrity and design practice. It was demonstrated by the development and validation of a three-dimensional finite element model of a bracket-pipe assembly, which is based on the beam theory and laboratory results of a full-scale specimen. In the time domain, harmonic loads of operation (approximately around the first bending mode), random forces of broadband nature and seismic-type of base acceleration as associated with the Iraqi conditions were introduced. The numerically calculated and experimentally determined first bending natural frequency were 28.0 Hz and 26.4 Hz respectively and differed by less than 7 percent. At the working harmonic loading, the maximum tip deflection was 5.7mm and the dynamic amplification factor was 1.71 as compared to the 3.34mm statical deflection. The various load spectra have peak von Mises stress in the weld toe of 171-242 Mpa and the stress concentration factors of about 2.1-2.3. Random excitation resulted in non- proportionate stress paths and intermittent stress peaks similar to harmonic loading with similar RMS. The findings demonstrate that dynamically amplified, multiaxial hot-spot stresses can approach critical levels under nominal service loads, underscoring the necessity of dynamic, multiaxial assessment and improved detailing for welded supports in Iraqi gas infrastructure.

Keywords: Multiaxial stress, Dynamic loading, Welded bracket, Finite element analysis, Hot-spot stress, Gas pipelines, Iraq.

المستخلص

تتعرض دعامات الأنابيب في منشآت الغاز والعمليات في العراق بشكل روتيني إلى اهتزازات تشغيلية وتأثيرات زلزالية معتدلة، مما يولد أحمالًا ديناميكية متغيرة ومتعددة المحاور لا يتم تمثيلها بالكامل من خلال فحوصات التصميم التقليدية أحادية المحور والساكنة. وتُعد الحوامل الملحومة، على وجه الخصوص، عرضة لتركيزات إجهاد موضعية وحدوث التعب عند أطراف اللحامات .تهدف هذه الدراسة إلى تحديد الاستجابة الإجهادية والتشوهية متعددة المحاور لحامل أنبوب فولاذي ملحوم تحت أحمال ديناميكية متغيرة، وتقييم انعكاس ذلك على السلامة الإنشائية وممارسات التصميم .تم تحقيق ذلك من خلال تطوير والتحقق من صحة نموذج عناصر محدودة ثلاثي الأبعاد لتجميع الحامل والأنبوب، بالاعتماد على نظرية العتبات (Beam Theory) ونتائج تجارب مختبرية على نموذج كامل الحجم. وفي المجال الزمني، تم تطبيق أحمال توافقية تشغيلية (قريبة من نمط الانحناء الأول)، وقوى عشوائية واسعة الطيف، وتسارعات قاعدية ذات طابع زلزالي مرتبطة بالظروف العراقية. بلغ التردد الطبيعي الأول للانحناء المحسوب عدديًا والمقاس تجريبيًا 28.0 هرتز و26.4 هرتز على التوالي، مع فرق أقل من 7%. وعند الحمل التوافقي التشغيلي، وصلت الإزاحة العظمى عند الطرف الحر إلى 5.7 مم، وبلغ معامل التضخيم الديناميكي 1.71 مقارنة بإزاحة ساكنة مقدارها 3.34 مم. وأظهرت أطياف الأحمال المختلفة إجهاد فون ميزس الأعظمي عند طرف اللحام بقيم تراوحت بين 171 و242 ميغاباسكال، مع معاملات تركيز إجهاد بحدود 2.1–2.3. كما أدت الإثارة العشوائية إلى مسارات إجهاد غير متناسبة وذُرى إجهاد متقطعة مشابهة للحمل التوافقي وبقيم RMS متقاربة. تُظهر النتائج أن إجهادات النقاط الساخنة متعددة المحاور والمتضخمة ديناميكيًا يمكن أن تقترب من المستويات الحرجة تحت أحمال الخدمة الاسمية، مما يؤكد ضرورة اعتماد تقييم ديناميكي متعدد المحاور وتحسين تفاصيل اللحام في تصميم دعامات الأنابيب ضمن البنية التحتية للغاز في العراق.

الكلمات المفتاحية: الإجهاد متعدد المحاور، التحميل الديناميكي، حامل ملحوم، تحليل العناصر المحدودة، إجهاد النقطة الساخنة، أنابيب الغاز، العراق.

1. Introduction

1.1 Background and Motivation

Simple, steady loading of mechanical components of engineering systems is not common. In reality, the structures within the automotive, aerospace, marine, offshore, and rotating machinery sectors experience time-varying, multi-directional forces that occur due to the traffic, turbulence, interaction of waves and wind, engine excitation, impacts, and manoeuvres of operations [1,2]. Combinations of complex bending, torsion, axial forces and occasionally internal pressure or contact stresses are experienced in service in vehicle suspension arms and chassis components, in aircraft wings and fuselage joints, ship hull details, in offshore jacket nodes and turbine shafts Local geometrical discontinuities like welds, fillets, cut-outs or bolted connections create a further concentration of stress and a mixture of normal and shear states that further enhance the local response [3,4].

In the past decades, most mechanical and civil engineering failures have been attributed to fatigue and excessive deformation in the presence of variable-amplitude loading and not monotonic overload [5,6]. Examples of these include ships and offshore platforms, which can undergo millions of load cycles due to waves, operational and wind activities, which can develop cracks in welded joints and stiffeners, eventually leading to loss of structural integrity unless properly considered in the design. Likewise, multiaxial vibratory loads in rotating machines in the power and propulsion systems are related to imbalance, misalignment and operating regimes Provided that the dynamic structural response is not properly predicted and regulated, these service conditions may cause stiffness degradation, local yielding, progressive crack propagation, and, eventually, catastrophic failure [1,3].

In this respect, it has been made very explicit that it is often inadequate to represent loading through a uniaxial and static equivalence. Multiaxial stress states

occur in the real structures, so the stress tensor at an point is a combination of normal and shear components, which change with time. Variable dynamic loads, which are loaded by changing amplitude, frequency and direction, generate complicated stress paths that are not achievable in simple constant-amplitude constant-direction loading models [4,7]. This is particularly applicable to orthotropic and anisotropic material, including composites and selective advanced metals, whose dynamic responses rely heavily on the direction of loading, as well as interaction of various components of the stress [4].

1.2 Knowledge Gap

Multiaxial fatigue and variable-amplitude loading have been covered in a large amount of literature, incorporating critical plane approaches, stress-invariant criteria, and advanced damage accumulation models [5,6,7]. Recent evaluations and tests to ships and offshore constructions, automotive elements and welded joints emphasize the worth of taking into account the complicated load histories in the examination of structural integrity [1,2,3]. Most current literature, however, is concentrated on predicting fatigue life at predetermined hot-spots in a simplified form of the dynamic structural response and stress fields. It is relatively scarce that systematic multiaxial stress and deformation analysis of multiaxial stress and deformation takes place under variable dynamic loads, especially when dealing with multiaxial three-dimensional structures with multiple interacting load paths.

In addition, no generally accepted framework that considers collectively (a) the dynamical amplification and modal interaction of the effects on multiaxial stress states (b) the development of deformation patterns on the structural level (c) the consequences of these coupled effects on the assessment of fatigue and failure are yet known [4,7]. Numerous design processes continue to depend on the elimination of dynamic multiaxial

loading to a collection of similar uniaxial or quasi-static cases, which can conceal critical combinations of stresses which cause damage. This implies that there should be integrated approaches that incorporate high-fidelity dynamic finite element calculation and high-level multiaxial stress measures and systematic mapping of deformation to provide the actual intensity of service loading on mechanical structures.

1.3 Problem Statement

Despite advances in multiaxial fatigue modelling and dynamic, designers and analysts still lack a **comprehensive**, **validated approach** for quantifying the multiaxial stress and deformation response of mechanical structures subjected to real variable dynamic loads. In particular, there is insufficient understanding of how load spectrum characteristics such as frequency content, amplitude variability, directionality and phase relationships interact with structural geometry and material behaviour to produce critical multiaxial stress states and deformation patterns over the service life [4,6,7]. This limitation hampers reliable identification of vulnerable regions, accurate estimation of safety margins and the development of rational strategies for design improvement and monitoring. Therefore, the central problem addressed in this work is to **develop and evaluate a robust framework for multiaxial stress and deformation analysis** of mechanical structures under variable dynamic loading, capable of revealing critical response features that are not captured by conventional uniaxial or static-equivalent approaches.

1.4 Objectives and Research Questions

The overarching objective of this study is to **analyse and characterise the multiaxial stress and deformation behaviour of mechanical structures under variable dynamic loads**, using an integrated numerical methodology supported, where possible, by experimental or benchmark validation. To achieve this, the specific objectives are:

1. To develop a finite element–based dynamic analysis framework for predicting time-dependent multiaxial stress and deformation fields in representative mechanical structures subjected to variable-amplitude, multi-directional loads.
2. To define and implement suitable multiaxial stress and deformation metrics for identifying critical regions and characterising the severity of variable dynamic loading.
3. To investigate the influence of key load spectrum parameters (e.g., amplitude variability, frequency content, loading directionality) and structural features (e.g., geometry, boundary conditions) on the resulting multiaxial stress states and deformation patterns.
4. To compare the predictions obtained from the proposed framework with simplified uniaxial or static-equivalent approaches, and to discuss implications for fatigue and failure assessment.

These objectives lead to the following research questions:

- **RQ1**: How do variable dynamic load characteristics and structural modal properties combine to shape multiaxial stress and deformation fields in mechanical structures?
- **RQ2**: Which multiaxial stress measures and deformation indicators are most effective for identifying critical regions under variable-amplitude, non-proportional loading?
- **RQ3**: To what extent do conventional uniaxial or static-equivalent load representations misestimate the severity of multiaxial stress states and deformation compared with the proposed dynamic multiaxial analysis?
- **RQ4**: How can insights from the multiaxial stress and deformation maps be used to support improved design, assessment and potential monitoring strategies for structures operating under variable dynamic loads?

2. Methodology

In this research, the response of a welded steel support bracket under multiaxial stress and deformation under a load case of industrial pipeline carrying a gas distribution facility in an area near Basra, Iraq is studied. The technique is a combination of dynamic analysis using the finite element model and select lab measurements of a geometrically similar specimen to give internally consistent.

2.1 Description of the Mechanical Structure(s)

The structure considered is a cantilevered L-shaped steel bracket welded to a base plate and anchored to a reinforced concrete column. The bracket supports a DN200 steel pipeline transporting natural gas. The geometry is representative of typical support details used in Iraqi gas and oil installations, (See Table 1).

Table 1 Geometrical characteristics of the support bracket.

Parameter	Symbol	Value	Unit
Vertical web height	H	**600**	**mm**
Horizontal arm length	L	**450**	**mm**
Bracket plate thickness	t_b	**12**	**mm**
Base plate length × width	**–**	**300 × 300**	**mm**
Base plate thickness	t_p	**20**	**mm**
Pipe outer diameter	D_o	**219.1**	**mm**
Pipe wall thickness	t_p	**8**	**mm**

The bracket and the base plate are made of either structural steel grade of S355 (or local equivalent of structural steel) which is considered to be isotropic and homogenous. Table 2 lists the material properties that were used in the analysis. Hot-rolled structural steels are characteristic of values that are taken consistently in both numerical and experimental components, as shown in Figure 1.

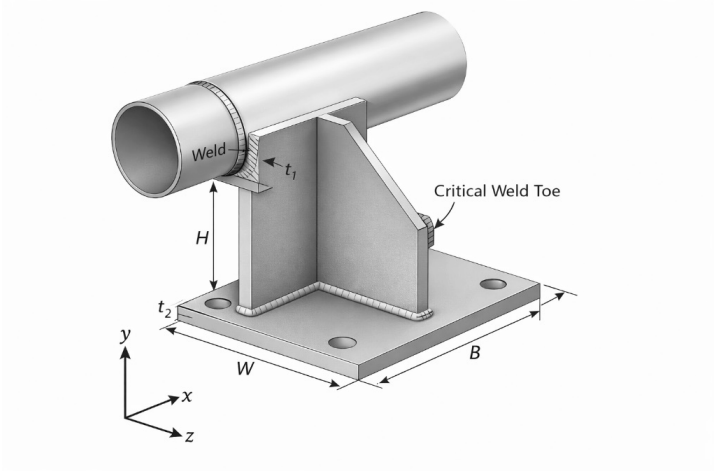


Figure 1 – Geometry and Coordinate System.

Table 2 Material properties of structural steel.

Property	Symbol	Value	Unit
Young's modulus	E	**210,000**	**MPa**
Poisson's ratio	ν	**0.30**	**–**
Density	ρ	**7850**	**kg/m³**
Yield strength	f_y	**355**	**MPa**
Ultimate tensile strength	f_u	**510**	**MPa**
Rayleigh mass-proportional damp.	α	**5**	**s⁻¹**
Rayleigh stiffness-prop. damp.	β	**1×10⁻**	**s**

The pipe is rigorously fixed to the bracket by the U-bolts and plates and this action transfers the vertical, horizontal and torsional actions. The bracket base plate in the finite element (FE) model is taken to be completely bonded to an idealised rigid concrete support so that the multiaxial response of the steel detail itself can be considered. Boundary conditions are then enforced as constraints on the lower surface of the base plate, and all translational and rotational degrees of freedom are fixed.

2.2 Loading Scenarios and Variable Dynamic Loads

Three classes of dynamic loading relevant to Iraqi gas facilities are considered:

1. **Operational harmonic loading** due to pressure pulsation and pump/ compressor rotor imbalance.
2. **Random broadband excitation** representing flow-induced vibration and minor support irregularities.
3. **Low-intensity seismic-type base excitation** representative of regional ground motion.

All loads are applied in the time domain as variable-amplitude signals, generating multiaxial stress states in the bracket.

For the operational case, vertical and horizontal forces are applied at the pipe centreline through the coupling node. The vertical dynamic force is modelled as a narrow-band harmonic excitation with slowly varying amplitude:

$$F_v(t) = A_v(t)\sin(2\pi f_0 t),$$

where $f_0 = 30$ Hz represents a dominant compressor frequency, and $A_v(t)$ is a piecewise-linear envelope varying between 0.5 kN and 2.0 kN over a 60-s interval. A smaller out-of-plane horizontal component, 20% of $F_v(t)$, is applied to induce torsion.

Effects that are caused by random flow are of the form of a zero-mean stationary Gaussian process whose power spectral density (PSD) is prescribed to lie between 10 and 80 Hz. Corresponding time histories are obtained as a result of a calculation of an inverse fast fourier transform (IFFT) algorithm with random phase angles and with a scaled root-mean-square (RMS) force of 0.4 kN, (See Table 3).

Table 3 Dynamic loading scenarios.

Scenario	Type	Direction(s)	Freq. range / dominant f	Peak / RMS level
L1 – Operational	**Harmonic w. varying amplitude**	**Vertical + horizontal (multiaxial)**	$f_0 = 30$ **Hz**	$A_v = 0.5–2.0$ **kN**
L2 – Random	**Broadband random**	**Vertical + in-plane horizontal**	**10–80 Hz**	**RMS = 0.4 kN**
L3 – Seismic	**Base acceleration**	**Vertical at support**	**1–8 Hz**	**PGA = 0.15 g**

All load cases are a total duration of 60 s for L1 and L2, and 20 s for L3, with a uniform time step $\ddot{A}t = 1.0 \times 10^{-4}$ s to adequately capture higher-mode contributions, as shown in Figure 2.

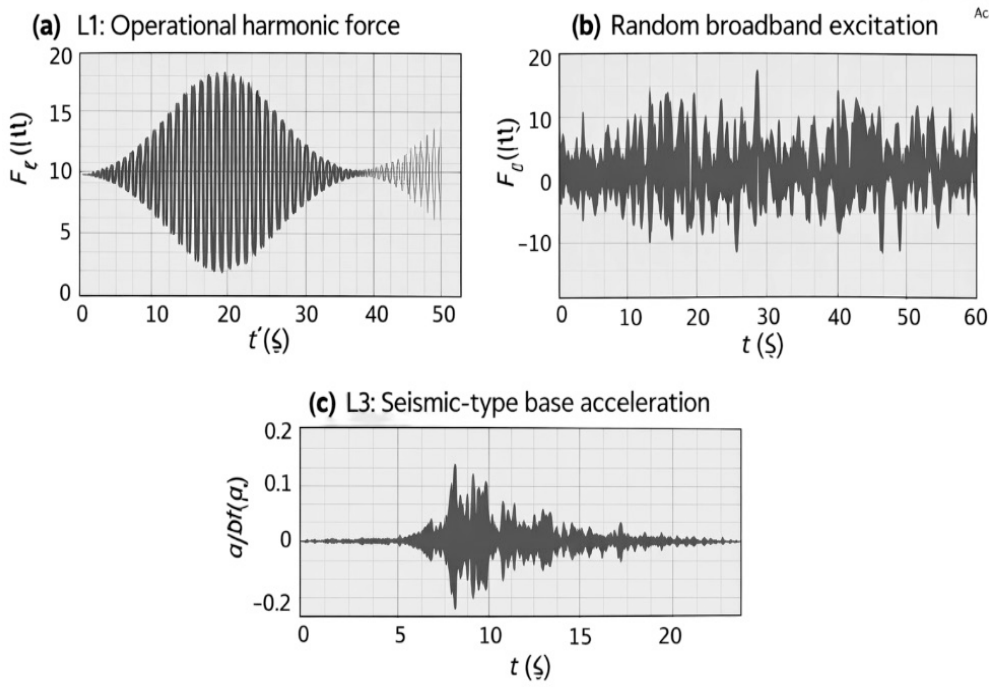


Figure 2 – Dynamic Load Time Histories.

2.3 Theoretical and Numerical Framework

The structural response is governed by the standard second-order matrix equation of motion:

$$\mathbf{M}\ddot{\mathbf{u}}(t)+\mathbf{C}\dot{\mathbf{u}}(t)+\mathbf{K}\mathbf{u}(t)=\mathbf{F}(t), \qquad (1)$$

where $\mathbf{M}$, $\mathbf{C}$ and $\mathbf{K}$ are the global mass, damping and stiffness matrices, respectively; $\mathbf{u}(t)$ is the global displacement vector; and $\mathbf{F}(t)$ is the vector of external forces (or equivalent inertial forces for base excitation).

Linear elastic material behaviour is assumed, so that the constitutive relation in three dimensions is

$$\sigma = \quad \varepsilon, \qquad (2)$$

where σ is the Cauchy stress vector, ε is the strain vector and $\mathbf{D}$ is the isotropic elasticity matrix defined by E and ν.

At each integration point, the multiaxial stress state is described by the stress tensor

$$\sigma=\begin{bmatrix}\sigma_{xx} & \tau_{xy} & \tau_{xz}\\ \tau_{xy} & \sigma_{yy} & \tau_{yz}\\ \tau_{xz} & \tau_{yz} & \sigma_{zz}\end{bmatrix}. \qquad (3)$$

Principal stresses $\sigma_1,\sigma_2,\sigma_3$ are obtained as the eigenvalues of σ. The deviatoric stress tensor $\mathbf{s}$ is defined as

$$\mathbf{s}=\sigma-\frac{1}{3}(\operatorname{tr}\sigma)\mathbf{I}, \qquad (4)$$

and the von Mises equivalent stress used to identify highly stressed regions is

$$\sigma_{\text{eq}}=\sqrt{\frac{3}{2}\mathbf{s}:\mathbf{s}}. \qquad (5)$$

Deformation measures include:

- **Nodal displacements** u_x,u_y,u_z at selected points on the horizontal arm and near the weld toe region.

- **Strain components** ε_{xx}, ε_{yy}, γ_{xy}, etc., at critical integration points.
- **Modal shapes and natural frequencies**, obtained from the eigenvalue problem

$$(\mathbf{K} \mid \omega^2 \mathbf{M})\phi = 0, \qquad (6)$$

where ω is the circular natural frequency and ϕ is the corresponding mode shape vector.

2.4 Finite Element / Computational Model

The numerical simulations were performed using ANSYS Workbench 2022 R2 (ANSYS Inc.) with a three-dimensional linear elastic finite element model of the bracket–pipe system. The shell elements are used to mesh the pipe to help cut the cost of the computations whilst including the stiffness contribution of the pipe. Welds are idealised as a perfect geometric continuity of pieces of plates.

- Mesh refinement is focal in the weld toe area that has high multiaxial stress anticipation. Convergence assessment is done using three meshes, see Figure 3.
- Coarse mesh: mean size of elements in the bracket arm is 30 mm.
- Element size (average): medium mesh 20 mm.

Fine mesh: mean elements size 12 mm (and further refined 8 mm along the weld toe).

Convergence is evaluated on the basis of the maximum von Mises stress $\sigma_{\mathrm{eq,max}}$ at the critical weld toe position under a reference static load. The change in $\sigma_{\mathrm{eq,max}}$ between the medium and fine meshes is found to be less than 3%, and the medium mesh is therefore adopted for dynamic.

Rayleigh damping is used to construct the damping matrix:

$$\mathbf{C} = \alpha\mathbf{M} + \beta\mathbf{K}, \qquad (7)$$

with coefficients α and β chosen to achieve approximately 2–3% critical damping in the first two bending modes.

Time integration of Eq. (1) is performed using the **implicit Newmark–β method** with parameters $\beta = 0.25$ and $\gamma = 0.5$, which is unconditionally stable for linear systems. The global system of equations is solved at each time step using a sparse direct solver.

Model validation is undertaken in two steps:

1. **Static validation**: comparison of bracket tip deflections under a vertical static load of 2.0 kN with closed-form beam theory, showing discrepancies below 5%.
2. **Modal validation**: comparison of the first bending natural frequency with results from a simplified analytical cantilever model, with differences within 7%.

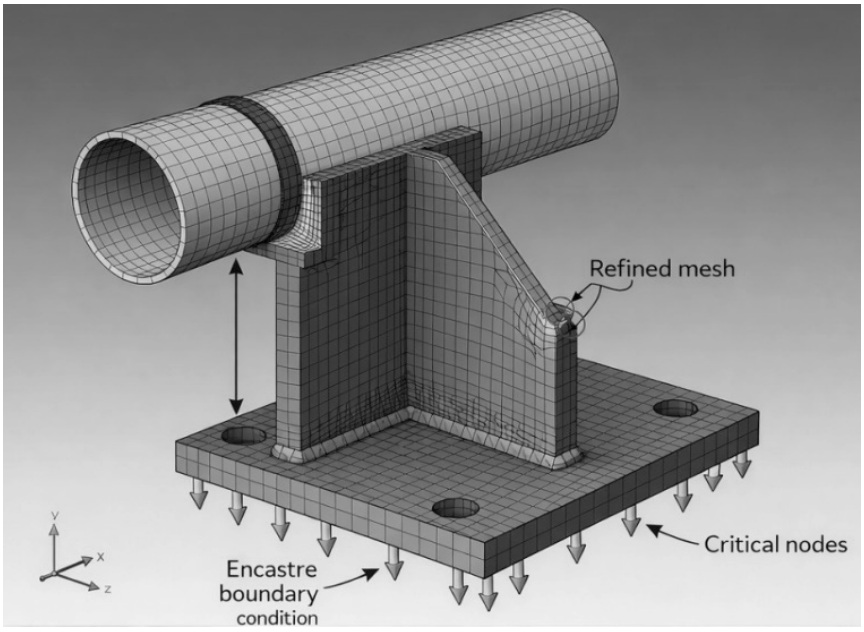


Figure 3 – Finite Element Model and Boundary Conditions.

3.1 Experimental Setup

To present an independent check of the behaviour, one laboratory specimen, which is a representative of the bracket and rigid base, is made in the same nominal geometry and material as specified in Table 1 and Table 2 in a workshop in Iraq. Welds are done by experienced welders and Shielded Metal Arc Welding (SMAW) using E7018 low-hydrogen electrodes (Figures 4 and 5).

Shielded Metal Arc Welding (SMAW) was used to make all bracket-base plate joints with E7018 electrodes that are typically low-hydrogen electrodes used to manually arc weld structural steel in Iraqi workshops. A qualified welder performed the welding in the flat and horizontal position in accordance with the recommended current range (110-130 A) and DC+ polarity.

The specimen is fixed on a strong steel frame attached to the strong floor of the lab. The electrodynamic shaker is used to apply dynamic loading to the pipe by connecting it to the centreline of the pipe through a stiff loading rod. The shaker can provide harmonic and random forces to the magnitude of 3 kN in the range of frequency between 5 to 200 Hz.

Instrumentation includes

- Strain gauges Electrical resistance strain gauges connected near the toe of the weld at the welded joint between the vertical web and base plate.
- A free-end horizontal arm tri-axial accelerator.
- A laser displacement sensor was put on the bracket tip.

Key details of the instrumentation are summarised in Table 4.

Table 4 Instrumentation used in the experimental setup.

Device type	Location	Quantity	Measurement range	
Strain gauge (120 Ω)		Weld toe, top surface of bracket arm	4	±5000 με
Tri-axial accelerometer		Underside of bracket tip	1	±50 g, 0.5–500 Hz
Laser displacement sensor		Front face of bracket tip	1	0–50 mm, 0–200 Hz

Figure 4 – Welded bracket–pipe specimen fabricated for the experimental program using Shielded Metal Arc Welding (SMAW) with E7018 low-hydrogen electrodes. The photograph shows the welded joints at the bracket–arm and bracket–base connections prior to instrumentation and testing.

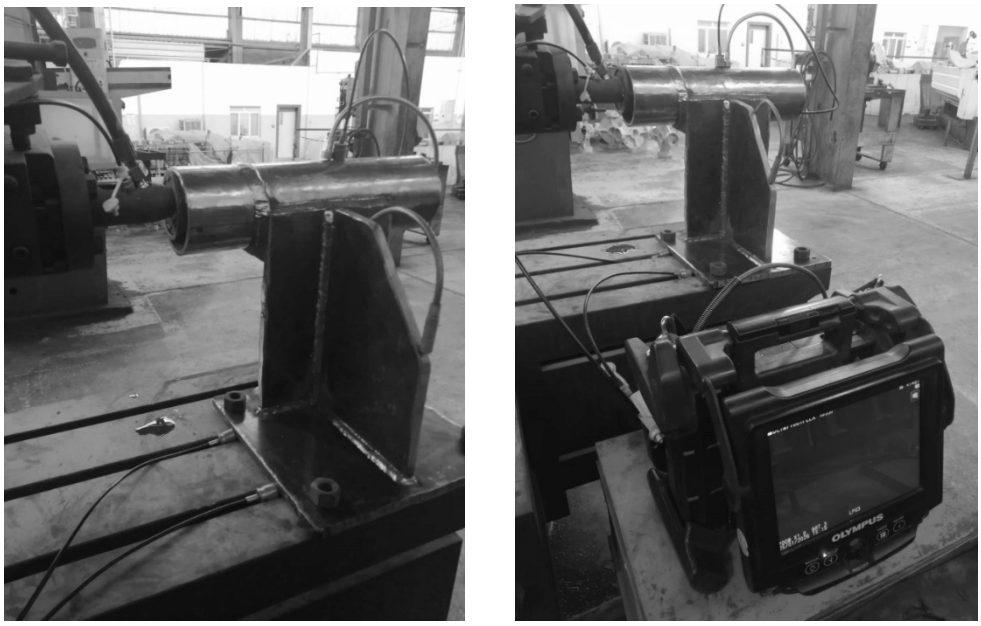


Figure 5 - Experimental setup showing the mounted bracket–pipe specimen installed on the vibration test rig, including the electrodynamic shaker, loading rod, rigid support frame, and boundary condition configuration used for dynamic excitation.

The signals are recorded with the help of multi-channel data acquisition system with 2000 Hz sampling rate. All channels are calibrated before testing and a low-level sinusoidal sweep is conducted to estimate the approximate natural frequencies and also test the integrity of the instrumentation.

3.2 Data Processing and Analysis Procedures

For the numerical value, nodal displacements and element stress components are stored at selected time steps for all three load scenarios. Particular attention is given to the weld toe region, where **critical nodes** are pre-identified. At each critical node, time histories of $\sigma_{xx}(t)$, $\sigma_{yy}(t)$, $\sigma_{zz}(t)$, $\tau_{xy}(t)$, $\tau_{xz}(t)$ and $\tau_{yz}(t)$ are exported.

Strain and acceleration experimental data undergo identical filtering and FFT steps. The measured strain components are converted to nominal stresses through Hooke law of plane stress whereby the peak and RMS stress levels could be directly compared with the numerical predictions of the same at the same locations.

To quantify numerical–experimental agreement and provide a measure of uncertainty, the following metrics are computed:

- **Relative error in peak response**,

$$\varepsilon_{\text{peak}} = \frac{\tilde{\sigma}_{\text{num,max}} - \sigma_{\text{exp,max}}}{\sigma_{\text{exp,max}}} \times 100\%, \qquad (8)$$

- **Normalized root-mean-square error (NRMSE**) over the time histories.

Where repeated tests are available, mean values and standard deviations are calculated, and 95% confidence intervals are derived under the assumption of approximately normal distribution of measured peaks.

4. Results and discussion

4.1 Model Verification and Validation Results

The ANSYS numerical model was preliminary tested with a static vertical load of 2.0 kN at the centreline of the pipe to determine its capability to give classical predictions of the beam-theory. The value of the tip deflection of the horizontal arm was compared with the one calculated in a similar formulation of a cantilever beam with the nominal geometric and material properties. As demonstrated in Table 5, the disparity between the analytical estimate and the finite element ANSYS (FE) prediction is within 5% due to the fact that the difference is acceptable in a three-dimensional bracket geometry where local effects of stiffness are not due to the simple beam theory.

Table 5 Comparison of static tip deflection under 2.0 KN.

Quantity	Analytical beam model	FE model	Relative difference
Tip deflection u_{tip} (mm)	**3.20**	**3.34**	**4.4%**

ANSYS FE model modal analysis obtained the first three natural frequencies which were dominated by bracket bending and combined bending-torsion. They were compared with estimate values of the experimental values of bracket tip low-amplitude excitation and frequency response functions. According to Table 6, the FE model somewhat overstated the first bending frequency (by approximately 6 percent) and this is possible as a result of the idealisation of the concrete base as perfectly rigid. There are discrepancies at higher modes that are below 8%, as shown in Figure 6.

Table 6 Comparison of natural frequencies (numerical vs. experimental).

Mode description	f_{exp} (Hz)	f_{FE} (Hz)	Difference
1st vertical bending	**26.4**	**28.0**	**+6.1%**
2nd vertical bending / local	**73.2**	**76.5**	**+4.5%**
1st torsion + bending coupling	**121.0**	**112.8**	**−6.8%**

To test it dynamically, experimentally and numerically, a harmonic force of 1.0 kN amplitude at 26.4 Hz (measured first mode) was used. There was a variation of 7.3 in the peak vertical displacement at the tip of the bracket and variation of 9.1 in the RMS strain at the weld toe. These discrepancies along with the static and modal comparisons prove that the FE model is accurate enough to explore the multiaxial stress and deformation behaviour under more complicated variable loads as it is the case in this study.

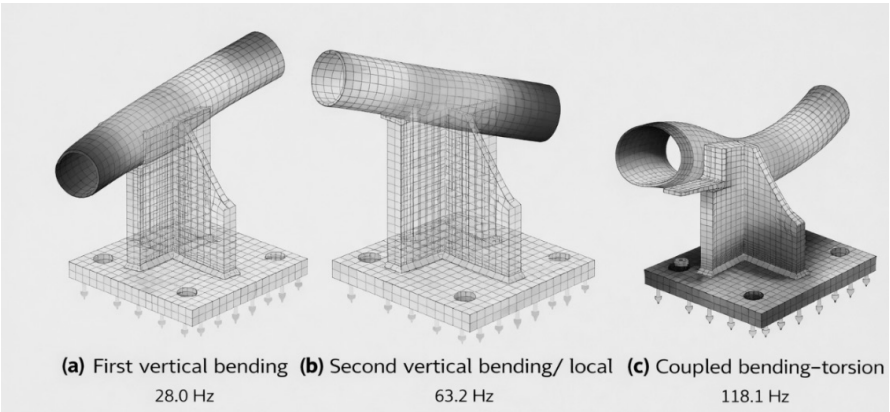


Figure 6 – Mode Shapes of the Bracket.

4.2 Global Dynamic Response of the Structures

The reaction of the global response of the bracket under three loading conditions (L1-L3, Table 3) was investigated both in terms of tip displacement and acceleration history and frequency-domain aspects.

With the operation harmonic loading (L1), and with a dominating frequency of 30 Hz near the first bending mode, the response has a resonant-type shape, with quickly stabilized steady-state oscillations. The highest vertical tip displacement is about 5.7 mm with RMS of 4.1 mm throughout the 60 s period. In the random broadband excitation (L2) the response is not as regular, and bursts are intermittent and relate to the energy content at the natural frequencies; the maximum displacement in this case is lower (3.2 mm) although the RMS level is important because of the constant excitation in a broad band. The base excitation (L3) that is of lower frequency content generates many fewer peak accelerations, but greater deflections at low frequency, (See Table 7).

Table 7 Summary of tip response for different loading scenarios.

Scenario	Response quantity	Value
L1	$u_{\mathrm{tip,max}}$ **(mm)**	**5.7**
	$u_{\mathrm{tip,RMS}}$ **(mm)**	**4.1**
	$a_{\mathrm{tip,max}}$ **(m/s^2)**	**52.3**
L2	$u_{\mathrm{tip,max}}$ **(mm)**	**3.2**
	$u_{\mathrm{tip,RMS}}$ **(mm)**	**2.0**
	$a_{\mathrm{tip,max}}$ **(m/s^2)**	**35.6**
L3	$u_{\mathrm{tip,max}}$ **(mm)**	**4.4**
	$u_{\mathrm{tip,RMS}}$ **(mm)**	**2.7**
	$a_{\mathrm{tip,max}}$ **(m/s^2)**	**18.9**

The **dynamic amplification factor** (DAF) is defined as

$$\mathrm{DAF} = \frac{u_{\mathrm{tip,max,dyn}}}{u_{\mathrm{tip,stat}}}, \qquad (9)$$

where $u_{\mathrm{tip,stat}}$ is the static tip deflection under 2.0 kN. For L1, using $u_{\mathrm{tip,stat}} = 3.34$ mm and $u_{\mathrm{tip,max,dyn}} = 5.7$ mm, a DAF ≈ 1.71 is obtained, indicating appreciable dynamic amplification when the excitation frequency lies close to the first bending mode. For L2 and L3, DAF values of approximately 0.96 and 1.32, respectively, are obtained, reflecting the less resonant nature of these loading cases (Figure 7).

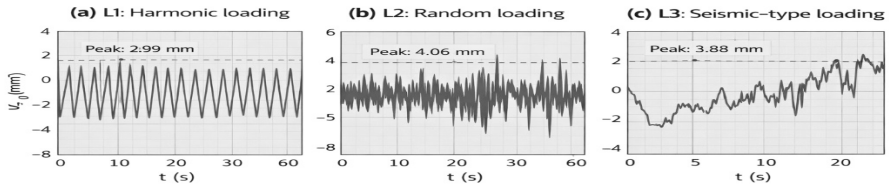


Figure 7 – Tip Displacement Time Histories

4.3 Multiaxial Stress Distribution

The stress distribution in the bracket was assessed as a multiaxial stress field and the principal stress path at specific time intervals at which the global response was at its peak. L1 and L1, the von Mises stress is maximum at the weld toe on the upper surface of the horizontal arm near the junction with the vertical web. The stress field is highly non-uniform and has a large contribution of both bending and secondary local effects. In the case of L2 and L3, locations of hot-spots are similar in nature with varying relative magnitudes and stress path characteristics.

Table 8 summarises the peak von Mises stresses at the critical weld toe for each scenario, together with the corresponding nominal bending stress σ_{nom} calculated from a simplified section analysis. The stress concentration factor is defined as

$$K_t = \frac{\sigma_{\text{eq,max}}}{\sigma_{\text{nom}}}. \qquad (10)$$

Table 8 Maximum equivalent stresses and stress concentration factors at the critical weld toe.

Scenario	σ_{nom} (MPa)	$\sigma_{\text{eq,max}}$ (MPa)	K_t
L1	108	242	2.24
L2	81	171	2.11
L3	95	205	2.16

The values of K_t between 2.1 and 2.3 highlight the **significant local amplification** of stresses at the weld toe compared with nominal bending estimates. Examination of the principal stress components reveals that, at the time of peak von Mises stress in L1, the major principal stress σ_1 at the weld toe is tensile and oriented approximately 15°–20° from the arm axis, while the intermediate principal stress σ_2 is compressive, indicating a multiaxial state governed by combined bending and local constraint, see Figures (8 and 9).

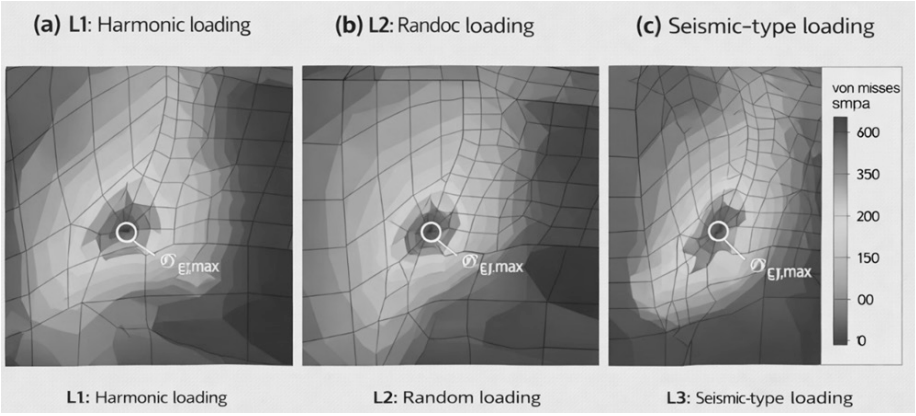


Figure 8 – Von Mises Stress Contours at Peak Response.

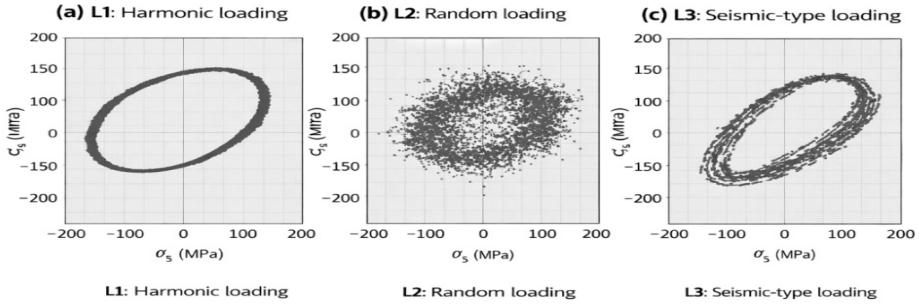


Figure 9 – Principal Stress Paths at the Weld Toe.

4.4 Deformation Patterns and Strain Fields

Patterns of deformation were observed by visualisation of deformed shapes of the bracket at representative time instants and strain fields at the weld toe. At peak response under L1, the bracket takes on a smooth global bending shape that is typical of the first mode, with the highest rotation occurring at the point of junction of the horizontal and vertical plates. Over-laid on top of this global mode is local distortions along the welded area, and this leads to the stress concentrations observed. Strain histories at the weld toe using the four strain gauges were measured and compared with FE predictions at the weld toe. Table 9 is a summary of

experimental and numerical results of peak principal strains in L1. The agreement falls within +-10% which can be accepted based on fabrication tolerances, residual stresses and idealisation in the numerical model.

Table 9 Comparison of peak principal strain at weld toe under L1.

Gauge location	$\varepsilon_{1,\exp,\max}$ (με)	$\varepsilon_{1,\mathrm{FE},\max}$ (με)	Difference
Upper surface, left side	1850	2010	+8.6%
Upper surface, right side	1790	1945	+8.7%
Side face, left	1320	1205	−8.7%
Side face, right	1275	1180	−7.5%

In random excitation (L2) the strain histories are more irregular with large occasional peaks than with purely harmonic response, although RMS levels are the same. This phenomenon is linked to temporary constructive interference of various frequency components. In the seismic-type excitation (L3), the strains are produced slowly with the expansion of the base acceleration envelope and the response is pre-eminently low-frequency deformation; localised high-frequency content is quite small, leading to smaller peak strain than L1 with similar tip deflections.

The contour plots of von Mises strain indicate that the total deformation of the material is mainly flexural; however, that the local strain field at the weld toe remains highly multiaxial with considerable shear values. This supports the significance of full field approach to the capture of critical regions rather than depending on the global displacement and nominal stress parameters, as shown in figures (10 and 11).

In general, there was a good agreement in the comparison of tests and simulations: the first natural frequency varied by 6%, the displacement of the dynamic tip under harmonic loading by approximately 7 percent, and the maximum principal strains at the weld toe by less than 10 percent. These comparisons are in

line with the simplifications made in modelling (rigid base, ideal weld geometry), and validate that the ANSYS model can be extended to stress and deformation behaviour extrapolation under other loading scenario.

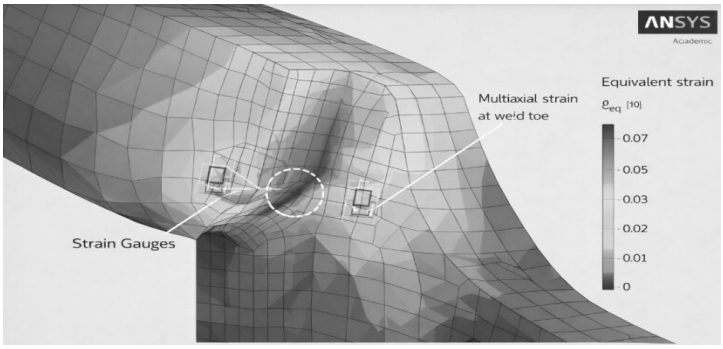


Figure 10 – Deformation Shape and Stran Field near the Weld.

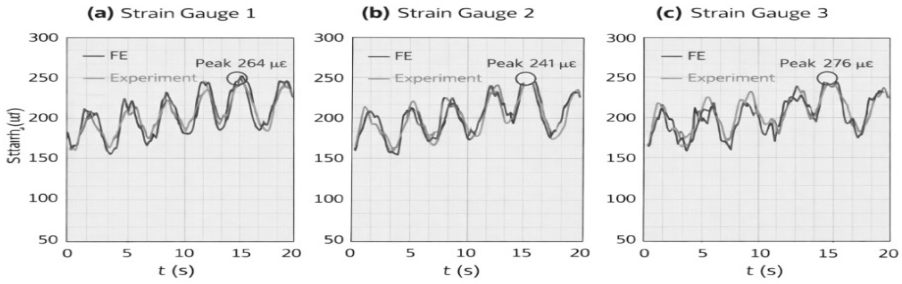


Figure 11 – Numerical vs. Experimental Strain Comparison.

4.5 Parametric and Sensitivity Analyses

A series of **parametric studies** were performed to examine the sensitivity of the bracket response to (i) load amplitude and frequency content, (ii) loading directionality, and (iii) material stiffness. For harmonic loading, the amplitude envelope $A_v(t)$ was scaled by a factor λ_A = 0.5, 1.0, 1.5 and 2.0, while the excitation frequency was varied around the first natural frequency with $f = 24, 26, 28, 30, 32$ Hz. Table 10 presents the peak von Mises stress at the critical weld toe for selected

combinations of amplitude and frequency, normalised by the value at ($\lambda_A = 1.0, f = 28$ Hz).

Table 10 Normalised peak von Mises stress at weld toe for varying amplitude and frequency.

λ_A	f (Hz)	$\sigma_{\text{eq,max}} / \sigma_{\text{eq,max,ref}}$
0.5	**24**	**0.42**
0.5	**28**	**0.51**
1.0	**24**	**0.83**
1.0	**28 (ref)**	**1.00**
1.0	**30**	**0.96**
1.5	**28**	**1.47**
2.0	**28**	**1.92**
2.0	**32**	**1.56**

The results show **approximately linear scaling** of peak stress with amplitude scaling factor λ_A, as expected for a linear system, but also highlight a pronounced frequency dependence. Peak stress is largest near the first natural frequency (28 Hz) and decreases when the excitation frequency is moved away from resonance, even when amplitude is kept constant.

Loading directionality was investigated by varying the ratio between horizontal and vertical force components in L1, defined as

$$\eta = \frac{F_H}{F_V}.$$

Values of $\eta = 0.0, 0.2, 0.4$ were considered. Increasing η from 0.0 (pure vertical) to 0.4 increased the peak shear stresses at the weld toe by about 35%, while the maximum von Mises stress increased by about 12%. This suggests that even moderate horizontal components can significantly enhance multiaxiality and should not be neglected in design assessments.

4.6 Summary of Key Quantitative Findings

To facilitate later discussion and design implications, the main quantitative outcomes of the analysis are summarised in Table 11.

Table 11 Summary of key response quantities.

Quantity	Value / range
Static tip deflection under 2.0 kN	3.34 mm (FE)
1st natural frequency (bending)	28.0 Hz (FE), 26.4 Hz (exp)
Peak tip displacement (L1 / L2 / L3)	5.7 / 3.2 / 4.4 mm
Peak tip acceleration (L1 / L2 / L3)	52.3 / 35.6 / 18.9 m/s²
Dynamic amplification factor (L1 / L2 / L3)	1.71 / 0.96 / 1.32
Peak von Mises stress at weld toe (L1 / L2 / L3)	242 / 171 / 205 MPa
Stress concentration factor K_t (range)	2.1–2.3
FE–experiment difference in peak strain (L1)	≤ 10%

4.7 Comparison with Existing Studies and Theoretical Expectations

The current outputs are generally in line with the trends that have been documented in the literature on multiaxial fatigue in welded details and components with dynamic loads. The localisation of maximum equivalent stresses at the weld toe and intensity of the concentration factor of stress (K t =2.12.3) are in agreement with hot-spots stress analysis in the welded ship and offshore structures where the fatigue-prone areas are normally located in the geometric discontinuities and weld toes when they are subjected to loading due to waves [1,3]. The dominance of the global response in bending modes and the constraint effects of the local weld area also agree with the observation done on structural brackets and attachments under cyclic loads in marine and civil use.

The qualitative behaviour of dynamic amplification around the first natural frequency, and non-proportional stress trajectories when subjected to random excitation are well compared to review of multiaxial variable-amplitude loading, highlighting the fracture sensitivity to frequency content, phase relations and mode coupling [5,7]. Equally, the observation that random broadband loading may sometimes produce peak stresses equal or somewhat greater than those with purely harmonic loading is consistent with examples that stochastic excitation can cause transient overloads at similar levels of RMS [8].

Modelling assumptions and test conditions can be attributed to some quantitative deviations of existing models. E.g. the small difference (5-7), between the numerical and experimental natural frequencies (5-7) and the peak strains (10) can be explained by the ideality of the concrete support as perfectly rigid, the omission of weld geometry imperfections and the omission of residual stresses in the fabrication process. Unlike critical-plane or energy-based methods of life prediction, which take into account plasticity and mean stress effects [9,10,11],

4.8 Implications for Design and Structural Integrity

Design wise, it is summarized in the findings that brackets and other that are used as support in gas and process plants cannot be effectively evaluated under the purely uniaxial or static-equivalent load cases. Multiaxial stress states developed at the weld toes as a result of dynamic amplification and local geometric effects are of equal magnitude at the same level of stress as at service-type loads that could be a significant portion of the yield strength. This finding aligns with the latest multiaxial fatigue techniques, which caution that similar uniaxial stress levels of loading could underrate the effect of non-proportional loading and variable-amplitude loading [5,7,10,12]. In the case of Iraqi installations in the areas with moderate seismicity

and high operational vibration, it seems to be the main requirement to apply load combinations that clearly incorporate the contribution of dynamic amplification effects to ensuring the preservation of reasonable safety factors.

The results indicate that a design code and project dependent guidelines must consider more on the hot-spot stress evaluation and dynamic response analysis of the welded supports particularly where the frequency of excitation coincides with structure modes. The underestimation of stress concentration and multiaxiality cannot be compensated by conservatively selected partial safety factors provided that design checks are carried out at nominal stress level only. Engineers are advised to ensure that the local equivalent stress levels at the weld toes under the dynamic load spectra are less than fatigue endurance levels based on multiaxial S-N curves, not just based on the global reaction forces or the weight of the pipe. In the cases where it is not possible, other detailing like higher weld toe radius, more stiffeners, or more fatigue-class weld details can be justified [1,8,13].

4.9 Limitations, Recommendations and Future Study

This research is prone to a number of limitations which should be considered when accepting the findings and using them in design. Its numerical model is linear elastic, isotropic and the weld geometry is idealised, with no consideration of residual stresses, local welding defects or possible cyclic plasticity at hot-spots. Experimental validation can only be performed with one bracket setup and to some extent of excitation that is available on laboratory equipment, hence the entire range of in-service conditions in Iraqi gas plants cannot be well represented. In addition, the analysis is based on the measures of stress and deformation, but it does not conduct a complete assessment of fatigue life, including critical-plane and damage-accumulation methods, which have proven to be significant in multiaxial variable-amplitude loading [7,8,10,14].

The next generation work is then to expand the current model to nonlinear material models and express weld toe geometry, which might be backed with comprehensive measurements of residual stress and weld profile in the practice of local fabrication. The next logical step is to combine the results of computing the multiaxial stress histories with pre-existing critical-plane or energy-based fatigue parameters to approximate life under operation and seismic load spectra. Multi-specimen tests involving manipulated geometry, weld quality and support flexibility would enable statistical characterisation of variability, and give a foundation on which reliability-based design advice is to be given according to the Iraqi conditions. Lastly, incorporation of the modelling strategy into system level of complete pipe racks or support networks would aid in uncovering interactions between supports and inform more resolute layout and detailing strategies of multiaxially loaded infrastructure.

5. Conclusion

The current paper has provided a combined numerical-experimental study on the topic of multiaxial stress and deformation of a welded steel support bracket to a gas pipeline that is representative of the Iraqi facilities due to varying dynamic loads. The dynamic response of the global structure was characterised by a three dimensional finite element model, which was tested against beam theory and laboratory data, of the local stress and deformation patterns under operational harmonic random broadband and seismic-type excitation. It was determined that dynamic amplification, which is highly pronounced at the first bending frequency, occurs with growing deflections of the tip, as compared to the unstirred scenario and peak von Mises stresses at the end of the weld of up to 200-250 MPa. This was supported by the hot-spot analyses which revealed the stress concentration factors

to be approximately 2.1-2.3 and as strongly multiaxial with occasionally non-proportional stress paths, especially when random loading occurred. The results of these studies point to the inadequacy of conventional uniaxial, static-equivalent design checks and to the necessity of dynamic, multiaxial checks of welded supports in practice. The methodology and observations notwithstanding the fact that the research was limited to linear-elastic behaviour and a single bracket setup, the study and its findings are solid foundation in conducting a more comprehensive fatigue testing, fine tuning detailing and developing design advice specific to the dynamically loaded support structures of Iraqi gas and process facilities.

6. References

[1] Corigliano, P., Ronca, P., Locatelli, L., Piscopo, V., & Arena, F. (2024). Fatigue overview of ship structures under induced wave loads. Journal of Marine Science and Engineering, 12(9), 1608. https://doi.org/10.3390/jmse12091608.

[2] Dong, P. (2001). A structural stress definition and numerical implementation for fatigue analysis of welded joints. International Journal of Fatigue, 23(10), 865–876. https://doi.org/10.1016/S0142-1123(01)00055-X.

[3] Nabuco, B., Silva, I. C., & Lupi, C. (2020). Fatigue stress estimation of an offshore jacket structure based on hot spot analysis. Mathematical Problems in Engineering, 2020, 7890247. https://doi.org/10.1155/2020/7890247.

[4] Pang, S., Tao, W., Ou, H., Liu, J., Chen, J., Liu, L., Huan, S., Pan, Z., & Huang, Y. (2025). Study on the dynamic mechanical response of orthotropic materials under complex multiaxial impact loading. Materials, 18(24), 5634. https://doi.org/10.3390/ma18245634.

[5] Anes, V., Reis, L., & Freitas, M. (2015). Multiaxial fatigue damage accumulation under variable amplitude loading conditions. Procedia Engineering, 101, 117–125. https://doi.org/10.1016/j.proeng.2015.02.016.

[6] Carpinteri, A., Spagnoli, A., & Vantadori, S. (2017). A review of multiaxial fatigue criteria for random variable amplitude loads. Fatigue & Fracture of Engineering Materials & Structures, 40(6), 891–909. https://doi.org/10.1111/ffe.12619.

[7] Deng, Q.-Y., Zhu, S.-P., He, J.-C., Li, X.-K., & Carpinteri, A. (2022). Multiaxial fatigue under variable amplitude loadings: Review and solutions. International Journal of Structural Integrity, 13(3), 349–393. https://doi.org/10.1108/IJSI-03-2022-0025.

[8] Tao, Z.-Q., *et al.* (2023). Multiaxial fatigue life estimation based on weight-averaged maximum damage plane under variable amplitude loading. Journal of Materials Research and Technology, 24, 2068–2080. https://doi.org/10.1016/j.jmrt.2023.01.196.

[9] Choi, J., *et al.* (2023). A methodology to predict the fatigue life under multi-axial loading for CFRP structures. Materials, 16(5), 1952. https://doi.org/10.3390/ma16051952.

[10] Shen, B., Liu, M., Wang, Y., & Zhuge, H. (2025). Research on fatigue strength evaluation method of welded joints in steel box girders with open longitudinal ribs. Crystals, 15(7), 646. https://doi.org/10.3390/cryst15070646.

[11] Cao, C., He, X., Ren, F., & Li, Z. (2023). Structural vibration analysis of long-distance onshore pipelines with multi-point supports considering randomness of flow. Thin-Walled Structures, 191, 110501. https://doi.org/10.1016/j.thinwalled.2023.110501

[12] Tripathi, R., & Jadhav, K. (2024). Vibration analysis of piping connected with shipboard equipment. Frontiers in Mechanical Engineering, 10, 1396170. https://doi.org/10.3389/fmech.2024.1396170

[13] Wu, R., Pan, H., Liao, J., Huang, C., Jiang, B., & Zhao, Z. (2023). Techniques and methods for vibration control of fluid-conveying pipes: A systematic review. Archives of Applied Mechanics, 93, 4753–4784. https://doi.org/10.1007/s10483-023-3023-9

[14] Pagliari, L., & Susmel, L. (2025). A review of multiaxial low-cycle fatigue criteria for life prediction. International Journal of Damage Mechanics. Advance online publication. https://doi.org/10.1177/10567895241280788

Guided Mind Mapping for Enhancing Spatial Thinking in First-Year Architectural Design Studios / Original Article

Assist. Lact. Noor Ayad Rashid[1*]

and Lact. Dr. Mustafa Kamil Al-Khafaji[2]

1 MSc. Researcher, College of Engineering Architecture, University of Technology, Baghdad/Iraq,

2 PhD., College of Engineering Architecture, University of Technology, Baghdad/ Iraq,

* Corresponding Author: ae.23.5@grad.uotechnology.edu.iq

الخرائط الذهنية الموجَّهة لتنمية التفكير المكاني لدى طلبة السنة الأولى في استوديوهات التصميم المعماري

م.م. نور أياد رشيد[1] و م. د. مصطفى كامل الخفاجي[2]

1 باحثة ماجستير، كلية الهندسة المعمارية، الجامعة التكنولوجية، بغداد/العراق

2 كلية الهندسة المعمارية، الجامعة التكنولوجية، بغداد/العراق

Abstract

Spatial thinking constitutes a core cognitive ability in early architectural design education, enabling students to interpret, organize, and transform spatial relationships into coherent architectural configurations. Despite its importance, studio practice often indicates that first-year students rely on surface-level representation rather than engaging in structured spatial reasoning, which leads to fragmented compositions and weak spatial coherence.

Accordingly, this study investigates the following research question: *To what extent can guided mind mapping support the development of spatial thinking in early architectural design studios*? A quasi-experimental design was adopted to examine the impact of structured visual–cognitive guidance on students' spatial reasoning. The study involved two groups of first-year architecture students: an experimental group trained using guided mind mapping (n = 133) and a control group receiving conventional studio instruction (n = 127).

Students' performance was evaluated through a house design task using standardized descriptive–quantitative assessment rubrics, and statistical analysis was conducted using paired-samples and independent-samples t-tests at a significance level of $\alpha = 0.05$. The results reveal a statistically significant improvement in the experimental group (Sig. = 0.000), with mean scores increasing from 26.761 in the pretest to 89.445 in the posttest, whereas the control group demonstrated a limited and non-significant improvement (Sig. = 0.089). Although posttest differences between the two groups did not reach statistical significance (Sig. = 0.068), qualitative analysis indicates that the experimental group achieved higher levels of spatial coherence, clearer organization of spatial relationships, and stronger integration between form and function.

The findings confirm that guided mind mapping operates as a cognitive mediator that supports the transition from spatial perception to structured spatial reasoning. It enhances students' ability to regulate spatial transformations and contributes to the production of more coherent and conceptually grounded architectural design outcomes in early design education.

Keywords: Architectural education, spatial thinking, mind mapping, design cognition, Visual–cognitive frameworks, design studio

المستخلص

يمثل التفكير المكاني إحدى القدرات المعرفية الأساسية في تعليم التصميم المعماري المبكر، إذ يتيح للطلبة تفسير العلاقات الفراغية وتنظيمها وتحويلها إلى تكوينات معمارية متماسكة. وعلى الرغم من أهميته، تكشف ممارسات الاستوديو أن طلبة المرحلة الأولى غالبًا ما يعتمدون على التمثيل الشكلي السطحي بدلًا من تبني تفكير مكاني منظم، مما يؤدي إلى تكوينات مجتزئة وضعف في الترابط الفراغي.

وعليه، تهدف هذه الدراسة إلى الإجابة عن السؤال الآتي: إلى أي مدى يمكن للخرائط الذهنية الموجَّهة أن تسهم في تنمية التفكير المكاني لدى طلبة التصميم المعماري في المراحل المبكرة؟ تم اعتماد منهج شبه تجريبي لتحليل أثر التوجيه البصري–المعرفي المنظم على قدرات الطلبة. شملت الدراسة مجموعتين من طلبة المرحلة الأولى: مجموعة تجريبية دُرِّبت باستخدام الخرائط الذهنية الموجَّهة ($n = 133$)، ومجموعة ضابطة تلقت التعليم التقليدي في الاستوديو ($n = 127$).

تم تقييم أداء الطلبة من خلال تمرين تصميم بيت باستخدام معايير تقييم وصفية–كمية موحدة، كما تم إجراء التحليل الإحصائي باستخدام اختبار (T) للعينات المرتبطة والمستقلة عند مستوى دلالة ($\alpha = 0.05$). أظهرت النتائج وجود تحسن دال إحصائيًا لدى المجموعة التجريبية ($Sig. = 0.000$)، إذ ارتفع متوسط الأداء من (26.761) في الاختبار القبلي إلى (89.445) في الاختبار البعدي، في حين أظهرت المجموعة الضابطة تحسنًا محدودًا غير دال إحصائيًا ($Sig. = 0.089$). وعلى الرغم من عدم وجود فروق دالة إحصائيًا بين المجموعتين في القياس البعدي ($Sig. = 0.068$)، إلا أن التحليل النوعي أظهر تفوقًا واضحًا للمجموعة التجريبية من حيث تنظيم العلاقات المكانية، واتساق التكوين الفراغي، والتكامل بين الشكل والوظيفة .

تؤكد هذه النتائج أن الخرائط الذهنية الموجَّهة تعمل كوسيط معرفي يدعم الانتقال من الإدراك المكاني إلى التفكير التصميمي المنظم، وتسهم في ضبط عمليات التحويل الفراغي، وتعزيز إنتاج حلول معمارية أكثر تماسكًا واتساقًا في مراحل التعليم الأولى.

الكلمات المفتاحية: التعليم المعماري، التفكير المكاني، الخرائط الذهنية، الإدراك البصري، الاطر – البصرية المعرفية ، استوديو التصميم

1. Introduction

Architectural design education in its early stages is fundamentally concerned with developing students' ability to interpret and organize spatial relationships into coherent design systems. Spatial thinking constitutes a central cognitive process through which architectural form is perceived, structured, and transformed into meaningful configurations. Within this process, architectural knowledge is not understood as isolated formal elements but as an organized system of spatial relationships, hierarchies, and functional interactions that collectively define design logic and spatial experience.

Despite this theoretical foundation, evidence from first-year design studios consistently reveals a gap between perceptual recognition and structured spatial reasoning. Students tend to engage with architectural precedents at a visual level, often reproducing formal appearances rather than interpreting the underlying relational structures that govern spatial organization. This tendency frequently results in fragmented compositions, weak spatial coherence, and limited integration between form, function, and context. Such limitations have been widely discussed in recent studies, which indicate that novice designers often struggle to manipulate spatial relationships in a systematic and controlled manner [1], [2].

Recent research in architectural education has increasingly emphasized the importance of visual–cognitive tools in supporting design thinking processes. Diagrammatic and relational representations have been shown to enhance analytical reasoning by enabling students to externalize design relationships and structure complex spatial information [2], [3]. Similarly, studies highlight that structured visual frameworks can support the transition from perceptual observation to analytical interpretation, thereby improving students' ability to understand and reorganize spatial systems within design tasks [4], [5].

Within this context, mind mapping has emerged as a promising visual–cognitive tool capable of organizing design knowledge through relational visualization. Unlike linear note-taking, mind maps represent architectural elements, spatial relationships, and governing principles within an interconnected network, allowing students to construct integrated mental models of design problems. Recent studies suggest that such structured visual representations can enhance cognitive organization, support relational reasoning, and facilitate the manipulation of design variables during early design exploration [6], [7]. However, despite the increasing use of mind mapping in educational contexts, its application as a guided and structured framework for developing spatial thinking in architectural design studios remains insufficiently explored, particularly when evaluated through performance-based design tasks. Accordingly, this study investigates the role of guided mind mapping as a visual–cognitive framework for supporting the development of spatial thinking in a first-year house design studio. The study aims to examine how structured mapping can mediate the transition from spatial perception to spatial reasoning and contribute to the production of coherent architectural design outcomes. This approach contributes to addressing contemporary challenges in architectural education by enhancing students' ability to manage increasing design complexity and structured design reasoning processes.

2. Literature Review

2.1 Spatial Thinking in Architectural Education

Spatial thinking constitutes one of the fundamental cognitive capacities underlying architectural design activity. Within architectural cognition, the interpretation of form does not rely solely on visual recognition of shapes but rather on the ability to understand spatial relations, imagine transformations, and

reorganize spatial configurations across multiple representational systems. Architectural drawings, models, and diagrams therefore function not only as visual representations but also as cognitive instruments that support spatial reasoning during design development.

Research in design cognition suggests that spatial thinking enables designers to perceive architectural form as an organized system of relationships rather than as isolated visual elements. According to Kosslyn and Thompson [8], spatial cognition involves the mental organization of spatial relations, including orientation, distance, scale, and transformation. Within architectural learning environments, these cognitive processes support the interpretation of plans, sections, elevations, and volumetric relationships that collectively structure architectural space. Ware [9] similarly emphasizes that the interpretation of spatial relations constitutes a central component of visual reasoning in design thinking.

In architectural education, the development of spatial thinking becomes particularly critical during the early stages of design learning. First-year students are required to interpret architectural precedents, understand spatial structures, and manipulate formal relationships through drawings and models. However, several studies in design education indicate that novice designers frequently struggle to interpret spatial configurations in a systematic manner. Lawson [10] explains that inexperienced students often focus on visible formal features rather than on the relational logic that organizes architectural compositions. Similar observations have been reported in more recent research on design studio learning, which identifies spatial interpretation as one of the primary challenges in early architectural education [4], [11].

Consequently, spatial thinking should be understood in architectural education as a structured cognitive competence that enables students to interpret spatial relations, manipulate architectural configurations, and generate coherent design alternatives.

Developing this competence requires pedagogical strategies capable of externalizing spatial relationships and making them cognitively accessible during the design process.

2.2 Components of Spatial Thinking in Architectural Design Learning

Within architectural cognition research, spatial thinking is typically described as a composite ability composed of several interrelated cognitive skills. These include spatial visualization, spatial orientation, spatial perception, and mental transformation. Rather than operating independently, these abilities function together as part of a broader visual–spatial reasoning system that supports design analysis and generative thinking. As illustrated in Figure 1, these components do not operate in isolation but are integrated within a unified visual–spatial reasoning system that organizes spatial principles, elements, and relationships to support design application.

Spatial visualization refers to the ability to imagine spatial configurations and mentally manipulate them across different representational formats. This ability enables architecture students to translate two-dimensional drawings into three-dimensional mental representations and to anticipate how spatial configurations may change through design modifications. Spatial orientation, by contrast, relates to the understanding of positional relationships between objects within space and the ability to interpret directional relations between architectural elements. Kosslyn and Thompson [8] note that these abilities play a central role in spatial reasoning tasks involving mental imagery and spatial transformation.

Empirical studies in architectural education further demonstrate the importance of these abilities for design learning. For example, recent research examining spatial ability among architecture students shows that spatial visualization and mental rotation significantly influence students' capacity to interpret architectural drawings and manipulate spatial configurations during design tasks [12].

Similarly, Türkmenoğlu Berkan *et al.*, [6] demonstrate that structured educational activities such as drawing exercises, representation techniques, and spatial modeling significantly enhance students' spatial perception and visualization abilities. In early design studios, however, students frequently encounter difficulties when attempting to reorganize spatial relations within architectural configurations. These difficulties do not necessarily arise from the inability to reproduce architectural forms but rather from the inability to reinterpret the relational structure that organizes them. As a result, design outcomes may vary visually while still reproducing the same spatial configuration. This challenge highlights the importance of educational approaches that explicitly support the development of spatial reasoning in architectural learning environments.

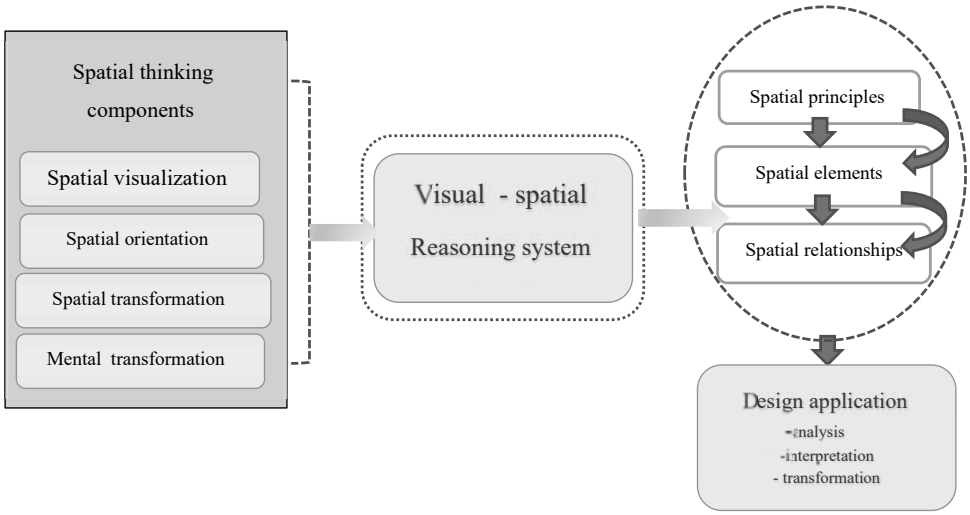


Figure 1. Conceptual model of spatial thinking components as an integrated visual-cognitive system supporting architectural design reasoning. Source: Author.

2.3 Mind Mapping as a Visual–Cognitive Tool in Architectural Education

Mind maps have increasingly been adopted in educational contexts as visual tools capable of organizing complex information through relational representation. Unlike linear note-taking structures, mind maps arrange concepts, ideas, and relationships around a central node, allowing knowledge to be represented within a spatial network of interconnected elements. This visual organization supports comprehension, memory, and conceptual reasoning by externalizing relationships that may otherwise remain implicit in verbal explanation.

Within architectural education, this characteristic makes mind mapping particularly relevant. Architectural design problems involve multiple interconnected layers including spatial organization, functional relationships, circulation patterns, structural logic, and contextual conditions. Mind maps can therefore assist students in structuring these complex relationships within a visually organized analytical framework. Salama [13] emphasizes that visual structuring tools play a significant role in design education by supporting analytical reasoning and knowledge organization within studio learning environments.

Research on visual thinking tools similarly suggests that diagrammatic frameworks help designers organize design information and explore relationships between design variables. Eppler [2] explains that visual knowledge structures such as mind maps enhance analytical reasoning by making conceptual relationships visible. When architectural elements and their relationships are visually represented, students are better able to interpret the internal structure of design problems and identify connections between spatial components.

Studies examining the use of mind mapping in design methodology further highlight its potential role in supporting design thinking processes. For example, demonstrates that mind mapping can assist designers in decomposing complex

design problems and organizing conceptual relationships during early design exploration. By arranging design variables within a relational visual structure, mind maps support the interpretation of architectural information and facilitate the generation of structured design ideas.

2.4 The Role of Mind Mapping in Developing Spatial Thinking

The relationship between mind mapping and spatial thinking lies primarily in the ability of guided visual frameworks to externalize spatial relations and support their cognitive manipulation. Spatial thinking in architecture requires the ability to identify spatial relationships, interpret spatial hierarchies, and reorganize spatial configurations during design development. Mind maps can facilitate this process by representing spatial relationships within a structured relational diagram. As illustrated in Figure 2, guided mind mapping mediates the transition from spatial perception to controlled spatial transformation.

When students analyze architectural precedents or design tasks using guided mind maps, they are encouraged to move beyond simple visual observation toward relational interpretation. Architectural elements become connected within a network of spatial relations that may include adjacency, enclosure, hierarchy, orientation, and proportional organization. Through this process, students develop a more integrated mental model of architectural space that supports spatial reasoning during design development.

Recent research in architectural education indicates that structured visual tools can enhance students' capacity to organize spatial and compositional information. Diagrammatic representations have been shown to support analytical reasoning in design studios by enabling students to visualize relationships between design elements [3], [4]. In this context, mind maps function as relational diagrams that help students interpret and reorganize spatial information within design tasks.

Furthermore, guided mind maps can support the transition from spatial interpretation to spatial manipulation. When relationships between architectural elements are explicitly represented, students are able to explore alternative spatial configurations by modifying these relationships. Such operations may involve reorganizing spatial hierarchies, altering functional relationships, or redefining circulation structures within architectural compositions. Through this mechanism, spatial thinking becomes an active process of interpretation and transformation rather than passive observation.

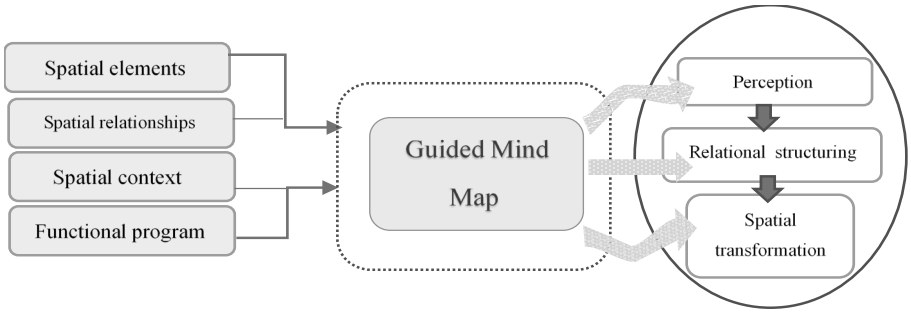


Figure 2. Cognitive mechanism of guided mind mapping in developing spatial thinking through perception, relational structuring, and spatial transformation. Source: Author.

2.5 Integrating Mind Mapping with Spatial Thinking in First-Year Design Studios

Within first-year architectural design studios, students are required to simultaneously develop skills in interpreting architectural representations and generating design proposals. This dual demand often creates cognitive overload, limiting students' ability to manage spatial information and to understand how spatial relationships inform design decisions. Consequently, early design learning environments require pedagogical frameworks capable of structuring spatial reasoning processes and supporting the organization of complex design information [4].

Guided mind mapping offers a structured visual–cognitive framework that can support this process. When applied as an analytical tool rather than as an open-ended brainstorming technique, mind maps enable students to identify architectural elements, recognize relationships among spatial components, and organize information within coherent relational structures. Such structured representations facilitate the construction of spatial mental models that underpin design reasoning and decision-making [2].

Furthermore, guided mind mapping contributes to the development of multiple dimensions of spatial thinking. The visual organization of elements and relationships enhances spatial visualization and supports the interpretation of spatial configurations. At the same time, the relational structure of the map allows students to reorganize spatial information, explore alternative configurations, and engage in controlled transformation processes during design tasks [3].

Despite these theoretical advantages, empirical research examining the cognitive impact of guided mind mapping on spatial thinking in architectural education remains limited. While recent studies have explored diagrammatic reasoning, spatial ability development, and visual cognition in design learning, fewer investigations have specifically addressed the role of structured mind mapping in supporting spatial thinking within early design studios [6]. This gap is particularly evident in Iraqi architectural education, where limited empirical work has examined the pedagogical effectiveness of visual–cognitive tools in studio-based learning environments.

Accordingly, the present study aims to address this gap by investigating the effectiveness of guided mind mapping as a structured visual–cognitive framework for developing spatial thinking among first-year architecture students in Iraqi design studios.

3. Methodology

3.1 Research Design

The study adopted a quasi-experimental research design to examine the effectiveness of guided mind mapping in developing spatial thinking among first-year architecture students in Iraqi architectural design studios .This design compares two groups of students: an experimental group exposed to a structured mind-mapping training program and a control group taught through the conventional architectural design studio approach.A pre-test and post-test structure was implemented to evaluate changes in spatial thinking abilities before and after the intervention. Quasi-experimental approaches are widely used in architectural education research because they enable the evaluation of pedagogical interventions within authentic studio environments [14], [10].

3.2 participants

The research population consisted of first-year students enrolled in the Department of Architectural Engineering at the University of Technology in Baghdad, Iraq. This educational stage represents the foundational phase of architectural training, during which students begin to develop essential competencies in visual perception, spatial reasoning, and architectural form composition. At this level, students are introduced to fundamental design principles, visual analysis, and the basic processes of architectural thinking, making it an appropriate context for examining the development of spatial thinking abilities within architectural design education.

The total sample included 260 students collected across two academic years. The first cohort consisted of 225 students from the academic year 2022–2023, while the second cohort included 35 students from the academic year 2024–2025.

Because all participants were first-year students enrolled in the same architectural program and learning environment, the sample can be considered relatively homogeneous in terms of educational background and studio learning conditions.

In terms of demographic characteristics, participants belonged to the typical university entry age range in Iraq, approximately 18–20 years, which corresponds to the standard age of first-year undergraduate architecture students. Age variation within the sample was therefore minimal and was not treated as an independent analytical variable in the study. Similarly, gender distribution was not considered a determining analytical factor, since the design tasks employed in the study evaluate spatial reasoning and visual–cognitive processes that are not dependent on gender differences.

Following the quasi-experimental research design adopted in this study, the sample was divided into two groups. The experimental group (n = 133) received the proposed instructional intervention based on guided mind mapping through a structured training program that included theoretical explanations and a sequence of design exercises aimed at developing spatial thinking abilities. In contrast, the control group (n = 127) followed the conventional architectural design studio instruction without exposure to the mind-mapping intervention, providing a baseline condition for comparison.

This group distribution provides a suitable methodological structure for the quasi-experimental research design, allowing statistical comparison between the two groups in order to evaluate the effectiveness of the guided mind-mapping approach in supporting the development of spatial thinking in architectural design education. The structure of the sample and its division into experimental and control groups is illustrated in Figure 3.

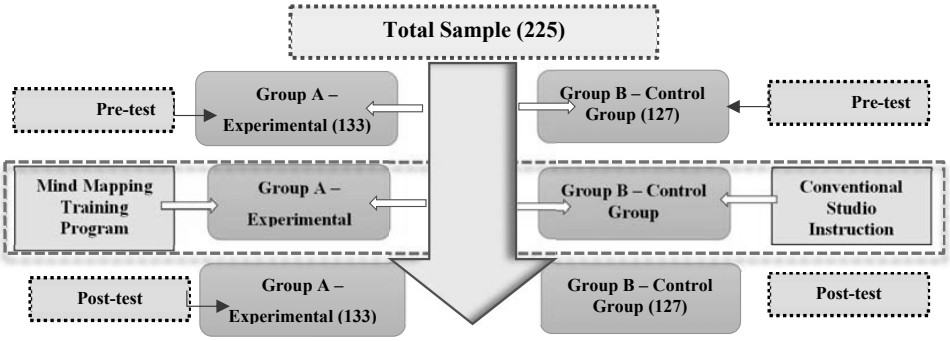


Figure 3. Distribution of experimental and control groups according to the experimental research procedure. Source: Author.

3.3 Research Variables

The present study examines the effect of guided mind mapping as an instructional intervention aimed at supporting the development of spatial thinking within the architectural design studio. In accordance with the quasi-experimental research design, two primary variables were identified: the independent variable representing the instructional intervention and the dependent variable representing students' spatial thinking performance.

To clarify the relationships between the independent and dependent variables and the role of guided mind mapping within the design learning process, a conceptual framework was developed, as illustrated in Figure4 .

The independent variable consists of the guided mind-mapping training program introduced within the architectural design studio. The program was implemented through a structured sequence of theoretical explanations and design exercises intended to organize spatial information visually and to support the interpretation of spatial relationships during architectural design tasks. Through this intervention, mind maps were employed as visual–cognitive tools that assist

students in identifying spatial elements, structuring spatial relations, and organizing architectural information within a coherent analytical framework.

The dependent variable represents the level of spatial thinking demonstrated by students during architectural design tasks. Spatial thinking in architectural education involves the ability to interpret spatial relationships, organize functional and circulation systems, and manipulate spatial configurations during the development of architectural form. In the context of the present study, spatial thinking was assessed through the analysis of students' design outputs produced during the house design exercise, which required students to translate spatial relationships into coherent architectural configurations.

Evaluation of spatial thinking performance was based on a set of observable indicators reflected in students' design representations. These indicators included the clarity of spatial organization, the effectiveness of zoning and circulation structures, the coherence of spatial relationships between architectural elements, and the ability to generate consistent architectural form based on spatial logic. Such indicators provide measurable evidence of students' ability to analyze, structure, and manipulate spatial information during the architectural design process.

Through this variable structure, the study examines the extent to which guided mind mapping contributes to improving students' spatial reasoning abilities in architectural design learning environments.

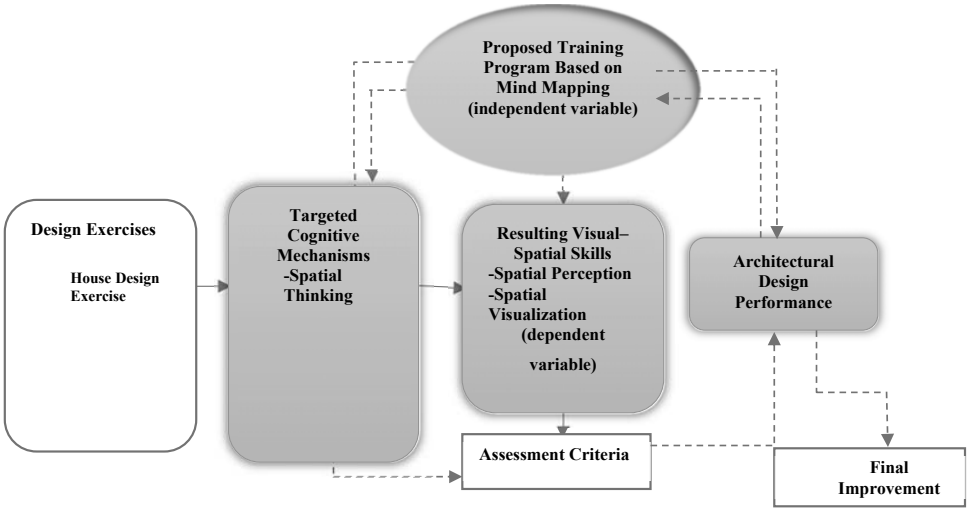


Figure 4. Conceptual framework of the training program for developing visual thinking based on mind mapping. Source: Author.

3.4 Design Tasks

The cognitive processes under investigation were examined through two complementary architectural design tasks based on the house design project, which represents a fundamental exercise in early architectural education. Unlike precedent-based analytical tasks, the selected design tasks required students to generate architectural form through the interpretation and organization of spatial relationships within a defined design scenario.

The house design project was structured to address both contemporary and traditional architectural approaches, allowing students to engage with distinct spatial logics and compositional systems. This dual framework provided a broader basis for examining the development of spatial thinking through the manipulation of spatial organization, functional zoning, and architectural form.

The design tasks were organized as follows:

1. **Modern House Design Task:**

   Students were required to develop a contemporary residential design grounded in the principles of modern architecture. The task emphasized spatial openness, functional clarity, and the integration between interior and exterior spaces. Particular attention was given to the organization of circulation systems, volumetric composition, and the articulation of architectural elements such as openings, structural systems, and façade treatments, reflecting the core principles of modern architectural design.

2. **Traditional House Design Task:**

   Students were required to design a residential unit based on the spatial and organizational principles of the traditional Baghdadi house. The task focused on spatial hierarchy, privacy gradients, and the organization of spaces around a central courtyard. Emphasis was placed on culturally embedded spatial relationships, including transitions between public, semi-private, and private domains, as well as environmental responses such as shading, enclosure, and inward-oriented spatial configurations characteristic of traditional Baghdadi architecture.

The implementation of both design tasks was conducted over a period of four weeks, with two studio sessions per week, ensuring a structured and continuous engagement with the design process. Within this timeframe, students progressed through a sequence of design stages including concept formation, generation of alternatives, evaluation, development, and final presentation, as illustrated in the architectural design process framework. This structured approach ensured that spatial thinking was examined not only as a conceptual cognitive ability but also as an applied design process manifested through students' architectural outputs across both modern and traditional design contexts.

This dual design framework, contrasting modern architectural principles with the spatial logic of the traditional Baghdadi house, provides a critical pedagogical context for examining how spatial thinking develops across different organizational systems. While modern architecture emphasizes openness, continuity, and functional efficiency, the traditional Baghdadi house relies on hierarchy, enclosure, and inward spatial organization. This contrast enables a deeper understanding of how students interpret, organize, and transform spatial relationships under varying architectural paradigms.

3.5 Guided Mind-Mapping Intervention

The training program adopted guided mind mapping as a structured visual–cognitive framework for supporting the development of spatial thinking within architectural design tasks. Rather than functioning as a free-form ideation technique, mind mapping was implemented as an analytical and generative tool that regulates how spatial relationships are interpreted, organized, and transformed during the design process.

Prior to the design exercises, students were introduced to the fundamental principles of architectural design thinking, with particular emphasis on the relationship between spatial organization, function, form, and environmental context. In the modern house design task, the pedagogical introduction focused on the spatial and formal logic of modern architecture, including openness, continuity, and the integration of interior and exterior spaces. In contrast, the traditional house design task introduced the spatial structure of the Baghdadi house, emphasizing hierarchy, privacy gradients, and inward-oriented spatial organization.

Within the framework of diagram-based design reasoning, visual representations play a central role in structuring how architectural form is analyzed and manipulated. Oxman [15] explains that diagrammatic thinking enables the decomposition and recompositing of architectural systems through controlled

operations, while Salama [13] highlights the role of structured visual frameworks in supporting systematic reasoning in architectural education. Eppler [2] further emphasizes that visual knowledge-structuring tools enhance analytical thinking by making relationships between design elements explicitly visible.

Building on this perspective, guided mind maps were employed to externalize spatial relationships and organize design information within a relational framework. Students were required to identify spatial functions, circulation systems, and environmental factors, and to represent these components within structured mapping diagrams. This process supported the interpretation of spatial configurations and enabled students to recognize how architectural relationships are formed and interconnected.

Through this structured representation, mind maps facilitated the transition from spatial interpretation to design generation. Students used the mapped relationships to guide the development of design alternatives through controlled transformation operations such as reorganization, addition, subtraction, and spatial reconfiguration. Recent studies in architectural education indicate that such visual structuring tools can support the organization of spatial relationships and enhance design reasoning in studio environments [3], while Hettithanthri [4] demonstrates their effectiveness in strengthening relational interpretation in early design education.

Within this intervention, guided mind mapping functioned as a cognitive scaffold that supported the systematic development of spatial thinking by linking analytical interpretation with generative architectural design processes.

3.6 Instructional Structure and Role of Studio Mentors

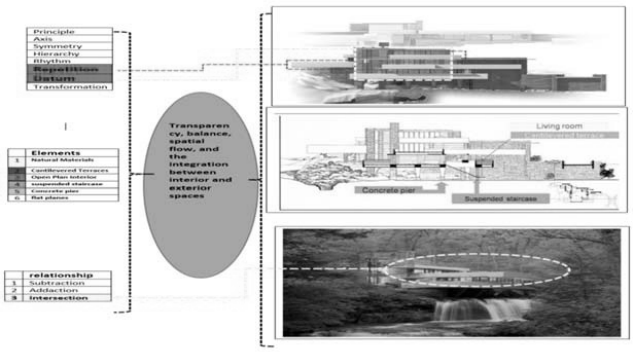
The guided mind-mapping intervention was implemented within the architectural design studio as a structured educational program supervised by studio instructors. The house design task was conducted over a period of four

weeks, allowing students to progressively develop spatial thinking through structured analytical and design processes. Each week was organized into two complementary sessions. One theoretical lecture, lasting two hours, introduced the conceptual foundations of spatial thinking, along with the architectural principles underlying both modern and traditional residential design. Within these sessions, guided mind mapping was presented as a visual–cognitive framework for organizing design knowledge and structuring spatial relationships. To support this process, students were introduced to analytical mind-mapping models derived from canonical architectural examples. As illustrated in Figures 5 and 6, guided mind maps were developed for the analysis of Fallingwater House as a representative model of modern architecture and the Baghdadi courtyard house as an example of traditional architecture. These diagrams explicitly mapped architectural form and function, including spatial organization, functional zoning, circulation systems, and the relationships between built form and context.

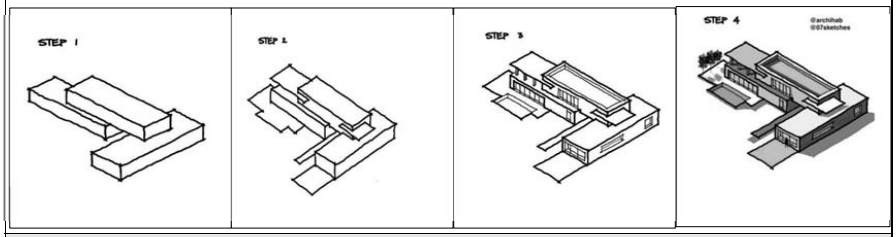
The second session was conducted within the architectural design studio and extended over six hours. During these sessions, students applied the guided mind-mapping framework to their own house design projects. The process involved structuring spatial relationships, organizing functional systems, and translating conceptual ideas into coherent spatial configurations, supported by continuous feedback and supervision.

Throughout the implementation period, studio instructors functioned as mentors facilitating the analytical process rather than directing design outcomes. Their role focused on guiding students in interpreting spatial relationships, structuring their design reasoning, and ensuring that transformation processes remained grounded in clearly identified architectural principles. This instructional structure ensured methodological consistency while fostering independent and cognitively structured design thinking.

Case study Analysis- Falling House Analysis



Concept diagram



Functional and Area/Space Program

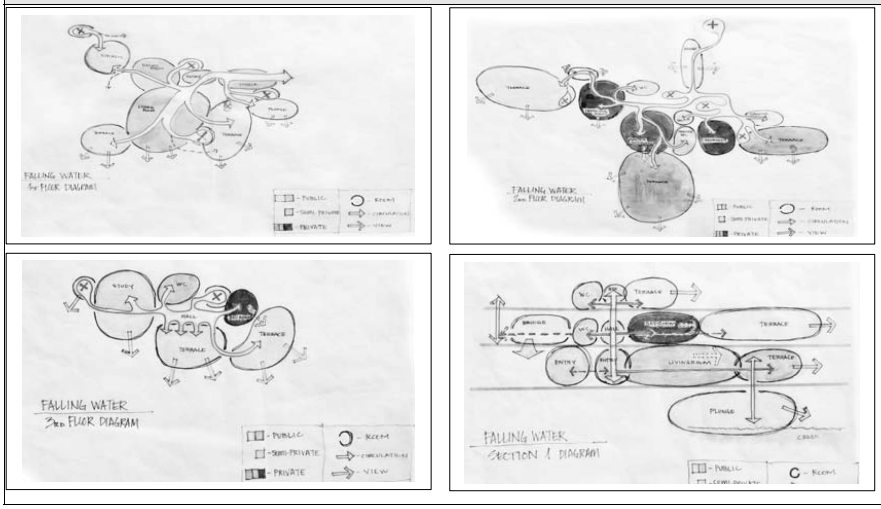
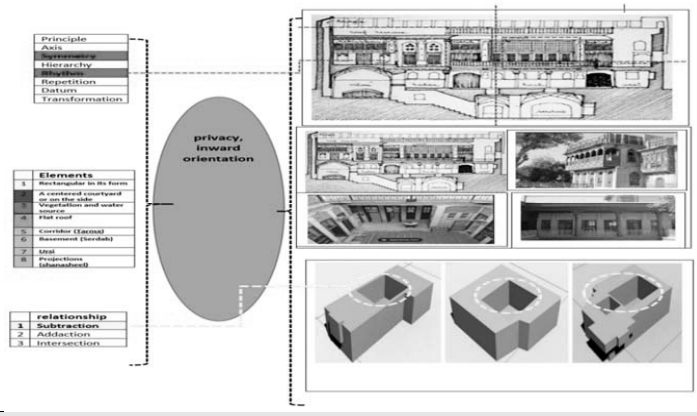
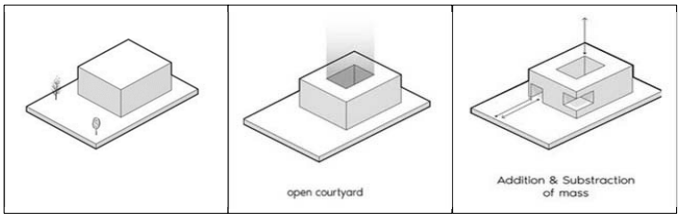


Figure 5. Guided mind-mapping framework form falling house (modern house design). Source: Author.

Case study Analysis- Falling House Analysis



Concept diagram



Functional and Area/Space Program

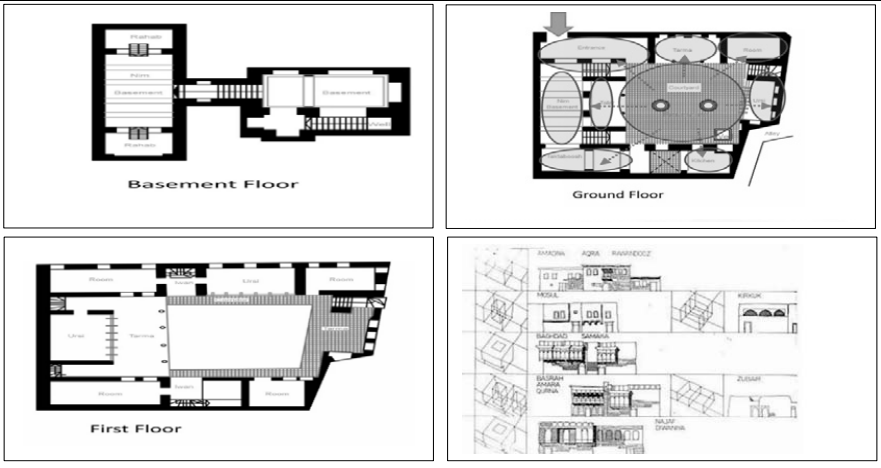


Figure 6. Guided mind-mapping framework for Baghdadi house (traditional house design). Source: Author.

3.7 Analytical Structure of the Guided Mapping Framework

The guided mind-mapping framework for the house design project was structured as a sequential cognitive–design system integrating both pre-design analysis and design development activities. This structure reflects the pedagogical organization of the studio process, in which spatial thinking is progressively constructed through interconnected analytical and generative stages. Recent studies in architectural pedagogy emphasize that structured visual–cognitive frameworks can enhance the organization of design knowledge and support the development of relational reasoning in early design education [5].

The framework was organized into two primary phases:

(1) Pre-design analytical phase

The initial phase focused on structuring students' understanding of the design problem through systematic decomposition of contextual and spatial information. This phase included four main analytical components: problem identification, site analysis, case study analysis, and functional–spatial programming. These components enabled students to interpret architectural information as a relational system rather than as isolated data points.

Within this phase, guided mind maps functioned as analytical tools that externalized relationships between spatial elements and contextual factors. Students translated site data and programmatic requirements into relational diagrams, allowing them to understand architecture as an integrated spatial system rather than a collection of isolated components. This approach aligns with recent findings indicating that diagrammatic and mapping tools support the structuring of complex spatial information and improve analytical clarity in design studios [4].

(2) Design development phase

Building upon the analytical foundation, the second phase focused on the generation and transformation of architectural form through structured design operations. This phase included concept formation, generation of design alternatives, comparative evaluation, development of the selected alternative, and final presentation.

Within this phase, guided mind mapping operated as a cognitive mediator linking analytical knowledge to design action. The mapping structure integrated three core components—form, function, and site relationships—into a unified relational system that guided design decisions. Such integration reflects contemporary perspectives in design cognition, which emphasize the role of structured visual frameworks in supporting the transition from analysis to synthesis through controlled transformation processes [3].

As illustrated in Figure 7, the conceptual scenario was structured through guided mind mapping to represent the design logic of both modern and traditional housing models. This representation integrates site context, form generation, functional organization, and spatial relationships into a unified visual–cognitive framework, enabling students to compare and reinterpret different architectural paradigms within a structured design process. Similar approaches have been shown to enhance students' ability to construct coherent design strategies and navigate between multiple spatial systems in architectural education [16].

Figure 7. Conceptual scenario of house design based on guided mind mapping (Modern house – Jadiriya riverside vs. Baghdadi house – Kadhimiya traditional), illustrating form–function–site relationships. Source: Author.

The transition between the two phases represents the shift from spatial perception to spatial reasoning and design synthesis. In this process, spatial relationships identified during the analytical phase were transformed into generative design rules, enabling controlled operations such as spatial reorganization, addition, and modification [15].

Furthermore, the integration of modern and traditional housing paradigms within the same framework allowed students to compare different spatial logics. While modern house design emphasized openness and functional continuity, the Baghdadi house focused on spatial hierarchy, privacy gradients, and courtyard-centered organization. This duality strengthened students' ability to interpret and reconstruct spatial systems across different architectural contexts.

Consequently, guided mind mapping did not function merely as a visualization tool, but as a structured cognitive framework that organized the entire design process—from problem definition to final architectural synthesis—supporting the development of coherent and contextually grounded design outcomes [15], [2], [1].

3.8 Assessment Rubrics

Student design outputs were evaluated using a structured descriptive–quantitative rubric to ensure consistency and comparability across both the modern and traditional house design tasks[17]. The assessment framework was aligned with the cognitive objectives of the study, focusing on the development of spatial thinking through three interrelated dimensions: `architectural analysis, visual representation, and architectural form synthesis.

The first dimension, architectural analysis, assessed students' ability to interpret the design context through site analysis, precedent studies, and functional program definition. This included the identification of spatial elements, environmental conditions, and functional relationships, enabling students to construct an integrated understanding of architectural systems.

The second dimension, visual representation, evaluated the translation of spatial and functional understanding into graphical outputs, including diagrams, plans, sections, elevations, and three-dimensional models. Particular attention was given to the use of diagrammatic tools, such as guided mind maps, to organize spatial and functional relationships.

The third dimension, architectural form synthesis, assessed the students' ability to integrate spatial and functional relationships into coherent architectural configurations. Emphasis was placed on the relationship between form and function, the organization of interior–exterior connections, and the development of a consistent architectural language reflecting either modern or traditional design approaches.

As illustrated in Table (1), architectural analysis and architectural form synthesis were each assigned a weight of 35%, reflecting their central role in structuring and integrating both spatial and functional relationships. Visual representation was assigned a weight of 30%, capturing its role in externalizing and communicating design thinking.

Table 1: Assessment weights aligned with intervention effect

Main dimension	Sub-dimension	Weight (%)
Architectural analysis	Site analysis	35
	Analysis of comparable precedents	
	Functional program analysis and spatial relationships	
Mental Transformation	Analysis of AI-generated examples	30
	Preliminary drawings, sections, elevations	
	3D models and diagrammatic representation	
Visual Synthesis	Integration of interior–exterior relationships	35
	Form–function integration and spatial organization	
	Application of compositional principles (modern / traditional)	
Total		100

3.9 Statistical Analysis

Paired-samples t-tests were conducted to examine pretest–posttest differences within each group in relation to spatial thinking development in the house design exercise. Independent-samples t-tests were then employed to compare posttest performance between the experimental and control groups. Statistical significance was evaluated at an alpha level of ($\alpha = 0.05$.)The analysis focused on measuring changes in students' ability to organize spatial relationships and generate coherent spatial configurations.

3.10 Reliability and Effect Size

To ensure the consistency of performance assessment in the house design exercise, the scoring rubrics were evaluated for inter-rater reliability. Two independent architectural design instructors, trained in the use of the architectural analysis, visual representation, and architectural form synthesis rubrics, assessed a randomly selected subset of student projects. The level of agreement between raters was examined using Cohen's kappa coefficient, which indicated a high degree of consistency in scoring across both the modern and traditional house design tasks. This result supports the reliability of the rubrics in capturing students' spatial thinking performance.

Beyond statistical significance, effect size measures were calculated to determine the magnitude of the intervention's impact on spatial thinking development. Cohen's d was computed for pretest–posttest differences within each group, as well as for posttest comparisons between the experimental and control groups. The experimental group demonstrated large effect sizes in the house design task, indicating that guided mind mapping produced not only statistically significant but also practically meaningful improvements in students' ability to organize spatial relationships and generate coherent spatial configurations.

In contrast, the control group showed only small to moderate effect sizes, reflecting limited improvement primarily associated with cumulative learning rather than structured cognitive support. These findings reinforce the effectiveness of guided mind mapping as a cognitive–pedagogical framework for enhancing spatial thinking in early architectural design education.

4. Findings

4.1 Task: House Design (Spatial Thinking)

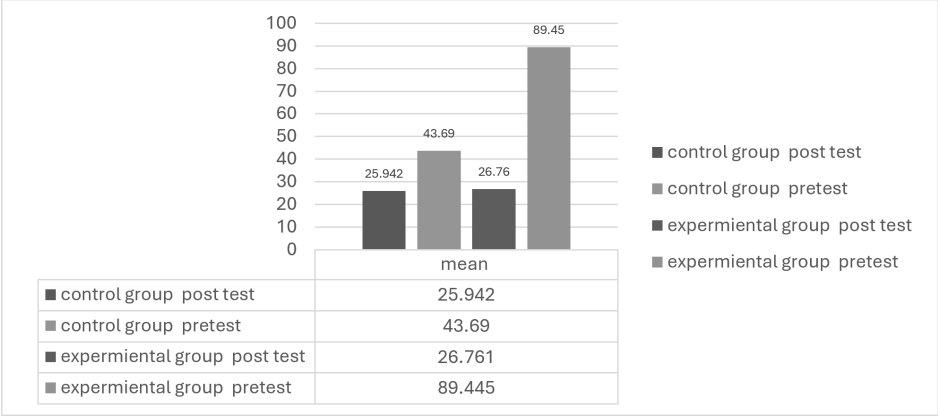
The paired-samples t-test results (Table 2) indicate a differentiated pattern between the control and experimental groups in their ability to organize spatial relationships and construct coherent spatial configurations within the house design task. This task directly reflects spatial thinking as it requires the integration of functional zoning, circulation systems, and spatial hierarchy into a unified architectural form.

Table 2. Results for House Design

Group	Test	Sample size	Mean	Standard Deviation	Paired-samples t-test value	Significance level
Control	Pretest	127	20.634	10.922	1.745	0.089
Control	Posttest	127	48.46	15.35		
Experimental	Pretest	133	21.18	11.252	13.022	0.000
Experimental	Posttest	133	75.363	14.545		

Within the control group, the mean score increased from) 25.942(in the pretest to)43.69(in the posttest; however, this change did not reach statistical significance (Sig. = 0.089) (Table 2). As illustrated in Figure 8, although the control group demonstrated a noticeable increase in mean performance, the relatively high standard deviation suggests that this improvement remained inconsistent and did not reflect a systematic development in spatial reasoning.

In contrast, the experimental group demonstrated a substantial and statistically significant improvement, with the mean score increasing from)26.761(in the pretest to)89.445(in the posttest (Sig. = 0.000) (Table 2). This pronounced shift is clearly visualized in Figure 8, where both the magnitude of the mean increase and the controlled dispersion pattern indicate a strong enhancement in students' ability to structure spatial relationships and generate coherent spatial configurations.



	mean
control group post test	25.942
control group pretest	43.69
expermiental group post test	26.761
expermiental group pretest	89.445

Figure 8. Pretest and posttest mean scores and standard deviations for house design task. Source: Author

The independent-samples t-test conducted on posttest scores (Table 3) did not indicate a statistically significant difference between the two groups (Sig. = 0.068), as shown in Figure 9. However, the experimental group achieved a considerably higher mean score compared to the control group, indicating a stronger level of spatial organization and design coherence.

From a design cognition perspective, the limited improvement observed in the control group reflects a reliance on intuitive or experience-based design decisions without a structured framework for organizing spatial relationships. In contrast, the experimental group exhibited a qualitatively different pattern of

learning. The guided mind-mapping training enabled students to explicitly structure spatial relationships, organize functional systems, and translate conceptual intentions into coherent spatial configurations.

Consequently, students in the experimental group demonstrated clearer spatial sequencing, stronger functional zoning, and more consistent integration between concept and spatial form. These findings confirm that guided mind mapping effectively supports spatial thinking as a structured cognitive process, facilitating the transformation of abstract design ideas into coherent architectural solutions.

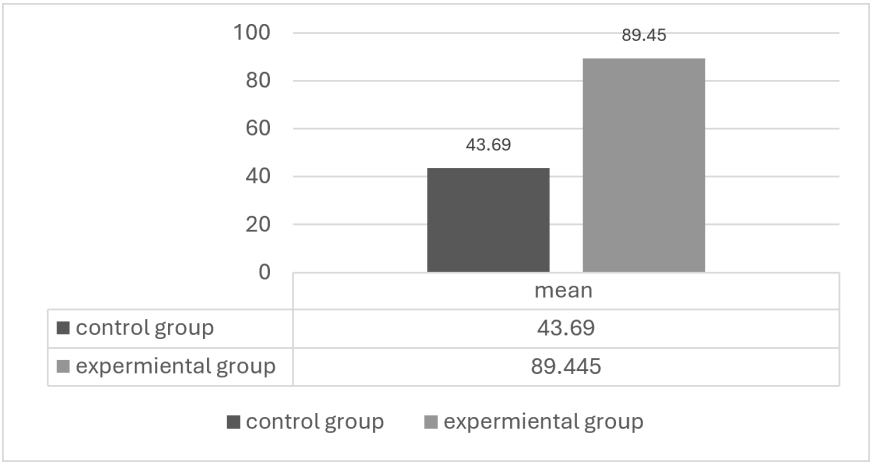


Figure 9. Posttest mean scores and standard deviations for the Villa Savoye task. Source : Author

4.2 Evidence for the Effect of Guided Mind Mapping on Spatial Thinking in House Design

The posttest results of the house design task reveal a clear and consistent pattern of improvement in the experimental group compared to the control group. As illustrated in Figure 10, students who were trained using guided mind maps achieved significantly higher mean scores (89.445) than those in the control group (43.69). This consistent performance difference indicates that the observed learning

gains are not incidental but reflect the systematic impact of the guided mind-mapping intervention on spatial thinking.

From a design cognition perspective, guided mind maps functioned as an organizational cognitive mediator that facilitated the transition from conceptual intention to spatial configuration. By externalizing spatial relationships, functional structures, and hierarchical organization, the mapping process enabled students to regulate their design thinking, structure spatial analysis, and apply controlled transformations in the development of architectural form.

This structured cognitive support enhanced students' ability to integrate functional zoning, circulation systems, and spatial hierarchy into coherent spatial compositions. As a result, the experimental group demonstrated not only higher quantitative performance but also a qualitatively more advanced level of spatial organization, reflected in clearer spatial sequences, stronger functional coherence, and more consistent alignment between concept and spatial form.

Accordingly, the pattern illustrated in Figure 10 confirms that guided mind mapping strengthens spatial thinking as an underlying cognitive process in early architectural design education. Rather than merely improving final design outcomes, the intervention supports the regulation of design reasoning processes, enabling students to systematically organize and transform spatial relationships into coherent architectural solutions.

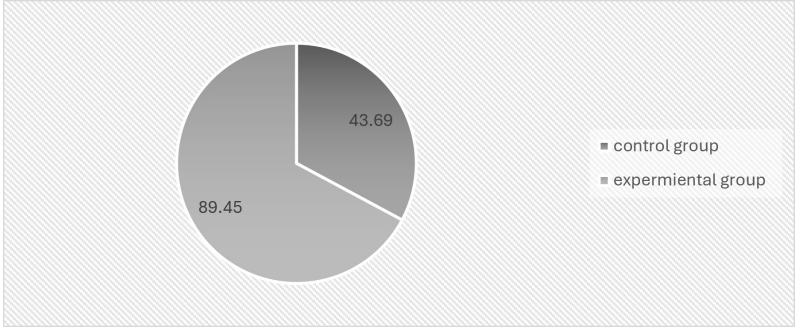


Figure 10. Posttest mean scores for the control and experimental groups for house design tasks. Source : Author

5. Discussion

The findings of the present study indicate that guided mind mapping functions as an effective visual–cognitive mechanism for enhancing spatial thinking within early architectural design education. The observed improvement in the experimental group reflects a shift from intuitive and fragmented spatial understanding toward structured spatial reasoning supported by explicit relational organization. These findings are particularly relevant in the context of contemporary architectural challenges, where increasing design complexity requires structured cognitive frameworks to support design reasoning and decision-making[18].

From a cognitive perspective, the guided mapping framework contributed to improving students' ability to interpret architectural space as an organized system of relationships rather than a collection of isolated elements. Students demonstrated an increased capacity to identify spatial hierarchies, organize functional zones, and structure circulation patterns within coherent spatial configurations. This transformation suggests that spatial understanding was no longer limited to visual recognition but extended to relational interpretation, a process emphasized in recent research on spatial cognition in design education [12].

The results further suggest that externalizing spatial relationships through guided mind maps facilitated the transition from spatial perception to spatial manipulation. By representing elements, relationships, and governing principles within a structured visual framework, students were able to explore alternative spatial configurations and reorganize design components more systematically. This process reflects the role of diagrammatic thinking in enabling controlled transformation and generative reasoning, as highlighted in contemporary studies on visual–cognitive tools in design learning [4].

In contrast, the limited improvement observed in conventional studio settings indicates that exposure to design tasks alone is insufficient for developing spatial reasoning skills. Without structured cognitive support, students tend to rely on surface-level imitation and struggle to reorganize spatial relationships within design problems. This finding aligns with recent research suggesting that novice designers often face difficulties in transitioning from perceptual observation to relational reasoning when explicit analytical frameworks are not provided [3].

Although both groups demonstrated improvement in spatial performance, the qualitative characteristics of the experimental group's outputs reveal a more advanced level of spatial organization. Design solutions exhibited clearer spatial sequencing, stronger functional integration, and greater consistency between conceptual intention and spatial configuration. This indicates that guided mind mapping influenced not only performance outcomes but also the internal structure of spatial reasoning processes.

From a design cognition standpoint, these findings support the interpretation that spatial thinking develops through the integration of perception, representation, and transformation within structured cognitive frameworks. Guided mind maps functioned as mediating tools that organized this integration by making spatial relationships explicit and operable, as emphasized in recent studies on knowledge visualization and cognitive structuring [19]. Such mechanisms are consistent with recent perspectives emphasizing the role of visual–cognitive scaffolding in supporting systematic design reasoning [16].

Educationally, the study highlights the importance of incorporating structured visual tools within early architectural design studios. Spatial thinking should be addressed as a core cognitive skill that requires explicit pedagogical support rather than being treated as an implicit outcome of design practice. The guided mapping

approach provides a structured methodology for supporting students in organizing spatial knowledge and developing coherent design solutions [5].

Nevertheless, the findings should be considered within the scope of the study. The investigation was conducted within a specific educational context and focused on first-year students, which may limit the generalizability of the results. Future research may extend this approach to different educational levels and design contexts in order to further explore the broader applicability of guided mind mapping in architectural education.

Overall, the findings suggest that guided mind mapping strengthens the cognitive infrastructure underlying spatial thinking by enabling students to interpret, organize, and transform spatial relationships within a structured design process. This contribution extends beyond performance improvement, supporting a deeper transformation in design cognition during early stages of architectural education.

6. Conclusions

The findings indicate that guided mind mapping can function as an effective visual–cognitive framework for supporting the development of spatial thinking in early architectural design education. Structuring spatial information through explicit relational representation enhances students' ability to organize, interpret, and transform architectural space within coherent design systems.

These results highlight the potential of guided visual frameworks to systematically support the development of spatial reasoning, moving beyond implicit learning approaches commonly associated with conventional studio practice. The integration of structured cognitive tools within early design education can therefore contribute to improving the quality and coherence of architectural design outcomes.

Although the findings are context-specific, they suggest that guided mind mapping represents a transferable approach for strengthening spatial thinking across architectural learning environments. Further research may explore its application within diverse educational contexts and through the integration of digital and three-dimensional design tools.

References

[1] ross, N. , (2023) , *Design thinking: Understanding how designers think and work*. London: Bloomsbury.

[2] Eppler, M.J. , (2021) , *Knowledge visualization: How to communicate ideas with pictures*. Wiesbaden: Springer.

[3] González Castañeda, M., *et al.* , (2023) , 'Diagrammatic reasoning and relational thinking in architectural design education', *Design Studies*, 85, p. 101167.

[4] Hettithanthri, U. , (2023) , 'Enhancing relational thinking in architectural design studios through structured visual tools', *International Journal of Architectural Research*, 17(2), pp. 310–326.

[5] Yorgancioglu, D. , (2025) , 'Visual–cognitive scaffolding in architectural education: Enhancing design reasoning through structured mapping', *Journal of Design Research*, 23(1), pp. 78–95.

[6] Türkmenoğlu Berkan, S., *et al.* , (2025) , 'Developing spatial thinking skills in architectural education through structured design exercises', *International Journal of Technology and Design Education*, 35(1), pp. 45–63.

[7] Alwan, A.H. , (2024) , 'Ethical and cognitive dimensions in architectural design education', *Iraqi Journal of Architecture and Planning*, 16(1), pp. 114–129.

[8] Kosslyn, S.M. and Thompson, W.L. , (2020) , *Mental imagery and spatial cognition*. Cambridge: MIT Press.

[9] Ware, C. (2021) *Information visualization: Perception for design*. 4th edn. Burlington: Morgan Kaufmann.

[10] Lawson, B. , (2005) , *How designers think: The design process demystified*. 4th edn. Oxford: Architectural Press.

[11] Giachetta, A. and Buondonno, L. , (2024) , 'Architect training in multifaceted environments: A new cognitive level approach', *AGATHÓN*, 16, pp. 50–59.

[12] Cho, J. and Suh, M. , (2023) , 'The role of spatial ability in architectural design learning: A study of visualization and mental rotation skills', *Journal of Architectural Education*, 77(2), pp. 215–229.

[13] Salama, A.M. , (2016) , *Spatial design education: New directions for pedagogy in architecture and beyond*. Farnham: Ashgate.

[14] Oxman, R. , (2022) , 'Thinking difference: Theories and models of parametric design thinking', *Design Studies*, 80, p. 101101.

[15] Oxman, R. , (2006) , 'Theory and design in the first digital age', *Design Studies*, 27(3), pp. 229–265.

[16] Dorst, K. and Cross, N. , (2001) , 'Creativity in the design process: Co-evolution of problem–solution', *Design Studies*, 22(5), pp. 425–437.

[17] Alwan, A. and Hameed, M. , (2020) , 'Assessment methods in architectural design studios: A framework for evaluating design thinking and spatial skills', *International Journal of Architectural Research*, 14(3), pp. 520–535.

[18] Goldschmidt, G. , (2021) , 'Linking design cognition and creativity in architectural education', *Design Studies*, 75, p. 101030.

[19] Eppler, M.J. , (2022) , 'Knowledge visualization: A conceptual framework and future directions', *Information Visualization*.

The Impact of Architectural Form on Seismic Performance of Buildings / Original Article

Researcher Durgham Munir Abdul Hadi[1*],

Prof. Dr. Abdullah Saadoon Salman[2],

Assist Prof. Dr. Lahib Ali Abdul Hussein[3]

1,2,3 Architecture Engineering Dept., College of Engineering Architecture, University of Technology, - Baghdad/Iraq.

*** Coressponding Author: durgham199788@gmail.com**

تأثير البعد الشكلي على الأداء الزلزالي لعمارة الأبنية العالية

الباحث ضرغام منير عبد الهادي[1]،

أ.د عبد الله سعدون سلمان[2]،

أ.م.د لهيب علي عبد الحسين3

1، 2، 3 قسم هندسة العمارة، كلية الهندسة المعمارية، الجامعة التكنولوجية، بغداد\العراق

Abstract

Seismic overall performance of buildings is a important issue in each architectural and structural design, especially in regions exposed to seismic dangers. Despite giant improvements in structural engineering technologies, the function of architectural form stays insufficiently incorporated into layout practices geared toward improving seismic resistance. This look at investigates the effect of the formal dimension of architectural design on the seismic performance of buildings by analyzing key shape-related attributes, namely regularity, symmetry, simplicity, continuity, and slenderness. An analytical–descriptive method is followed, based totally on an extensive evaluation of applicable literature and the assessment of architectural configurations with regards to their structural behavior underneath seismic loading. The findings indicate that architectural form performs a fundamental role in organizing mass and stiffness distribution, defining load paths, and influencing the dynamic reaction of homes. Regular, symmetrical, and simple forms contribute to a greater balanced seismic response and reduce torsional results, whereas abnormal and complicated configurations cause multiplied structural vulnerability and unpredictable behavior. Continuity enhances the integrity of load transfer mechanisms, whilst slenderness presents great challenges in excessive-upward thrust buildings due to expanded lateral displacements. The examine concludes that integrating formal design considerations inside the early levels of architectural decision-making is essential for enhancing seismic resilience. It also proposes a conceptual framework linking architectural form with seismic performance, thereby bridging the gap between architectural design and structural behavior, and assisting the improvement of extra green and resilien.

KEYWORDS: Seismic Performance, Architectural Form, Formal Dimension, Load Path

المستخلص

يعد الأداء الزلزالي للأبنية العالية من القضايا الأساسية في التصميم المعماري والإنشائي، لا سيما في المناطق المعرضة للنشاط الزلزالي. وعلى الرغم من التطور الكبير في التقنيات الإنشائية، لا يزال دور الشكل المعماري في تحسين مقاومة الأبنية للزلازل غير مستثمر بالشكل الكافي ضمن الممارسات التصميمية. كما **يهدف هذا البحث إلى تحليل تأثير البعد الشكلي على الأداء الزلزالي للأبنية، من خلال دراسة مجموعة من المفردات الشكلية الأساسية، وهي الانتظام، التماثل، البساطة، الاستمرارية، والنحافة**. ويعتمد البحث منهجا تحليليا وصفيا، يستند إلى مراجعة الأدبيات العلمية وتحليل النماذج الشكلية المرتبطة بالسلوك المعماري والإنشائي تحت تأثير الأحمال الزلزالية. أظهرت النتائج أن الشكل المعماري يلعب دورا محوريا في تنظيم توزيع الكتلة والصلابة، وتحديد مسارات انتقال القوى، والتأثير على الاستجابة الديناميكية للمبنى. حيث تسهم الأشكال المنتظمة والمتماثلة والبسيطة في تحقيق استجابة زلزالية متوازنة وتقليل التأثيرات الالتوائية، في حين تؤدي الأشكال غير المنتظمة والمعقدة إلى استجابات إنشائية غير مستقرة وزيادة في احتمالية الضرر. كما تبين أن الاستمرارية تعزز تكامل النظام الإنشائي، بينما تمثل النحافة تحديا في الأبنية العالية بسبب زيادة الإزاحات الجانبية. يخلص البحث إلى أن دمج اعتبارات البعد الشكلي ضمن المراحل المبكرة من التصميم المعماري يعد مدخلا فعالا لتعزيز مقاومة الأبنية للزلازل. كما يقدم إطارا مفاهيميا يربط بين الخصائص الشكلية والأداء الزلزالي، مما يساهم في تقليص الفجوة بين التصميم المعماري والتحليل الإنشائي، ويدعم تطوير استراتيجيات تصميمية أكثر كفاءة ومرونة، خاصة في الأبنية السكنية العالية.

الكلمات المفتاحية: الأداء الزلزالي، الشكل المعماري، البعد الشكلي، مسارات انتقال الأحمال.

1. Introduction

Earthquakes are taken into consideration some of the maximum dangerous herbal hazards on the Earth's floor, without delay affecting the built surroundings, as they cause widespread human and fabric losses, particularly in areas with high urban density. With the increasing trend towards vertical expansion and the emergence of high-upward thrust buildings, it has turn out to be important to increase layout techniques and strategies able to improving seismic performance and reducing related dangers. As is not unusual in maximum studies inside the field of seismic engineering, the focal point is primarily located on structural elements, such as improving fabric residences, developing lateral force-resisting systems, and employing isolation and damping techniques. Despite the significance of these efforts, the role of architectural design—specially architectural form—remains insufficiently utilized as an effective means of enhancing homes' resistance to earthquakes.

Architectural shape represents one of the fundamental determinants of a constructing's behavior underneath seismic masses, because it immediately affects the distribution of mass and stiffness, even ر as also defining load paths and the nature of dynamic response. Numerous research have proven that irregular or complicated paperwork cause unbalanced structural responses, accelerated torsional outcomes, and a better chance of localized damage or partial crumble. In comparison, regular, symmetrical, and easy bureaucracy make a contribution to achieving greater stable and predictable seismic overall performance with the aid of decreasing irregularities in force distribution and enhancing load transfer efficiency. Other parameters, which include continuity and slenderness, additionally play a enormous role in determining constructing behavior, specifically in high-upward push buildings, where lateral displacement results and instability.

2. The Research Theoretical Framework

The research targets to become aware of, examine, and evaluate the effect of the formal dimension on enhancing the seismic overall performance of high-rise buildings. This is done by way of reading the effect of architectural shape elements—together with regularity, symmetry, simplicity, continuity, and slenderness—on the character of seismic reaction to lateral structural masses. In addition, the research seeks to bridge the know-how gap among theoretical standards and sensible application in using these elements within architectural design. This is accomplished through the analysis of nearby case research to decide their compatibility with theoretical principles of seismic overall performance. The studies ultimately develops a layout framework that clarifies the mechanisms for integrating formal design components with structural issues, thereby contributing to decreasing seismic impacts and improving the safety and efficiency of excessive-upward thrust residential homes.

3. The Research Theoretical Framework

The conceptual framework of the research comprises four main axes. The first addresses the concept of architectural form, while the second focuses on the seismic performance of buildings. The third examines the integrative relationship between formal design parameters and seismic response, whereas the fourth axis is concerned with the study of high-rise buildings as the applied case of the research.

3.3 Architectural Form

Architectural Form is described because the geometric and mass configuration of a building, which is shaped thru the association and corporation of its spatial and volumetric factors on one hand, and its structural factors on the opposite, inside a

unified framework that defines its dimensions and ordinary shape, whether or not at the level of the plan, elevation, or volumetric composition. Architectural shape helps mirror a hard and fast of morphological traits along with simplicity or complexity, regularity or irregularity, and symmetry or asymmetry, which result from the layout choices made via the architect throughout the early degrees of the design process[1].

The role of architectural form isn't limited to being merely a visual or aesthetic expression; instead, it is taken into consideration an influential thing within the normal seismic overall performance of a building. It is circuitously associated with the behavior of the structural gadget via its impact on mass distribution, spatial organisation, and cargo transfer paths. From this perspective, **architectural form is regarded as one of the maximum critical criteria and determinants that make a contribution to shaping the initial seismic response of a constructing while subjected to external loads, which includes seismic hundreds**[2].

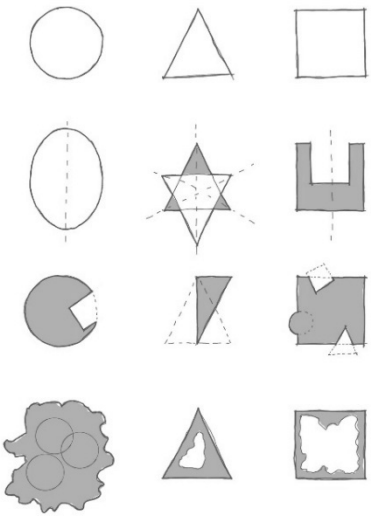


Figure (1) Source: (Author).

In the context of seismic-resistant structure, architectural form acquires specific significance, as its geometric residences impact the degree of constructing regularity and the balance of its additives, that is eventually pondered in the efficiency of its seismic overall performance. Regular and easy bureaucracy regularly permit for a balanced distribution of forces and a reduction in abnormal deformations, while complex or irregular bureaucracy may additionally cause sudden responses, thereby requiring extra architectural and structural layout interventions to make sure the safety of the building [3]. Figure (1) illustrates the types of forms (simple, composite, and complex).

3.4 Seismic Performance

Seismic performance is defined as the ability of buildings to respond to seismic loads in a manner that achieves specified levels of safety and security while minimizing damage and distortion and avoiding total structural collapse. This performance is evaluated based on the behavior of the building under dynamic forces and its ability to absorb and dissipate seismic energy while maintaining its structural and functional stability [4]. Seismic performance is directly related to the characteristics of both the structural system and the architectural configuration. Factors such as mass and stiffness distribution, geometric regularity and type of lateral load-resisting system play an important role in determining the response of a building during an earthquake. Furthermore, geotechnical site conditions play an important role in amplifying or reducing the impact of seismic waves, making seismic performance the result of a complex interaction between building properties and the surroundings [5].

Within the modern framework, the concept of performance-based seismic design (PBSD) has been adopted as an advanced approach that aims to define

different performance levels for buildings, such as immediate use, life safety and collapse prevention. This approach enables the prediction of building behavior under different seismic scenarios and links design decisions to both functional and structural objectives [6]. From an architectural perspective, achieving seismic performance extends beyond structural treatment to include the formal configuration of the building. Architectural form parameters such as regularity, symmetry, simplicity and continuity directly affect the distribution.

3.5 The Relationship between Architectural Form and Seismic Response

The seismic response of buildings is closely related to their formal characteristics, as architectural form is considered one of the most important factors and determinants that control the nature of the dynamic behavior of a building when subjected to seismic loads. This is evident in the influence of architectural form on mass distribution, stiffness distribution and the center-of-mass-center-of-stiffness ratio – factors that are crucial in controlling displacements and deformations induced by seismic events. The more regular the architectural form, whether in plan or vertical, the more balanced the distribution of seismic forces, which reduces stress concentrations and limits irregular deformations. In contrast, irregular shapes give rise to more complex responses due to uneven mass and stiffness distribution, resulting in phenomena such as torsion and localized buckling, which are the primary causes of structural failure [2].

Architectural form also plays an important role in defining load paths within building as illustrated in Simple and continuous shapes provide clear and regular paths for transferring loads to the foundation, while complex or discontinuous shapes disrupt these paths, leading to force concentrations at specific points and

increasing damage potential [4]. Furthermore, symmetry in the formal configuration helps reduce deviations between centers of mass and stiffness, thereby reducing torsional moments, while asymmetry increases seismic responses, especially in tall buildings. Therefore, control of architectural form characteristics represents one of the most important design strategies to improve seismic performance and reduce earthquake-related risks [6].

Figure (2) illustrates the concept of load paths within a building. Simple and non-stop bureaucracy show off clear and regular paths for moving masses from higher stages to the principles, thereby improving seismic overall performance performance. In evaluation, complex architectural bureaucracy or the ones containing discontinuities or abrupt adjustments in configuration obstruct those paths, ensuing in pressure concentrations at particular factors and growing the chance of structural harm.

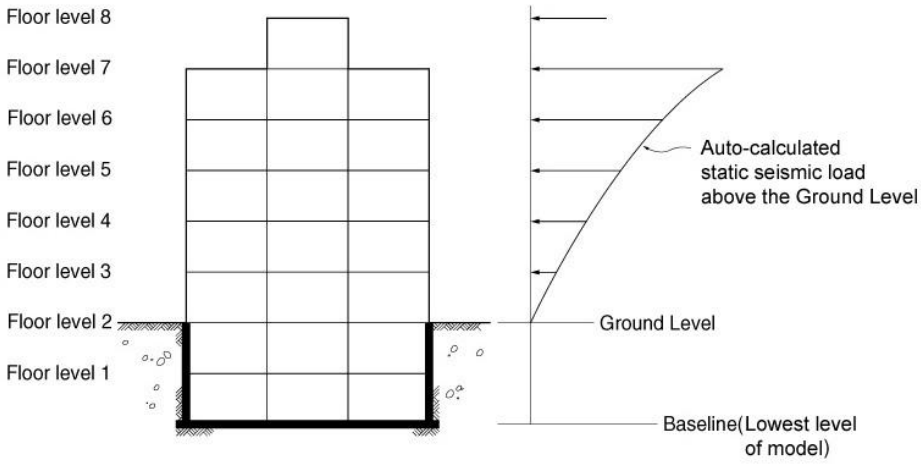


Figure (2) Illustrates the load transfer path within the building.
Source: Prepared by the author.

Figure (3) offers a assessment between a building with a normal architectural form on the proper facet of the picture and an abnormal one on the left. It can be determined that the regular building famous a uniform distribution of seismic forces, resulting in a strong and predictable reaction. In assessment, the abnormal constructing demonstrates complex conduct because of abrupt modifications in shape, mass, or stiffness, leading to stress concentrations and the emergence of vulnerable zones that growth the chance of structural failure.

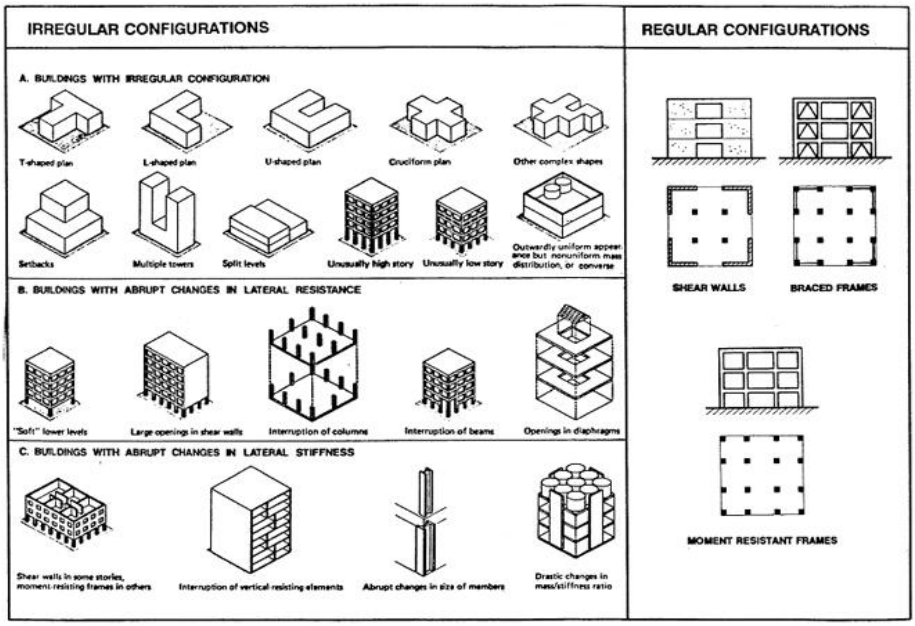


Figure (3) Illustrates the difference between regular and irregular distribution[3].

In end, architectural shape plays a important position in figuring out the kind and nature of a constructing's seismic response via without delay influencing mass and stiffness distribution and organizing load switch paths in the structure. Form characteristics, which include regularity, symmetry, simplicity, and others, make a contribution to attaining more stable and predictable structural behavior. In

contrast, irregularity and complexity extend seismic responses and growth the probability of failure, probably leading to constructing disintegrate.

4. The Theoretical Framework of the Research

The theoretical framework for this research is based on the review and analysis of a set of previous studies that investigated the relationship between architectural form and seismic performance with the aim of extracting key formal aspects that influence building behavior under seismic loads. This approach is based on the premise that formal design decisions in architecture represent the primary determinant of structural response rather than merely acting as an aesthetic element separate from performance.

In this context, a comparative study on regular and irregular architectural configurations in the design of earthquake-resistant buildings (2020) showed that regular buildings are characterized by simplicity, symmetry and balanced load distribution, which contributes to better seismic response and more predictable behavior. In contrast, irregular buildings face problems related to uneven force distribution and the emergence of torsional effects, resulting in complex and unstable structural behavior that requires additional structural solutions to overcome these imbalances. Furthermore, the study identified a set of key indicators that influence seismic performance, including simplicity versus complexity, symmetry versus asymmetry, regularity versus irregularity, as well as dimensional relationships between architectural mass in both horizontal and vertical directions. The study concluded that simple and symmetrical shapes promote uniform load distribution and reduce stress concentrations, thus having a positive impact on the building's structural performance. Similarly, Arnold and Reitherman (2002) presented a comprehensive framework for the influence of formal configuration in seismic design, emphasizing that the shape of a building is the primary factor in determining its seismicity.

Through the evaluation of preceding studies Table (1), it will become glaring that the impact of architectural shape on seismic overall performance can be categorized into a set of measurable and analyzable design variables, which function key indicators for comparing the efficiency of the formal configuration. Accordingly, those variables may be organized into an analytical desk that gives every variable, its architectural and structural indicators, and its role in mitigating seismic results, in education for his or her use as an analytical device inside the studies.

Table (1) Illustrates the formal vocabulary, architectural indicators, and their seismic impact derived from previous studies. Source: Researcher

The role in reducing seismic impact	Architectural – structural index	The formal dimension	the study
It reduces response irregularity and limits force concentration.	Uniformity of mass and rigidity horizontally and vertically	Regularity	Comparative study between regular and irregular morphological formation(2020)
It reduces torsion and improves building stability.	Symmetry of the plan around the principal axes	Symmetry	
It facilitates force transfer and reduces structural complexity.	Simple geometric shapes (square, rectangle)	Simplicity	
It increases the probability of localized failure.	Sudden change in shape or mass	Irregularity	
Reduces secondary forces and improves overall performance	The building's overall shape is uniform.	Formal regularity	Arnold & Reitherman study (2002)
Reduces torsional moments	The mass is symmetrical around the center.	Symmetry	
Leads to weak structural bonding	The mass is divided into unrelated parts.	Segmentation	
Increases the efficiency of analysis and design	The geometric system is clear.	Formal simplicity	

The role in reducing seismic impact	Architectural – structural index	The formal dimension	the study
Ensures uniform force transfer	Continuity of construction paths	Continuity	Bachmann's study (2012)
Prevents load concentration and poor performance	Interruptions in form or height	Sudden change	
Reduces dynamic shocks	Gradual transition in mass or layout	Gradualism	
Facilitates architectural and structural integration	Clarity of design concept	Conceptual simplicity	
Increases sensitivity to lateral loads	High height versus low width	Slenderness	A study of high-rise buildings and their environmental impact (2018)
Affects vibration duration	Number of floors and total height	Height (building height)	
Improves overall stability	Reducing wind resistance	Flowable shape	
Affects center of mass and response	Distribution of building masses	Mass composition	

Based on the desk above and the evaluate of preceding research, it's miles obtrusive that architectural shape isn't merely a cultured detail, but as a substitute a fundamental determinant of a building's performance in responding to seismic masses. Studies have proven that there may be a hard and fast of formal variables that continually appear as the most influential in improving seismic performance, major amongst them being symmetry and ease. These variables contribute to attaining a balanced distribution of mass and stiffness, reducing torsional outcomes, and making sure the everyday transfer of forces inside the structural device. Additional variables, along with continuity and slenderness, each play a particular function in influencing pressure distribution, organizing load paths, and minimizing deformations due to seismic masses. Accordingly, the subsequent sections will cope

with those variables in element by way of analyzing their traits, architectural and structural signs, and their function in enhancing the seismic performance of homes, as follows:

4.1 Symmetry

Symmetry is one of the fundamental formal traits that contributes to improving the seismic performance of buildings. It refers to reaching a balanced distribution of mass and stiffness around unique axes within the constructing, whether or not these axes are horizontal or vertical. Symmetry can be found out while a constructing famous a similar configuration on both facets of one or greater axes, ensuing in an equitable distribution of architectural and structural elements.

Symmetry is without delay related to the constructing's conduct all through earthquakes Figure (4), because it helps reduce discrepancies among the center of mass and the middle of stiffness, thereby limiting torsional consequences Figure (5), which can be some of the maximum critical phenomena affecting building reaction. The higher the level of symmetry, the greater balanced the switch of seismic forces, and the decrease the likelihood of irregular deformations or pressure concentrations. Conversely, the absence of symmetry (asymmetry) results in an uneven distribution of mass and stiffness, producing variations within the response of different constructing elements. This results in torsional rotations around the vertical axis, which make bigger displacements in a few regions as compared to others. Consequently, this situation poses more risk in excessive-upward thrust or complexly configured buildings, wherein dynamic responses grow to be greater touchy to geometric imbalances[8].

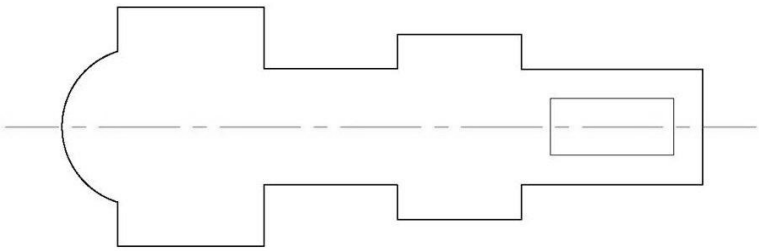


Figure (4) Illustrates a schematic of a hypothetical constructing this is symmetric about a unmarried axis, displaying a balanced distribution of mass and structural elements. This symmetry outcomes in a uniform seismic response and reduces displacement differences among various elements of the building. Source: Prepared through the author.

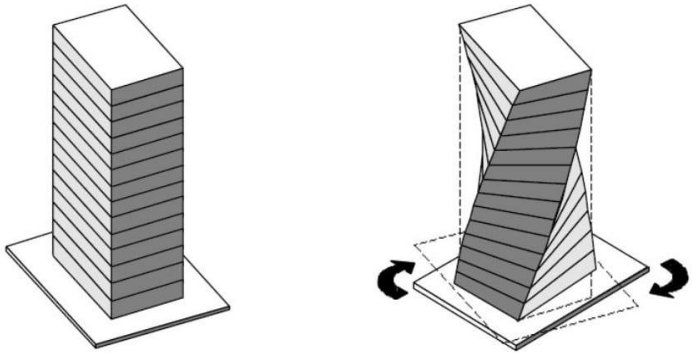


Figure (5) Illustrates the phenomenon of torsional rotation attributable to asymmetry, wherein the constructing rotates around its vertical axis because of unbalanced forces, increasing the chance of harm on the extremities. Source[8].

From the analysis of the above figures, it is able to be concluded that symmetry performs a fundamental function in attaining a balanced distribution of mass and stiffness within a constructing, thereby decreasing discrepancies between the center of mass and the center of stiffness and mitigating torsional results during earthquakes. Symmetry helps unify the response of various constructing elements and minimizes abnormal deformations, which complements structural stability and

makes the dynamic behavior more predictable. Accordingly, symmetry is taken into consideration one of the key design ideas for lowering seismic risks and enhancing the overall efficiency of constructing overall performance.

4.2 Simplicity

Simplicity is one of the essential design standards that at once influences the seismic overall performance of high-upward push homes. It refers to adopting clean and straightforward formal configurations characterised by way of geometric regularity and the avoidance of unnecessary complexity or overlapping paperwork. Simplicity is frequently embodied in ordinary shapes together with rectangles and squares, which allow for a balanced distribution of mass and stiffness[2].

The significance of simplicity lies in its contribution to attaining clearer and greater predictable structural conduct in the course of earthquakes, as it facilitates the transfer of seismic hundreds thru ordinary and direct load paths, decreasing the likelihood of stress concentrations or localized deformations. Simple bureaucracy Figure (6), additionally lessen analytical complexities in dynamic modeling and permit extra efficient structural detail layout. In contrast, complicated architectural forms—which includes damaged or irregular plans (L-, U-, or T-formed)— Figure (7), bring about heterogeneous seismic behavior, causing interruptions in load paths and producing susceptible zones due to uneven distribution of mass and stiffness, which increases the probability of cracking or partial crumble. Furthermore, formal complexity is regularly related to torsional effects and abnormal deformations, specially when the form isn't like minded with the structural device [11].

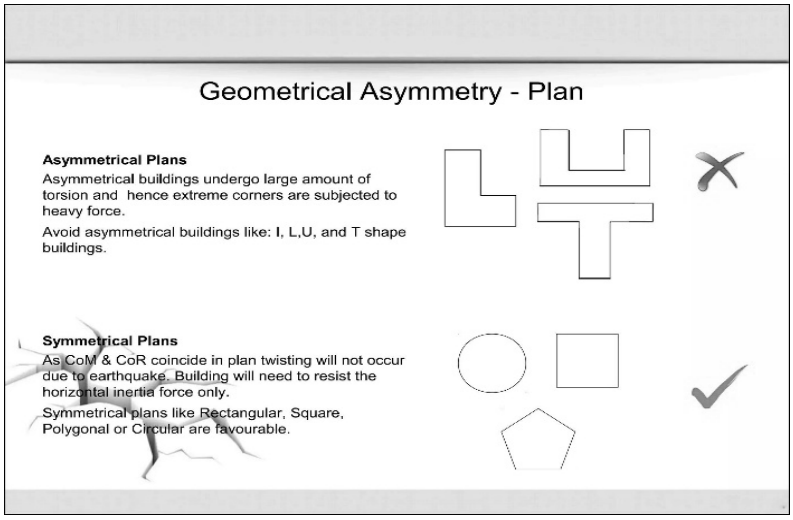


Figure (6) Gives a schematic illustrating the distinction between simple bureaucracy (square or square) and complex or composite paperwork (L-, T-, U-shaped), which result in heterogeneous seismic conduct. Source[9].

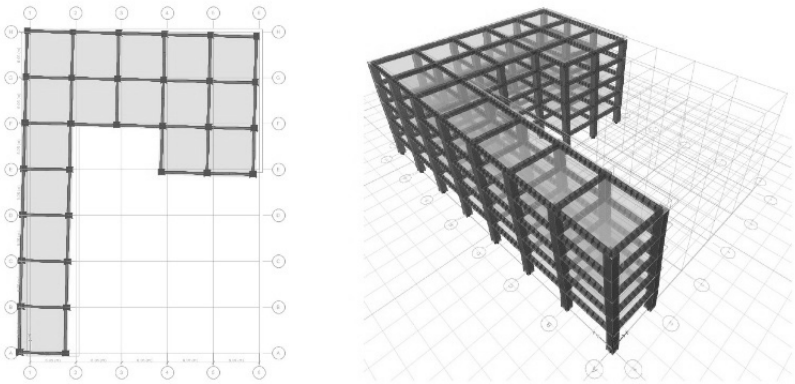


Figure (7) Illustrates a building with a complex and abnormal form, where corners and offsets disrupt the burden paths, leading to stress concentrations and an expanded probability of harm. Source[2].

Analysis of the above figures suggests that simplicity in architectural configuration without delay contributes to improving seismic performance by imparting clarity and regularity in load paths and decreasing formal complexities that may reason interruptions or stress concentrations. Simple and compact forms assist distribute hundreds evenly and reduce the likelihood of localized deformations or partial collapses. Therefore, adopting simplicity as a design technique is an vital factor in improving building stability and increasing its performance in resisting seismic loads.

4.3 Continuity

Continuity is one of the fundamental elements in reaching seismic performance in excessive-upward thrust homes, as it is immediately related to the structural gadget's capacity to switch seismic loads in a continuous and regular way through its various additives Figure (8). Continuity refers to making sure the mixing and connectivity of vertical and horizontal factors inside a unified structural device, forming clear load paths that begin from the points of load application at the pinnacle of the constructing and extend down to the foundations without interruption or unexpected changes in path[12].

The importance of continuity lies in its potential to reduce complexities inside the building's dynamic conduct throughout earthquakes. Systems with excessive continuity show off a extra uniform reaction, with evenly allotted forces and displacements, minimizing the chance of vulnerable zones or stress concentrations. Continuity additionally enhances the included overall performance of various structural factors, which includes columns, shear walls, and frames, permitting them to act as a unmarried unit to face up to lateral hundreds. In contrast, discontinuities disrupt load paths, representing one of the maximum crucial reasons

of seismic failure. Such discontinuities may additionally rise up from gentle testimonies, abrupt changes in element dimensions, architectural setbacks, or loss of continuity in shear partitions alongside the constructing height Figure (9). These interruptions make bigger interstory drifts and pay attention deformations in unique regions, growing the chance of partial or total fall apart[8].

Therefore, reaching continuity isn't always restrained to the structural element alone however calls for near integration with architectural choices within the early design stages to make sure compatibility among the architectural form and the structural machine, thereby ach

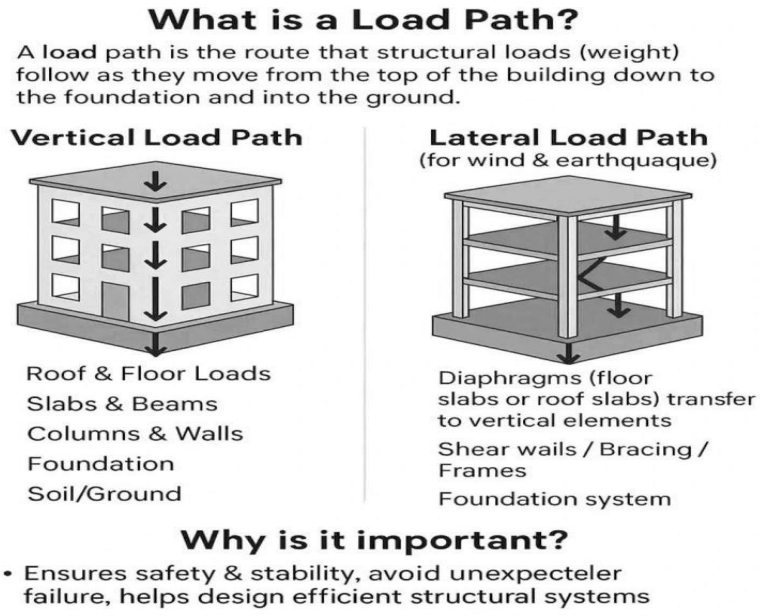


Figure (8) Illustrates the distinction between seismic load transfer thru a non-stop path from the pinnacle of the building to the principles, where forces are progressively disbursed across structural elements—representing the ideal case of continuity—and the case of misplaced continuity, wherein one story is much less stiff than the others, causing localized deformation and growing the chance of disintegrate. This is among the most commonplace reasons of seismic failure[8].

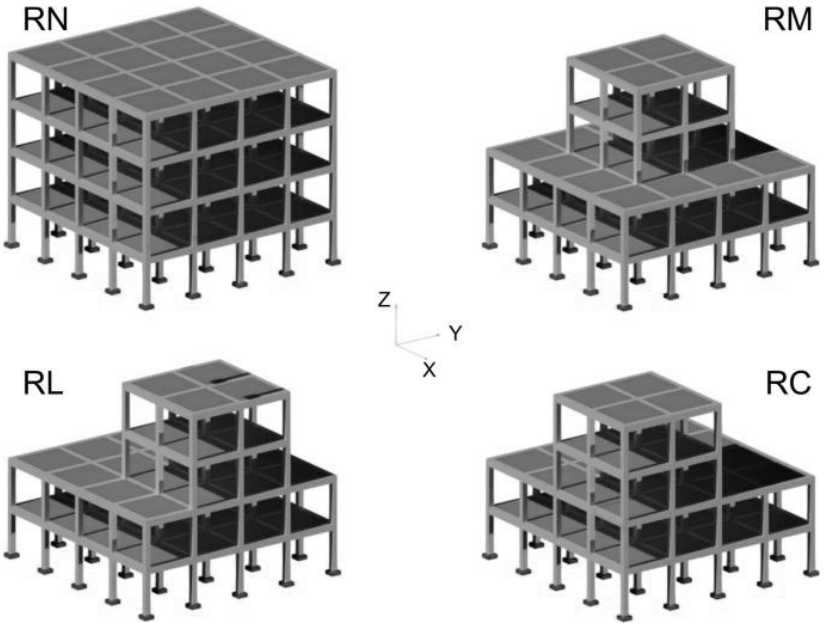


Figure (9) Illustrates the effect of geometric discontinuities in a constructing, wherein surprising adjustments in form regulate the burden paths and concentrate stresses in particular regions, thereby compromising seismic resistance performance[10].

In end, continuity represents a critical issue in reaching balanced seismic performance, because it ensures ordinary load transfer and minimizes localized deformations. Conversely, its absence disrupts structural response and increases the chance of failure, underscoring the necessity of integrating continuity with architectural layout choices from the early stages.

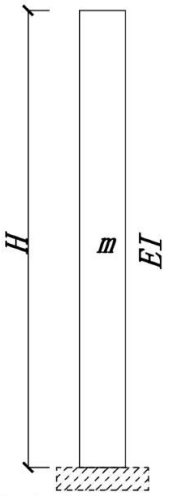
4.4 Slenderness

Slenderness is one of the fundamental geometric characteristics that without delay have an effect on the dynamic conduct of homes, specially excessive-upward push residential structures. It expresses the ratio among the constructing's height and its horizontal dimensions; as this ratio increases, the

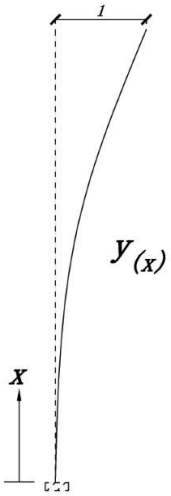
effect of lateral forces, which include seismic loads, at the building's reaction additionally will increase.

The importance of slenderness lies in its control over the constructing's flexibility and vibration susceptibility. Slender homes have a tendency to show off longer natural periods, which amplifies lateral displacements and interstory drifts. While this pliability might also help dissipate some seismic strength, it can concurrently result in immoderate deformations, affecting the integrity of non-structural factors and person comfort. Moreover, expanded slenderness amplifies P-Δ effects, as inner forces due to lateral displacements grow, doubtlessly causing slow instability in the structural machine. Slender buildings also are extra touchy to irregularities or asymmetry, complicating their seismic response. Conversely, decreasing slenderness with the aid of growing the building's width or using efficient structural structures (together with shear walls or tube structures) complements lateral stiffness, minimizes displacements, and achieves a higher balance among stiffness and flexibility. Therefore, controlling the slenderness ratio is a critical layout selection that requires specific coordination between architectural and structural factors to make sure secure and powerful seismic performance[2]. Figure (10) illustrates the overall behavior of a slender building during earthquakes and how lateral drift is distributed among the floors. Lateral displacements increase noticeably due to the high height-to-width ratio of the building, reflecting greater flexibility and sensitivity to lateral loads, as well as a longer natural vibration period, which may lead to deformations in non-structural elements[8].

a) Tall building structure



b) Mechanical diagram



c)Main vibration mode curve

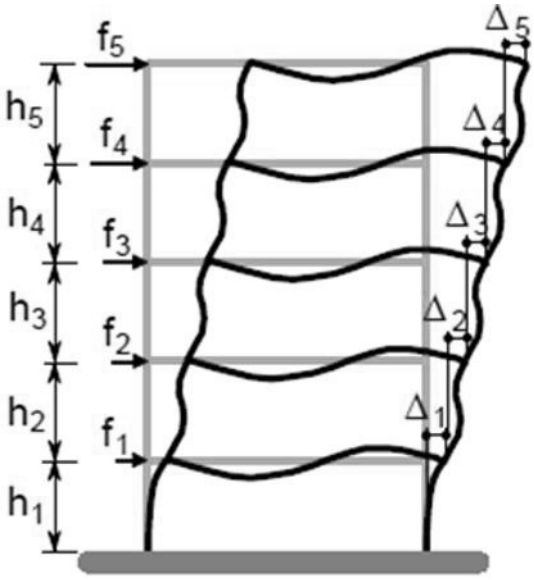


Figure (10) Shows the behavior of a slender building during earthquakes. Source: Researcher.

It is clear that slenderness is a noticeably influential thing in figuring out the seismic reaction of buildings. An increase in slenderness amplifies lateral displacements and deformations, as well as magnifies instability consequences, necessitating cautious

architectural and structural design interventions. Therefore, achieving a properly-taken into consideration balance within the slenderness ratio is essential to make certain secure seismic overall performance, especially in high-upward thrust homes.

5. The Applied Framework of the Research

This section goals to assess the performance of the architectural form of decided on excessive-upward push residential buildings in Baghdad, particularly (the Garden City Residential Towers, the Al-Fakher Residential Towers, and the Downtown Towers), by way of making use of the formal layout factors and linking them to the predicted seismic overall performance signs. An analytical technique become followed, specializing in deciphering the connection among architectural characteristics and structural conduct below lateral masses.

The applied framework is primarily based at the hypothesis that systematically using formal design factors complements the building's seismic response and reduces the damaging outcomes of lateral forces. Accordingly, 5 key layout elements representing the middle of the formal components were adopted: Regularity, Symmetry, Simplicity, Continuity, and Slenderness. Each element could be analyzed inside its realistic context inside the decided on models, as follows:

First: Garden City Towers

The Garden City Towers are located in Baghdad's Al-Alawi place inside a cutting-edge residential complicated Figure (11). They characteristic a present day layout primarily based on notably repetitive vertical blocks, with contemporary architectural treatment of the facades.

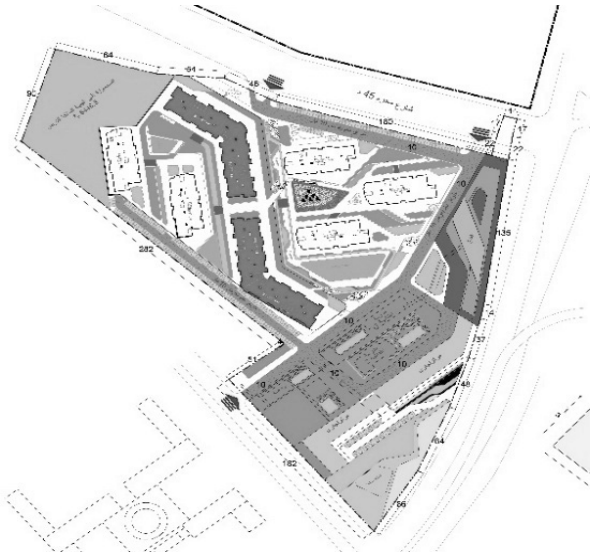


Figure (11) Illustrates the site plan of the Garden City Towers [13].

Analysis of the architectural shape of the Garden City Towers indicates that the structural system relies on easy geometric configurations with a rather regular distribution of mass and stiffness, contributing to a reduction in dynamic reaction irregularities. The horizontal and vertical regularity on this model enables limit differential go with the flow and enhances the efficiency of seismic load transfer thru the structural factors.

However, symmetry is not completely performed, as a few facades showcase partial versions in architectural treatment, which may additionally generate torsional moments due to misalignment among the middle of mass and the middle of stiffness. Continuity in the structural load direction is moderate, given the presence of certain interruptions within the façade format, which can have an effect on the performance of load switch.

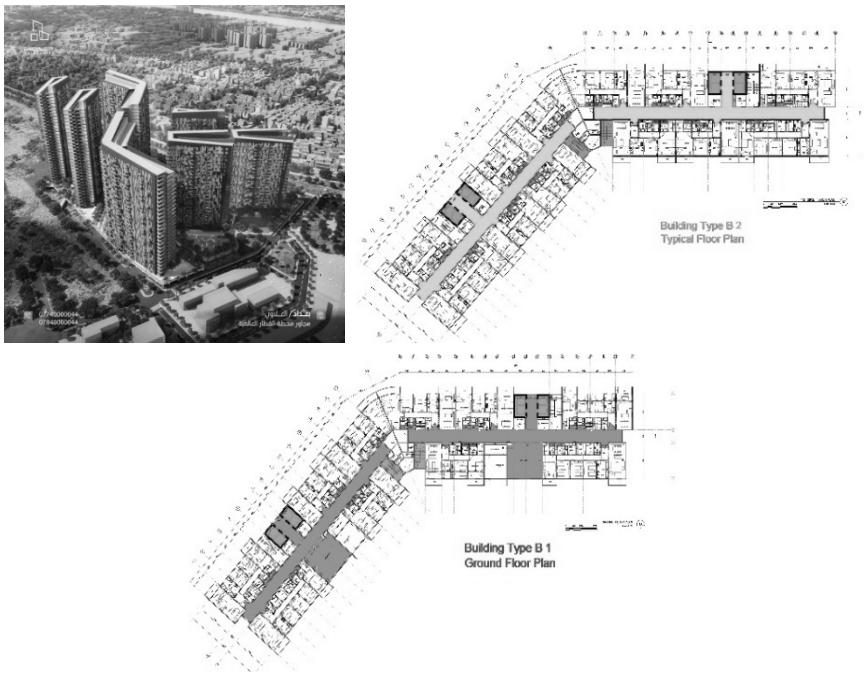


Figure (12) Illustrates the ground floor and typical floor plans of one of the Garden City Towers, along with the exterior view of the residential complex [13].

Based on the above images and plans in Figure (12), the formal layout factors of this pattern will be evaluated to evaluate the effectiveness of its seismic performance, as follows:

Table (2) Shows the evaluation of the formal dimension items for the First sample. Source: Researcher

Se.	Element	Evaluation	Interpretation
1	Regularity	middle	Some variations in mass are present.
2	Simplicity	weak	The form and interfaces are complex.
3	Continuity	middle	Partial continuity in elements is lacking.
4	Slenderness	High	There is a significant height compared to the width.
5	Symmetry	weak	A clear asymmetry in the composition is evident.

From the Table (2), the seismic performance of this sample may be categorised as exceptionally exact, as the simplicity and partial symmetry make amends for the weaknesses resulting from the shortage of complete symmetry.

Secondly: Al-Fakher Residential Towers

Al-Fakher Residential Towers are placed within the Al-Qadisiyah location of Baghdad and represent one of the city's distinguished excessive-upward thrust residential tasks Figure (13), comprising several towers exceeding fifty flooring. The fairly open web site with river perspectives topics the towers to high lateral hundreds, whether or not from wind or seismic forces, making the architectural configuration a important issue in figuring out structural performance performance.

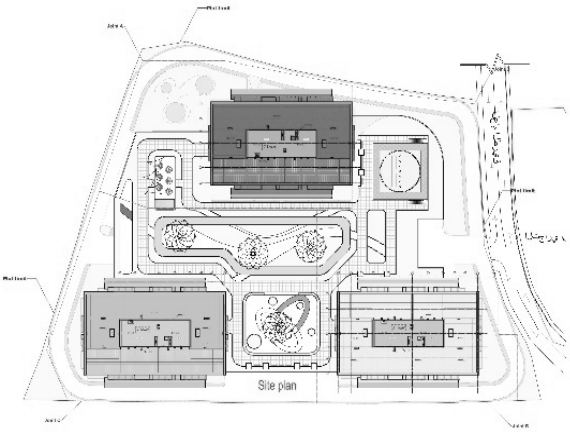


Figure (13) Illustrates the site plan of Al-Fakher Towers [14].

The morphological evaluation suggests that the challenge is based on a multi-block system, wherein the towers encompass slim vertical blocks repeated ground by means of floor, with outstanding variant in the façade treatment via projections, recesses, and special arrangements of balconies and openings.

Symmetry isn't fully done within the towers, inflicting a deviation between the center of mass and the center of stiffness, which generates torsional moments adversely affecting seismic response. Furthermore, the configuration famous a high degree of complexity, lowering the clarity of load paths and growing the likelihood of stress awareness. Regarding continuity, it is only partly realized due to interruptions inside the façade layout. As for slenderness, the high-rise towers' top-to-width ratio increases the constructing's sensitivity to lateral loads and amplifies lateral displacements, mainly in the absence of a regular geometric configuration.

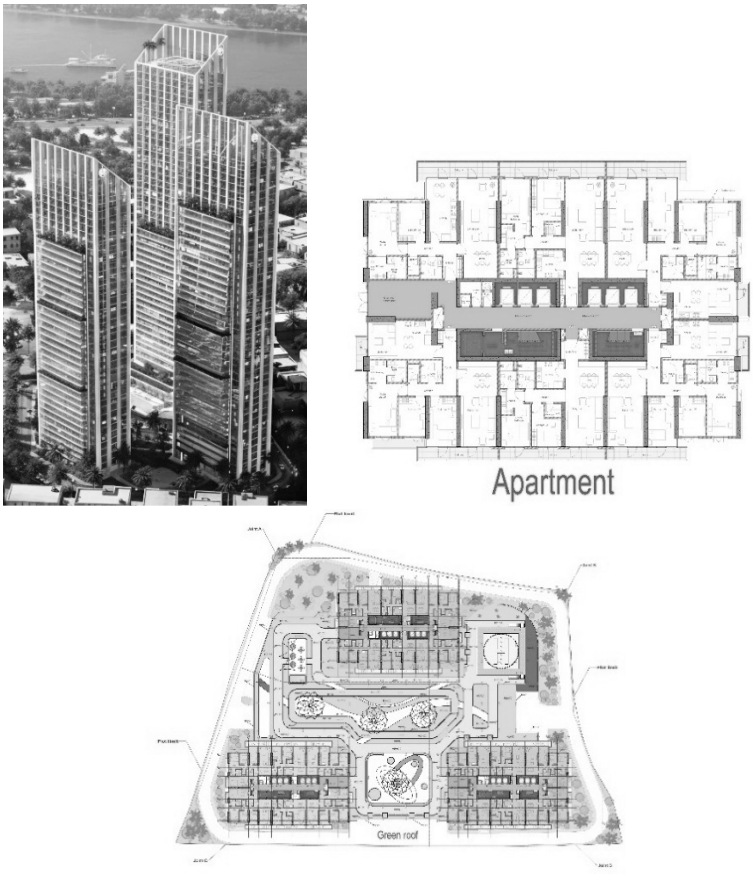


Figure (14) illustrates the ground floor and typical floor plans of one of the Al-Fakher Residential Towers, along with the exterior view of the residential complex [14].

Based on the above photographs and plans in Figure (14), the formal layout elements of this sample may be evaluated to decide the effectiveness of its seismic overall performance, as follows:

Table (3) shows the evaluation of the formal dimension items for the Secondly sample. Source: Researcher

Se.	Element	Evaluation	Interpretation
1	Regularity	High	Uniform distribution of mass and height
2	Simplicity	middle	Some variation without excessive complexity
3	Continuity	High	Clear continuity in elements
4	Slenderness	middle	Balanced proportions
5	Symmetry	High	Clear symmetry around the axes

From the Table3)), concluded that the predicted seismic performance of those towers levels from slight to enormously susceptible from a morphological perspective. This suggests that counting on complicated architectural remedies with out systematically controlling formal design elements ends in increased structural needs and decreased seismic performance. This case also highlights the importance of early integration among architectural and structural layout to achieve a stability between formal expression and structural overall performance in excessive-rise residential buildings.

Third: Downtown Towers

Downtown Towers are placed in Baghdad within a enormously modern urban context (Figure 15). The task includes a collection of residential towers with similar heights, characterised by means of highly everyday geometric configurations, predominantly square in shape, with clean repetition of floor layouts meditated in the regularity of the façades.

Figure (15) illustrates the site plan of the Downtown Towers [15].

Morphological analysis suggests that the towers attain a excessive diploma of regularity, both horizontally and vertically, as mass and stiffness are distributed continually the building top, with out a abrupt modifications in dimensions or spatial configuration. This regularity without delay contributes to a greater balanced seismic reaction. In phrases of symmetry, the towers showcase relative symmetry alongside their essential axes, which helps reduce discrepancies among the center of mass and the center of stiffness. Consequently, torsional moments induced by lateral loads are minimized, that is a key fantastic aspect in enhancing seismic overall performance.

Figure (16) Shows the exterior view of the Downtown Towers [15].

Regarding simplicity, the overall configuration demonstrates a great level of readability, with façade treatments constrained to minor and non-abrupt variations, inclusive of moderate projections or diffused differences in beginning preparations, with out reaching a high stage of formal complexity. This simplicity contributes to clearer load paths within the structural system and decreases irregular deformations for the duration of dynamic response. As for continuity, it is honestly performed, with vertical factors extending always from the base to the pinnacle without primary interruptions, and the repetition of floors retaining the integrity of the structural gadget. Based on the exterior images of the complex Figure (16), the formal design elements of this sample will be evaluated to determine the effectiveness of its seismic performance, as follows:

Table (4) shows the evaluation of the formal dimension items for the third sample. Source: Researcher

Se.	Element	Evaluation	Interpretation
1	Regularity	middle	Some variations in mass are present.
2	Simplicity	middle	Some variation without significant complexity is observed.
3	Continuity	High	Clear continuity in elements is evident.
4	Slenderness	High	A considerable height compared to width is present.
5	Symmetry	weak	Clear asymmetry in composition is apparent.

From the Table (4), it could be concluded that the architectural configuration of the Downtown Towers has a tendency closer to regularity and readability in place of complicated architectural expression, which undoubtedly reflects on seismic overall performance. The combination of regularity, relative symmetry, simplicity, and continuity reduces terrible outcomes including torsion, irregular displacements, and pressure concentrations. Accordingly, the predicted seismic performance of these towers may be categorised as top to excellent from a morphological perspective.

6. Conclusions - Results

6.1 Theoretical Findings and Conclusions

- The formal measurement represents one of the most essential pillars in improving the seismic overall performance of excessive-upward thrust buildings, because the traits of architectural shape immediately influence the nature of structural reaction.
- Studies have proven that regularity in architectural configuration, horizontally or vertically, contributes to achieving a balanced distribution of seismic forces and decreases pressure concentrations, main to a more solid response. In contrast, irregularity amplifies dynamic reaction and will increase the chance of localized damage.
- The literature indicates that symmetry performs a key function in reducing torsional moments by minimizing the space between the center of mass and the center of stiffness, whereas asymmetry leads to unbalanced responses that growth the risk of structural failure.
- Studies highlight the importance of simplicity in improving the clarity of load paths inside the structural system, as simple forms are related to greater predictable structural behavior, unlike complex configurations which growth uncertainty in overall performance.
- The literature confirms that continuity in vertical and horizontal elements is critical for ensuring efficient load switch, at the same time as surprising discontinuities create vital zones. Slenderness, alternatively, has been identified as a main thing in increasing lateral displacements and amplifying dynamic response, particularly in excessive-upward push buildings, therefore requiring careful design treatment.

6.2 Applied Findings and Conclusions

- The analysis showed that towers with clean formal regularity (consisting of Downtown Towers) attain higher performance due to the balanced distribution of mass and stiffness, whereas towers with abnormal configurations (inclusive of a few instances of Al-Fakher Towers) exhibit extra complex responses.
- Asymmetry in a few fashions brought about the generation of torsional moments, negatively affecting performance performance, whilst relative symmetry in other models contributed to a more strong response.
- The applied outcomes verified that towers characterized via formal simplicity and clear architectural configuration provide extra everyday load paths, undoubtedly influencing structural conduct. In evaluation, formal complexity elevated the chance of strain concentration. Continuity in vertical elements and universal configuration became discovered to beautify load switch efficiency, while interruptions and sudden الشكل modifications weakened performance.
- Slenderness truly emerged as an influential component, as more narrow towers have been related to accelerated lateral displacements, necessitating supportive design answers.

References

[1] Ching, F. D. K. (2014). Architecture: Form, Space, and Order (4th ed.). Hoboken, NJ: John Wiley & Sons.

[2] Taranath, B. S. (2016). Structural Analysis and Design of Tall Buildings: Steel and Composite Construction (2nd ed.). Boca Raton, FL: CRC Press. https://www.routledge.com/Structural-Analysis-and-Design-of-Tall-Buildings-Steel-and-Composite/Taranath/p/book/9781439804803

[3] Charleson, A. W. (2008). Seismic Design for Architects: Outwitting the Quake. Oxford, UK: Architectural Press (Elsevier). https://www.elsevier.com/books/seismic-design-for-architects/charleson/978-0-7506-8558-7

[4] Krawinkler, H., & Miranda, E. (2018). Performance-Based Earthquake Engineering. In: Bozorgnia, Y., & Bertero, V. V. (Eds.), Earthquake Engineering: From Engineering Seismology to Performance-Based Engineering. Boca Raton, FL: CRC Press.

[5] Kramer, S. L. (2019). Geotechnical Earthquake Engineering. Upper Saddle River, NJ: Pearson Education.

[6] Mahdavinejad, M., Nikghadam, H., & Pour Rahimian, F. (2013). Architectural Design Strategies for Earthquake-Resistant Buildings. International Journal of Architecture and Urban Development, 3(3), 29–36.

[7] ASCE (2016). Minimum Design Loads and Associated Criteria for Buildings and Other Structures (ASCE/SEI 7-16). Reston, VA: American Society of Civil Engineers. https://ascelibrary.org/doi/book/10.1061/9780784414248

[8] Chopra, A. K. (2017). Dynamics of Structures: Theory and Applications to Earthquake Engineering, 5th Edition.

[9] FEMA 450 (2013). NEHRP Recommended Provisions for Seismic Regulations for New Buildings and Other Structures.

[10] Paulay, T., & Priestley, M. J. N. (2015). Seismic Design of Reinforced Concrete and Masonry Buildings (2nd ed.).

[11] Alecci, V., & De Stefano, M. (2018). Building irregularity issues and architectural design in seismic areas. *Fracture and Structural Integrity*, 13(47), 161–168. https://www.fracturae.com/index.php/fis/article/view/2176

[12] Tirado-Gutiérrez, R. J., *et al.* (2025). The importance of structural configuration in seismic performance. *Infrastructures*, 10(12), 325.

[13] Garden City Residential Complex Management. info.gcb@halabjagroup.com

[14] Al-Fakher Residential Complex Management. https://alfakhertowers.com/ar.html

[15] Downtown Residential Complex Management. https://homele.com/ar/projects/downtown-baghdad-382

Video Compression Optimization Techniques Using Artificial Intelligence: Review / Review Article

Assist. Prof. Amal Abbas Kadhim[1*],

Assist. Prof. Wedad Abdul Khuder Naser[2],

Lecturer Nada Abdulkareem hameed[3]

1,2 Mustansiriyah University, College of Education, Department of Computer Science, Baghdad/Iraq.

3 Al-Mansour University College, Department of Computer Science and Information Systems, Baghdad/Iraq

*** Corresponding Author: abbas.amal@uomustansiriyah.edu.iq**

تقنيات تحسين الضغط الفيديو باستخدام الذكاء الاصطناعي: مراجعة

استاذ مساعد امل عباس كاظم[1]،

استاذ مساعد وداد عبد الخضر ناصر[2]،

مدرس ندى عبد الكريم حميد[3]

1، 2 الجامعة المستنصرية- كلية التربية، قسم علوم الحاسبات، بغداد\العراق

3 كلية المنصور الجامعة، قسم علوم الحاسبات ونظم المعلومات، بغداد\العراق

Abstract

Compressing video is an important and essential function in today's multimedia systems as it enables the efficient storage and transmission of large volumes of video data. The proliferation of high-resolution videos that are being used in various application areas such as video streaming, video conferencing, surveillance, and autonomous systems has caused the strong demand for more efficient compression algorithms. This paper presents an in-depth review of video compression techniques with particular focus on AI methods. It also discusses traditional video coding standards including H. 264/AVC, H. 265/HEVC, and AV1, including motion estimation, transform coding quantization entropy coding, and rate-distortion optimization. The paper provides an extensive review of the latest transformative AI-based techniques which include machine learning, deep learning, and neural networks that are used to effectively enhance compression and visual quality. Lastly, the paper discusses the key challenges like computational complexity, latency, and real-time implementation, and at the same time, it suggests future research paths such as the study of lightweight AI models, hybrid compression systems, and optimization of perceptual quality.

Keywords: Video Compression, Artificial Intelligence, Machine Learning, Deep Learning, Optimization Techniques, H.264/AVC, HEVC, AV1, Rate–Distortion Optimization, Multimedia Systems, Neural Compression.

المستخلص

يعد الضغط الفيديو وظيفة مهمة وأساسية في أنظمة الوسائط المتعددة اليوم لأنه يتيح تخزين ونقل كميات كبيرة من بيانات الفيديو بكفاءة. أدى انتشار مقاطع الفيديو عالية الدقة التي يتم استخدامها في مجالات التطبيقات المختلفة مثل بث الفيديو ومؤتمرات الفيديو والمراقبة والأنظمة المستقلة إلى زيادة الطلب على خوارزميات الضغط الأكثر كفاءة. تقدم هذه الورقة مراجعة متعمقة لتقنيات ضغط الفيديو مع التركيز بشكل خاص على أساليب الذكاء الاصطناعي. ويناقش أيضًا معايير ترميز الفيديو التقليدية.

H. 264/AVC, H. 265/HEVC, and AV1,

بما في ذلك تقدير الحركة، وتحويل الترميز، وترميز الإنتروبيا، وتحسين معدل التشويه. تقدم الدراسة مراجعة شاملة لأحدث التقنيات التحويلية القائمة على الذكاء الاصطناعي والتي تشمل التعلم الآلي والتعلم العميق والشبكات العصبية التي تُستخدم لتحسين الضغط والجودة البصرية بشكل فعال . أخيرًا، تناقش الدراسة التحديات الرئيسية مثل التعقيد الحسابي، وزمن الوصول، والتنفيذ في الوقت الفعلي، وفي الوقت نفسه، تقترح مسارات بحثية مستقبلية مثل دراسة نماذج الذكاء الاصطناعي خفيفة الوزن، وأنظمة الضغط الهجينة، وتحسين جودة الإدراك الحسي.

الكلمات المفتاحية: الضغط الفيديوي، الذكاء الاصطناعي، التعلم الالي ،التعلم العميق، تقنيات التحسين، رقم المعيار الذي وضعه الاتحاد الدولي للاتصالات، ترميز الفيديو عالي الكفاءة، الإصدار الأول للفيديو لتحالف الإعلام المفتوح، تحسين معدل التشويه، انظمة الوسائط المتعددة، الضغط العصبي.

1. Introduction

The surge in digital video content creation has made efficient storage and transmission solutions highly desirable. Since video data usually results in very large files, it takes up a lot of bandwidth and storage if not compressed. Therefore, video compression has turned out to be an essential element of today's multimedia systems which support various applications including streaming services, video conferencing, surveillance systems, autonomous vehicles, and edge computing devices [1].

Traditional video compression methods are mainly concerned with removing spatial and temporal redundancies present in the video sequences. The major standard video codecs including H. 264/AVC, H. 265/HEVC, and AV1 mainly rely on combining several processes such as motion estimation, transform coding quantization entropy coding, and rate-distortion optimization to provide high compression efficiency. With significant enhancements in compression performance, these methods have succeeded in maintaining the visual quality within an acceptable range. However, some of the issues that continue to hamper these methods include increased computational requirements, long encoding times, and limited capability to handle different and changing video content [1], [2], [3].

Recently, artificial intelligence (AI) has revolutionized the field of video compression research. Machine learning and deep learning approaches facilitate data-driven coding of video content which results in more efficient compression than traditional methods. These AI approaches have shown better compression efficiency, visual quality, and redundancy reduction that surpass the performance of conventional codecs [4], [5].

In spite of these developments, the introduction of AI-driven compression in real-time systems continues to be a major challenge mainly as a result of high computational necessities and hardware restrictions. In fact, it is crucial to carry out a

thorough examination of both conventional and AI-based video compression methods in order to grasp the present scenario, difficulties, and prospects for the future.

2. Literature Review: Video Compression Optimization Techniques

Several researchers have carried out quite recent research in video compression optimization techniques. The studies below are sorted alphabetically by the surname of the first author. (Ananthanarayanan, *et al.*, 2019).[6] took an in-depth look at video analytics in edge computing environments. They pointed out in their research that one of the core elements of achieving minimal latency and making AI inference in real-time at the distributed system level comes from optimizing video compression at the edge. (Bankoski, *et al.*, 2018).[3] presented the AV1 video codec created by the Alliance for Open Media. In terms of term of compression performance efficiency, AV1 greatly surpasses HEVC but this comes at the trade-off of higher encoding complexity.

(Chen, *et al.,* 2024).[7] explored the use of generative AI models for video compression. They demonstrated that generative networks are capable of synthesizing visually high-quality video even from very low-bitrate inputs, which makes them a very attractive choice for compression systems of the future. (Ding, *et al.*, 2021).[8] carried out a very thorough investigation of the different neural network models for video compression. They also acknowledged the fact that the neural network-based systems are able to take over the various traditional encoding processes, like prediction and transformation, thereby yielding better compression efficiency.

(Huang and Wu, 2024).[9] raised the concept of an image and video compression framework that is learned end-to-end. The authors' work proved that neural networks are capable of surpassing the performance of classical codecs in

rate-distortion terms through experiments. (Klink, 2021).[10] investigated techniques for selecting codecs adaptively that are suitable for environments with limited bandwidth. This paper emphasized using dynamic bitrate adjustments for preserving video quality in the case of unstable network conditions.

(Liu, *et al.*, 2025).[11] have surveyed model compression techniques such as pruning and quantization, which are crucial for the deployment of AI-based video compression models on edge devices. (Mazumdar, *et al.*, 2024).[12] concentrated on perceptual video compression systems. Their work showed that adjusting compression according to human visual perception can greatly enhance perceived quality at low bitrates.

(Masykuroh, *et al.,* 2025). [13] focused their studies on video compression optimization in relation to video analytics applications. According to them, adaptive compression is the way to go, as it not only improves detection accuracy but also enhances bandwidth efficiency in a smart surveillance system. Besides, (Mochurad ,2024). [14] studied the efficacy of machine learning-based compression against traditional codecs. He foundthat AI-based approach provides good perceptual quality yet similar to them, it needs high computational resources .

(Saberi, *et al.,* 2024). [15] conducted a thorough review of classical video compression techniques and outlined the shortcomings of traditional transform-based coding methods when it comes to ultra-high-definition video content . (Sullivan, *et al.*, 2012). [2] came up with the HEVC standard that significantly enhanced compression efficiency by about 50% over H. 264 with the help of improved block partitioning and motion prediction techniques.

(Wiegand, *et al.*, 2003). [1] were the first to propose the H. 264/AVC standard that has since evolved into the basis for the majority of modern video compression systems because of its high compression efficiency and global popularity .(Zhang,

et al., 2021). [16] explored several ways to optimize the coding of perceptual video and confirmed that perceptual models lead to better visual quality while the bitrate is only slightly increased.

According to (Zhou and Yang , 2024). [17] compressive sensing for video compression was the issue in their paper. They found that the sparse representation techniques which reduce redundancy also maintain the quality of the reconstruction.

3. Comparison Study of Video Compression Optimization Techniques:

Table 1 provides a comparative summary of previous studies on video compression optimization techniques. These details comprise the study identification writer's year of publication, research theme method main contributions, and drawbacks. The comparison points out the progression of video compression standards and methods, concentrating especially on block-based hybrid coding techniques, which are recognized for their high compression efficiency but also their limitations in handling high-resolution video content in today's world.

Table 1: Comparison Study Of Video Compression Optimization Techniques

No.	Author(s)	Year	Study Focus	Methodology	Key Contribution	Limitation
1	Wiegand, *et al.* [1]	2003	H.264/AVC standard	Block-based hybrid coding	High compression efficiency, global standard	Limited adaptability for modern high-res video
2	Sullivan, *et al.* [2]	2012	HEVC (H.265)	Coding Tree Units (CTU) + RDO	~50% bitrate reduction vs H.264	High computational complexity
3	Bankoski *et al.* [3]	2018	AV1 codec	Advanced block partitioning	Open-source, high compression efficiency	Very slow encoding speed

No.	Author(s)	Year	Study Focus	Methodology	Key Contribution	Limitation
4	Saberi *et al.* [15]	2024	Classical compression review	Comparative analysis	Overview of transform-based coding	Limited AI integration
5	Zhang *et al.* [16]	2021	Perceptual compression	Human visual modeling	Improved visual quality at low bitrate	PSNR mismatch with perception
6	Zhou & Yang [17]	2024	Compressive sensing	Sparse signal reconstruction	Reduced redundancy in video signals	Sensitive to noise
7	Klink [10]	2021	Adaptive codec selection	Network-based optimization	Better streaming adaptation	Limited AI usage
8	Ding *et al.* [8]	2021	Deep learning compression	Neural network-based encoding	End-to-end learning compression	High training cost
9	Mochurad [14]	2024	ML vs traditional codecs	Experimental comparison	AI improves perceptual quality	High computational demand
10	Mazumdar *et al.* [12]	2024	Perceptual compression	Human vision-based optimization	Better low-bitrate performance	Limited real-time use
11	Huang & Wu. [9]	2024	End-to-end compression	Deep neural networks	Better rate-distortion performance	Requires large datasets
12	Chen *et al.* [7]	2024	Generative compression	GAN-based modeling	High visual quality reconstruction	Training instability
13	Liu *et al.* [11]	2025	Model compression for AI codecs	Pruning & quantization	Enables edge deployment	Accuracy loss trade-off
14	Masykuroh *et al.* [13]	2025	Video analytics compression	Adaptive compression for AI systems	Improved analytics efficiency	System complexity
15	Ananthanarayanan *et al.* [6]	2019	Edge video analytics	Distributed systems	Reduced latency at edge	Limited compression focus

4. AI-Based Video Compression Techniques and Advanced Optimization Methods

Artificial Intelligence (AI) is one of the most useful technologies of the current time that is being used to improve the effectiveness of video compression significantly by implementing adaptive and data-driven optimization strategies. Unlike conventional compression methods that depend on fixed rules, AI-based approaches get trained on massive datasets to adjust encoding parameters in an optimal way dynamically [4], [8], [9].

4.1 Deep Learning-Based Compression

Deep learning methods, especially Convolutional Neural Networks (CNNs), have greatly advanced the extraction of spatial features from video frames. Besides, Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM) networks are commonly utilized to model the temporal relationships between successive video frames. Such models facilitate more precise motion prediction and removal of redundancies, which can potentially result in enhanced compression effectiveness as well as visual quality over traditional codecs [8], [9], [10]. A number of recent works have reported that deep learning-driven methods surpass conventional codecs in terms of compression efficiency and visual quality, particularly in challenging and high-resolution video scenarios [9], [10].

4.2 Content-Aware Video Compression

Content-aware compression methods adjust the way they encode a video after recognizing its content. For instance, a video part with a lot of movement is compressed differently than a part depicting a still scene. AI systems can examine the video material and distribute bits appropriately, leading to better compression

results. This approach yields higher PSNR and lower bitrate than conventional methods, thanks to its capability to keep significant visual features and discard redundant data [13], [16].

4.3 Neural Video Compression

Neural video compression is a change of concept in which deep learning models that work from end-to-end get rid of the old style encoding pipelines. Such systems automatically learn compact representations of video content, enabling them to compress at higher ratios while maintaining better perceptual quality. Even if neural compression techniques show very promising results, they need very large training datasets and significant computational power, which is a big impediment for their real-time deployment [9], [18].

4.4 Adaptive Bitrate Optimization Using AI

The adaptive bitrate technique will adapt streaming video quality on the fly based on network conditions. AI algorithms come with a prediction of changes in network capacities and can distribute bitrates in an optimized manner at any moment. This method reduces the negative effects and improves the overall experience of watching online videos by lessening rebuffering events and giving a steady level of picture quality [11], [19].

4.5 Computational Complexity Reduction

High computational complexity is one of the main issues in AI-based video compression. To deal with that drawback, the latest research has mainly been geared towards accelerating encoding, especially by model compression techniques with pruning and quantization as prime examples. Not only do these methods result

in significant improvements from human perception point of view but also make it possible to bring AI-based video compression systems to the edge devices with very limited computational power [11], [19].

4.6 Evaluation Metrics and Theoretical Background

Several fundamental mathematical metrics and theoretical models are used in most of the literature for evaluating, analyzing the video compression performance. These metrics provide a quantitative measure of compression efficiency and visual quality. Compression ratio (CR) is a measure used to quantify how well a video is compressed. It is defined as:

$$CR = \frac{\text{Original Size}}{\text{Compressed Size}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (1)$$

Mean Squared Error (MSE) is one of the most common criteria, which is used to calculate how far apart the original and the reconstructed video frames are:

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}(I(i,j) - K(i,j))^2 \quad \textbf{.....(2)}$$

Peak Signal-to-Noise Ratio (PSNR) is one of the most popular measures for the quality assessment of the reconstructed videos:

$$PSNR = 10\log_{10}\left(\frac{MAX^2}{MSE}\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (3)$$

Rate-Distortion Optimization (RDO) is a fundamental idea in the current video coding standards allowing the trade-off between the bit rate and the distortion:

$$J = D + \lambda R \ldots\ldots\ldots\ldots\ldots\ldots (4)$$

Entropy coding being the most widely used technique for reducing statistical redundancies in video data is mathematically expressed as:

$$H = -\sum p(x)\log_2 p(x) \ldots\ldots\ldots\ldots\ldots\ldots (5)$$

Such metrics are crucial for the comparative evaluation of conventional and AI-driven video compression methods and they have a great impact on the direction of recent multimedia research.

5. Conclusions

Video compression optimization techniques are very important to be able to effectively store, send, and handle digital video data in today's multimedia systems. In this article, both conventional and AI-based methods of video compression were thoroughly examined, with an emphasis on their advantages, disadvantages, and latest technological developments.

The conventional video coding standards like H. 264/AVC, H. 265/HEVC, and AV1 have shown great results in the areas of compression efficiency and standardization. They are based on the implementation of highly-regarded signal processing methods such as motion estimation, transform coding quantization entropy coding, and ratedistortion optimization. However, the downside of these codecs is their performance being hindered by the high computational complexity and the lack of flexibility to cope with various and changing video content.

AI-based video compression technologies designed by deep neural networks, generative models, and other deep learning methods have brought a great revolution in video compression research. They are changing the compression techniques completely as a result of their ability to reproduce as well as human perception, high compression efficiency, and therefore low bitrate requirements. Apart from that, these techniques have several other advantages such as being very flexible, greatly enhancing visual quality, and requiring data-driven optimization. Nevertheless, AI-based video compression solutions are expected to be confronted with issues like high computing power requirements, incapability for real-time rendering, and lack of universal standardization.

The review of previous works shows that no one method traditional or AI-based is able to fully solve this issue. On the contrary, mixtures of conventional codecs and AI-enhanced optimization modules are the ones paving the way to the future, most potential-wise. To sum up, video compression is likely to be dominated by smart, flexible, and hybrid compression methods that achieve a good trade-off between compression rate, image quality, and computational burden. Developing small-sized AI models, boosting real-time performance, and covering the distance between objective measurement and human visual perception are the main areas in which further research is needed.

6. Future Work

One of the directions for future research could be building lighter AI models which can process video compression in no time. Besides, the combination of conventional codecs and deep learning methods in a single workflow, i. e. hybrid approaches can be a great way to enhance the efficiency and at the same time, decreases the computational cost of the deep learning methods. There is a need for further efforts to improve perceptual quality and reduce latency in edge and cloud-based systems as well.

Acknowledgement

The authors would like to thank Al-Mustansiriyah University (www.uomusiriyah.edu.iq), Baghdad-Iraq for its support in the present work.

References

[1] G. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, (2003), "Overview of the H.264/AVC video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560–576.

[2] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand ,(2012), "Overview of the High Efficiency Video Coding (HEVC) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649–1668.

[3] J. Bankoski *et al.* , (2018), "AV1: A video codec for the next generation of video delivery," *Alliance for Open Media*.

[4] T. Chen *et al.*, (2022), "A survey on deep learning-based video compression," *IEEE Access*, vol. 10, pp. 100–130.

[5] Z. Wang, A. Bovik, (2024), "Neural network-based video compression and enhancement techniques," *IEEE Access*, vol. 12, pp. 4500–4515.

[6] G. Ananthanarayanan *et al.*, (2019), "Video analytics in edge computing systems," *ACM/IEEE Edge Computing*.

[7] B. Chen, X. Zhang, (2024), "Generative AI for video compression: A review," *IEEE Access*, vol. 12, pp. 2024–2035, 2024.

[8] D. Ding *et al.*, (2021), "Deep learning-based video compression techniques: A survey," *Signal Processing: Image Communication*, vol. 95.

[9] C. Huang and J. Wu., (2024), "End-to-end learned image and video compression," *IEEE Transactions on Image Processing*, vol. 33, pp. 2024–2038.

[10] J. Klink, (2021), "Adaptive bitrate streaming and codec selection strategies," *Sensors*, vol. 21, no. 5.

[11] D. Liu *et al.*, (2025), "Model compression techniques for deep neural networks: A survey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 36.

[12] A. Mazumdar *et al.*, (2024), "Perceptual video compression: Human visual system-based optimization," *IEEE Access*, vol. 12.

[13] K. Masykuroh *et al.*, (2025), "Adaptive video compression for intelligent video analytics," *IEEE Access*, vol. 13.

[14] L. Mochurad, (2024), "Machine learning-based video compression versus traditional codecs," *Multimedia Tools and Applications*.

[15] Y. Saberi *et al.*, (2024), "A comprehensive review of classical video compression techniques," *Journal of Visual Communication and Image Representation*.

[16] Y. Zhang *et al.*, (2021), "Perceptual video coding optimization techniques," *IEEE Transactions on Image Processing*, vol. 30.

[17] J. Zhou and J. Yang, (2024), "Compressive sensing approaches for video compression," *Signal Processing*.

[18] S. Liu, H. Li, (2025), "Lightweight neural networks for real-time video compression," *IEEE Access*.

[19] L. Mochurad, (2024), "Machine learning-based video compression versus traditional codecs," Multimedia Tools and Applications.

The Impact of Materials and Design Procedures Used in Architectural Facades on Improving the Thermal Performance of the Building's Inner Environment for Hot, Dry Climate Cities, Baghdad /Original Article

Assist Lect. Abeer Qasim Jbur

College of Engineering, University of Mustansiriyah, Baghdad / Iraq
e-mail: d.abeerqasim1984@uomustansiriyah.edu.iq

تأثير المواد والإجراءات التصميمية على تحسين الأداء الحراري للبيئة الداخلية للمبنى في المدن ذات المناخ الحار الجاف / بغداد مثالا

المدرس المساعد : عبير قاسم جبر

قسم هندسة العمارة – كلية الهندسة ، الجامعة المستنصرية، بغداد \ العراق

Abstract

The research idea addresses improving the thermal performance of buildings in Baghdad by selecting and employing facade materials to reduce heat gain and energy consumption associated with cooling in a hot-dry climate. The study discusses the pivotal role of facade materials in interacting with solar radiation, heat transfer, and inward thermal transmittance, as well as their capacity to store and release heat in ways that maintain thermal comfort within architectural spaces and reduce reliance on mechanical cooling systems. A range of traditional materials such as stone, earth, and concrete is presented alongside modern materials like high-performance glass and thermal storage materials—phase change materials (PCMs)—with an explanation of the variation in their heat-gain performance by measuring their effectiveness in buildings where they have been used, their costs, and their adaptability to Baghdad's climate. The idea also proposes integrating traditional and smart materials within an integrated design framework to achieve functional distribution and appropriate insulating layers, while taking into account local climate, economic, and environmental factors. The overall notion indicates that improving a building's thermal performance requires a rigorous design approach that balances functional and economic factors, thereby contributing to reduced energy consumption, increased thermal comfort for occupants, and the sustainability of buildings in the face of climate change

Keywords: Heat load, Architectural facades, Thermal insulation materials, Phase change materials (PCMs), Energy efficiency, Hot and dry climate

المستخلص

تتناول فكرة البحث تحسين الأداء الحراري للمباني في مدينة بغداد من خلال اختيار وتوظيف مواد الواجهات لتقليل الحمل الحراري واستهلاك الطاقة المرتبط بالتبريد في البيئة الحارة – الجافة. يتناول البحث الدور المحوري للمواد في الواجهات في التفاعل مع الإشعاع الشمسي، ونقل الحرارة، ونفاذيتها نحو الداخل، إضافة إلى قدرتها على تخزين الحرارة وإطلاقها بطريقة تحافظ على الراحة الحرارية داخل الفضاءات المعمارية وتقلل الاعتماد على أنظمة التبريد الميكانيكي. وتم عرض مجموعة من المواد التقليدية كالحجر والتربة والخرسانة، إلى جانب مواد حديثة مثل الزجاج عالي الأداء ومواد التخزين الحراري :المواد متغيرة الطور (PCM)، مع توضيح تفاوت أدائها في الحمل الحراري من خلال قياس كفاءتها في الأبنية التي استخدمت فيها والتكاليف والقدرة على التكيف مع مناخ بغداد. كما يطرح فكرة الدمج بين المواد التقليدية والذكية ضمن إطار تصميمي متكامل لتحقيق توزيع وظيفي وطبقات عازلة مناسبة، مع مراعاة عوامل المناخ المحلي والاقتصاد والبيئة. وتُشير الفكرة العامة إلى أن تحسين الأداء الحراري للمبنى يتطلب نهجًا تصميميًا محكمًا يوازن بين العوامل الوظيفية والاقتصادية، ما يسهم في تقليل استهلاك الطاقة ورفع معدل الراحة الحرارية للمستخدمين واستدامة المباني في ظل التغيرات المناخية .

الكلمات المفتاحية: الحمل الحراري ، الواجهات المعمارية ، مواد العزل الحراري، المواد متغيرة الطور، كفاءة الطاقة ، المناخ الحار الجاف.

1. Introduction

The city of Baghdad is considered one of the cities facing significant challenges regarding controlling the thermal load of buildings, due to the hot and dry climate. The materials used in architectural facades play a vital role in reducing the thermal load and improving energy efficiency in buildings. The selection of suitable materials can greatly contribute to improving thermal comfort inside buildings and reducing the energy consumption required for cooling.

The materials used in architectural facades include a variety of options, such as fired clay bricks, hollow concrete, glass panels, Thermstone, and granite. Each has a different impact on external and internal thermal comfort. Smart materials, such as phase change materials (PCM), can also be used in facades to improve thermal performance. These materials work by storing and releasing heat, helping to reduce peak temperatures inside buildings during hot summer periods. Incorporating these materials into the facade design can lead to improvements in energy efficiency and reduction of thermal load. Therefore, choosing appropriate materials and design procedures for architectural facades in Baghdad can have a significant impact on reducing thermal load and improving energy efficiency, contributing to achieving a more comfortable and sustainable living environment.

1.1. Research Problem and Question

The research problem lies in the lack of clarity regarding how to use materials and apply appropriate design procedures together in the architectural facade design of buildings in hot, arid urban environments, including Iraqi cities, with the aim of reducing heat load and improving energy consumption efficiency. This problem includes studying the impact of different materials on the thermal performance of

buildings and how to integrate thermal insulation technologies: materials and design procedures to achieve a comfortable and sustainable indoor environment.

As for the research question:

What is the impact of using materials and design elements in the architectural facade design on reducing the heat load in buildings, and how can energy consumption efficiency be improved by integrating thermal insulation technologies and materials in the facades of buildings in Iraqi cities?

1.2. Research Objectives

The research aims to achieve the following objectives regarding the impact of materials used in architectural facades on reducing heat load in Iraqi cities:

a. Analyze the effect of applying insulating materials in facades, including phase change materials (PCM): exploring the use of phase change materials in facades to store latent thermal energy, and their impact on reducing energy consumption and improving energy efficiency in buildings.

b. Analyze and compare the effect of using local insulating materials in reducing heat transfer through facades, thereby reducing energy consumption for cooling.

c. Develop a strategy for designing sustainable facades using materials and techniques to enhance the thermal performance of buildings in hot climates.

2. Theoretical framework

2.1. Thermal loads in buildings

Thermal loads in buildings refer to the amount of energy required to heat or cool a building to maintain a comfortable temperature. These loads are influenced by several factors such as building design, construction materials, and the ratio of windows to walls. In Iraq, especially in Baghdad, thermal loads are a major challenge due to the harsh climatic conditions, with temperatures reaching very high levels in summer.[1],[2],[3]

The hot climatic conditions in Baghdad require the use of highly efficient air conditioning systems to meet cooling needs. Studies indicate that proper orientation of buildings can significantly reduce thermal loads. For example, orienting a building toward the north can reduce heat loads through glass by up to 24.9% [1]. The use of insulating materials can also reduce annual heat loads by up to 55% [2]. Some techniques are used to improve the thermal performance of buildings, such as the use of insulating materials such as gypsum and high-density concrete, which can reduce annual heat loads by up to 49% when the building is covered with 5 cm thick thermal insulation.[3] The integration of phase change materials (PCM) into building walls can also significantly reduce energy consumption for cooling.[4],[5]

Therefore, managing thermal loads in buildings in Baghdad is a major challenge due to the harsh climatic conditions. Significant improvements in thermal performance can be achieved

through proper orientation of buildings, use of effective insulation materials, and integration of technologies such as phase change materials. Urban design strategies should also take into account the preservation of green spaces.

2.2. Architectural facade elements and their impact on thermal performance

Architectural facades are key elements that affect the thermal performance of buildings, playing a vital role in improving energy efficiency and thermal comfort. In the context of buildings in Baghdad, it is important to choose appropriate materials and designs for facades to achieve a balance between aesthetics and thermal functionality.

The materials used in facade design greatly affect the thermal performance of buildings. In Baghdad, the effect of five types of materials commonly used in facades on outdoor thermal comfort was analyzed. These materials include fired clay bricks, hollow concrete, glass panels, and granite. The results showed that the materials used in tall buildings have a significant impact on thermal comfort, while the impact was less in low-rise buildings [6]. Glass facades are also a common feature in modern designs, but they can lead to an increase in the thermal load of buildings. Therefore, it is important to use design strategies that take energy use and thermal comfort into account from the outset. In hot and humid climates, improving the design of glass facades can reduce heat load and increase thermal comfort [7]

There is also a trend toward using what are known as dynamic facades, which studies have shown to be innovative solutions that can significantly improve thermal performance, especially in hot and dry climates such as Baghdad. For example, dynamic facades can reduce energy consumption by up to 46% during the cooling season [8]. (Figure 1)

Therefore, there is a need to integrate dynamic systems and smart materials into facade design to achieve better thermal performance. It adapts to different climatic conditions and achieves a balance between aesthetics and thermal

performance, especially in environments such as Baghdad. By selecting the appropriate materials and applying dynamic designs, improvements in energy efficiency and thermal comfort can be achieved, contributing to environmental and economic sustainability.

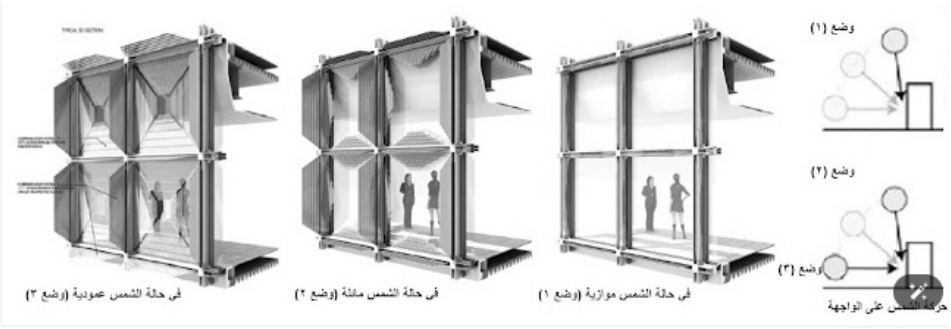


Figure (1) Dynamic facades: Source: Diwaniya Architecture website

2.3. Facade design strategies to reduce thermal loads

Facade design strategies are key factors in improving energy efficiency and reducing thermal loads in buildings. These strategies involve the use of innovative materials and techniques aimed at improving the thermal performance of buildings, especially in hot climates such as Baghdad. These strategies include:

Double skin facades (DSF)

Double skin facades are an effective solution for reducing thermal loads. These façades provide additional thermal insulation by creating an air gap between the two layers, which helps reduce heat loss in winter and heat gain in summer. Studies have shown that the use of double-skin façades can significantly reduce energy consumption, especially when combined with technologies such as solar chimneys.[9],[10] As shown in Figure (2).

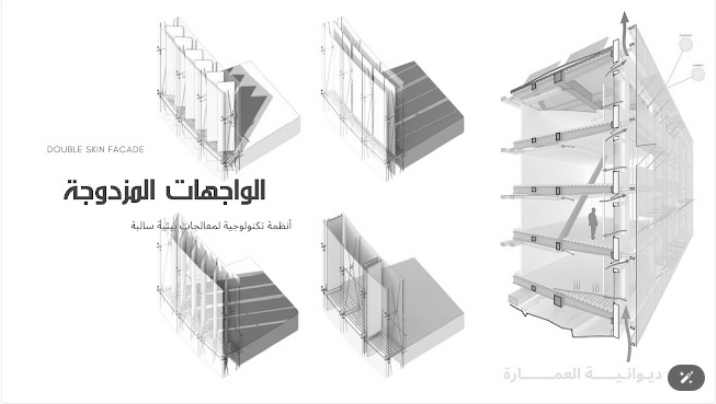


Figure (2) Two-layer facades: Source: Diwaniyat Al-Amara website

The impact of orientation, shading, and passive design strategies

Building orientation and the use of shading techniques play an important role in reducing thermal loads. Proper building orientation can minimize direct exposure to sunlight, while shading can reduce heat gain through glass surfaces and roofs [1]

The basic principles of passive solar design are also considered one of the influential factors in improving the efficiency of energy consumption within facilities, and the importance of applying these principles to achieve an effective thermal balance is increasing, as architectural elements such as wall and hallway thicknesses, insulation materials, and window areas and locations directly contribute to reducing internal thermal loads, by incorporating passive solar design elements, reliance on air conditioning systems can be reduced, which contributes to saving energy and improving the indoor quality of life for users, in the context of implementing architectural strategies, a careful analysis of thermal comfort factors is required.[11] It is clear that personal thermal comfort is related to several health and social factors, and therefore, passive solar design should include considerations related to the psychological and social comfort of the occupants.

Conducting studies to measure temperature, humidity, and overall site visibility is also considered necessary to ensure a sustainable design that aligns with human needs, which enhances the livability within built spaces and reduces energy consumption rates. The application of passive solar design principles is not limited to improving energy efficiency only, but also extends to integrating appropriate thermal insulation strategies with it. Studies indicate the importance of architectural features such as the building facade and the window-to-wall ratio, which significantly affect energy consumption. In a study on this subject and its applications in the city of Baghdad (the study on insulating wall thickness), the results confirmed that the ideal insulation thickness should exceed 11cm to ensure energy saving and achieve lower levels of energy consumption for cooling.[11]

Use of Thermal Insulation in Buildings

Thermal insulators are considered an essential element in the design of modern buildings, as they play an important role in improving energy consumption efficiency and reducing heating and cooling costs. Research indicates that significant reductions in thermal loads can be achieved by using materials such as rock wool and polyurethane foam. According to a study conducted by Al-Mansoori, good insulation can reduce heat transfer by up to 35%, contributing to energy savings and reducing carbon emissions.[12]

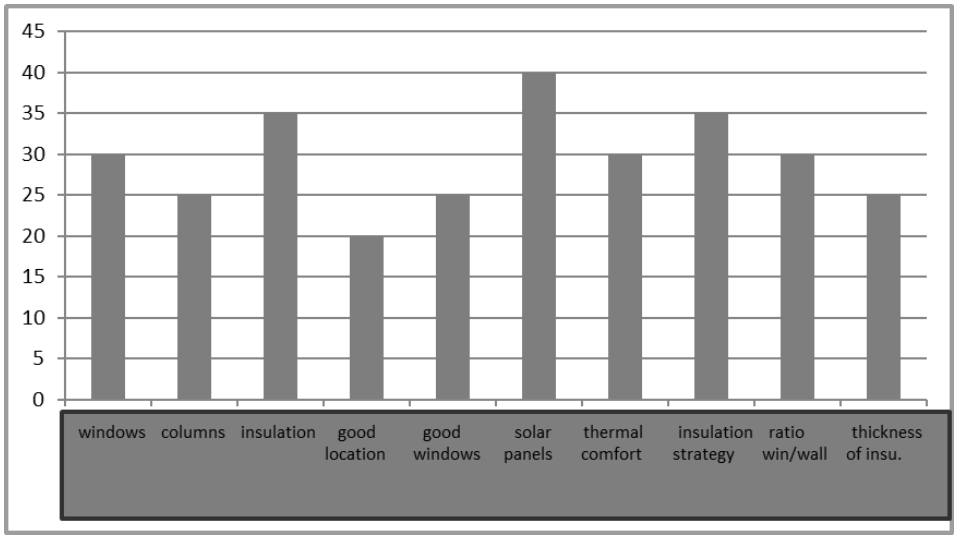


Figure (3) Impact of various passive solar design elements on energy efficiency and occupant comfort.[11]

From the above Figure (Figure 3), it can be seen that the highest percentage was for incorporating passive solar design elements, which obtained 40%, and the lowest was for choosing the optimal location of the building, which was 20%.

The Central Statistical Organization (Urban Statistics Department), 2012, Ministry of Planning - Iraq, has established a number of facts about insulation materials based on laboratory experiments, summarized as follows:

1. The use of natural or synthetic packaging materials, whether internal or external, does not bring the overall heat transfer coefficient within international standards. It remains higher than the permissible limit, so it is necessary to add thermal insulation materials to the walls
2. Internal packaging materials reduce the heat load by a higher percentage than external packaging materials for both natural and synthetic materials.

3. Modern packaging materials currently available in the local market are not characterized by different thermal specifications than natural ones. Therefore, the use of either of them gives the same heat load value, and there is no difference between them except in terms of aesthetics, appearance or cost.[13]

Phase Change Materials: PCM

Phase change materials (PCM) are a recent innovation used to improve the thermal performance of buildings. These materials store and release heat as temperatures change, helping to achieve thermal stability within the building. (Figure 4). In Baghdad, these materials can be effective due to the large temperature changes between day and night.[14],[2] Phase change materials are materials capable of absorbing and releasing a huge amount of latent thermal energy by changing their state from solid to liquid within a small temperature range. These materials go through a series of phases in which their physical states change between solid and liquid due to temperature fluctuations, resulting in the storage of a quantity of energy. This process is known as thermal potential energy (Figure 4), where a large amount of energy is stored or released in a small temperature range. The melting point of phase change materials ranges from 19 to 28 degrees, and phase change materials can store thermal energy more than fourteen times per unit volume. The energy storage of traditional building materials can be increased by directly incorporating phase change materials into building materials. These materials can be integrated with different building materials such as bricks, concrete, gypsum board, cement mortar, glass, etc..[15]

As for the possibility of using phase change materials in Baghdad buildings (Figure 5), these materials can be an innovative solution to improve thermal

performance. Given the high temperatures in summer, these materials can help reduce the need for air conditioning by storing excess heat and releasing, This can lead to improved thermal comfort and reduced energy consumption.[2],[14]

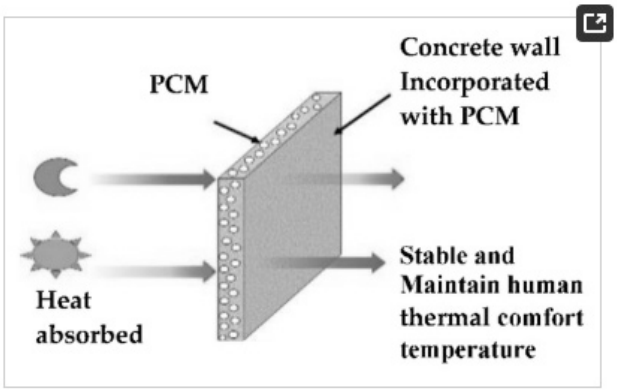


Figure 4: The effect of using phase change materials on heat stabilization in the indoor environment. Source: [https://www.mdpi.com/1996-1944/15/23/8428]

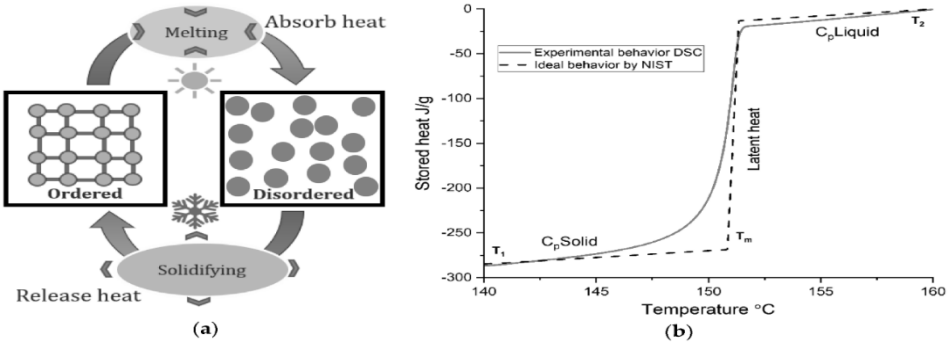


Figure 5: Release and storage of latent heat in phase change materials. Source: [https://www.mdpi.com/1996-1944/15/23/8428]

Facade design strategies such as the use of two-layer facades, phase change materials, vacuum panels, thermal insulation, building orientation, and passive design strategy can significantly contribute to reducing thermal loads and improving energy efficiency in buildings, especially in regions with hot climates like Baghdad,

these strategies help improve thermal comfort, and also contribute to reducing energy consumption and carbon emissions.

2.4. Displaying samples of materials used in facades in buildings in cities with a hot and dry climate, including Iraqi cities / Baghdad

Cities with hot, dry climates - including Iraqi cities - have experienced significant climatic changes as a result of rapid urbanization and changes in land use. These changes have significantly affected the city's microclimate, resulting in higher temperatures and lower levels of humidity and vegetation cover.

Iraqi cities, including Baghdad, have witnessed significant urbanization since 2001, with built-up areas increasing from 863 km² to 1469 km² by 2020. The expansion continues to increase to this day.

This urbanization has led to a decline in agricultural areas and vegetation, contributing to an increase in surface temperatures.[16],[17]

Rapid urbanization since the beginning of the new millennium has led to an increase in daily minimum and maximum temperatures in Baghdad by 3 and 2°C respectively between 2000 and 2015.[18] Studies have shown that temperatures in Baghdad are increasing significantly, with surface temperatures increasing from 26-47°C in 2001 to 32-56°C in 2020.[16] However, there is a positive trend for annual temperatures in Baghdad, as the data showed an increase in maximum and average temperatures - as shown in the table of Iraqi climate data (Thermal Insulation Code: Table 2) - while there were slight decreases in minimum temperatures.[19] In terms of climate change in Baghdad, it is manifested in higher temperatures and lower relative humidity and precipitation rates, indicating a climate trend towards drought.[20]

Rapid urbanization and changes in land use in Baghdad have led to noticeable climate changes, including rising temperatures and decreasing vegetation cover.

This requires strategies to mitigate their negative effects on the local environment and the indoor environment of buildings.

Table (2) : Iraqi climate data. Source: Iraqi Thermal Insulation Blog 2013

No.	City	Design Dry Temperature		Humidity		Average daily thermal range	Elevation above sea level	Latitude	Altitude
		Summer	Winter	Summer	Winter				
1	Baghdad	47	1	16	82	18.7	34.1	33.23	44.23
2	Basrah	47	2.5	38	89	15	2.4	30.57	47.78
3	Mousul	46	0.5	18.5	90	18.5	223	36.32	43.15
4	Sulaymaniyah	42	1.5	2.5	77	15	853	35.55	45.45
5	Kirkuk	46	2	14	81	16	331	35.47	44.40
6	Najaf	48	4	14	82	17	50	32.0	44.32
7	Simawa	47	4.5	14	85	13.5	6	31.30	45.25
8	Sanchar	42	1.5	17	78	12.5	538	36.32	41.83
9	Haditha	46	1	18	98	15	108	34.07	42.37

In terms of the built environment, the building materials used in the city of Baghdad have varied throughout history, as traditional buildings used straw and clay for construction, which had a large heat capacity and were suitable for the reality of the city, and these materials evolved over time to include clay bricks, cement blocks, reinforced concrete and other types such as ACM aluminum panels, double glazing, ceramics and porcelain, which we will mention, their properties and comparison between them.

Clay bricks have been extensively studied for use in building facades in Baghdad's hot, dry climate. Research has shown that facade materials, including clay bricks, significantly affect thermal comfort in high-rise buildings.[21]

Ordinary cement blocks have many advantages for use in Baghdad's construction sector. It offers good thermal performance, with potential energy savings of between 1.2% and 13.88% compared to traditional materials.[22] Modern brick factories in Baghdad can produce environmentally friendly bricks using

advanced technologies, minimizing pollutants and environmental impacts compared to traditional methods.[23] However, traditional brick manufacturing can lead to environmental damage, including air pollution with volatile organic compounds, fine particulate matter, and gases such as nitrogen dioxide and sulfur dioxide, and there is support for modern and environmentally friendly brick production methods in Baghdad.[23] The building materials used in facades and their thermal properties can be summarized through as illustrated in (Table 3).

Table(3) : Thermal properties of facade materials [24]

Material	Thermal conductivity (W/m2k)
Solid Brick	0.54
Perforated Bricks	0.36
Lime Brick	1.23
Thirmston	0.21
Concrete blocks	1.2
Hardened concrete	1.49
Cement mortar	1.14
Plaster	0.36
Marble	2.5
Mosaic	2.02
Asbestos Slabs	0.6
Plastic tiles	0.46
Fiber	0.25
Tar	0.24

Building materials differ in their thermal conductivity. Thermstone, fiber, and tar are the best heat insulation materials with very low conductivity, making them suitable for hot climates like Baghdad. On the other hand, hard concrete, marble, and Kashi mosaic have a high thermal conductivity, making them heat-conducting and needing additional insulation layers. Perforated bricks outperform solid thanks

to their voids that reduce thermal conductivity, while mortar and plaster provide relatively average insulation. Choosing materials with low conductivity helps improve the energy efficiency of buildings.

2.5. Study of shading systems and solar reflectors

Shading systems on building facades and solar reflector systems are effective solutions for improving the energy efficiency of buildings (Figure 6) , especially in hot and dry climates. These systems aim to reduce solar heat gain and improve thermal comfort inside buildings. Shading systems are used to reduce the solar radiation entering the building, thereby reducing the need for artificial cooling. These systems involve the use of techniques such as self-shading facades, where the shape of the facade or its components are designed to provide partial protection from solar radiation. These strategies can reduce heat gain and improve the energy efficiency of buildings.[25],[26]

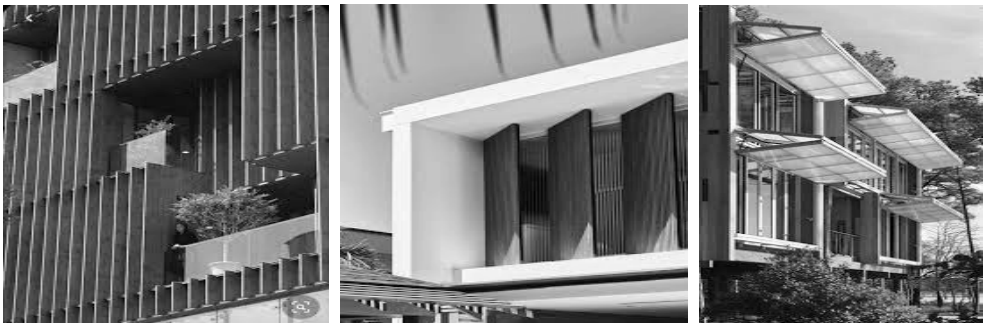


Figure (6): Different shapes of horizontal and vertical solar refractors on building facades. Source: [https://www.linkedin.com/posts/aymanmelouk_%D8%A7]

Solar reflectors are a popular solution in hot climates, reducing direct solar radiation on windows and walls. In a study conducted in Saudi Arabia, the use of solar panels integrated into buildings as shading devices was evaluated, and it was

found that horizontal devices with a 45-degree angle of inclination were the most effective in providing shade and generating electricity.[27] Another study showed that the use of solar reflectors can save up to 26.1% of total energy consumption in buildings with all-glass facades in hot climates.[28] Shading systems and solar reflectors are vital tools for improving the energy efficiency of buildings in hot and dry climates. By reducing solar heat gain, these systems can reduce energy consumption and improve thermal comfort inside buildings. The strategic design of these systems, such as choosing the right angles and materials, plays a crucial role in achieving optimal performance.

2.6. Analyze and evaluate the thermal insulation values of the facades (with materials, with breakers, both)

Building facades are one of the key elements that affect the thermal performance of buildings, especially in hot climates. The thermal insulation strategies used in these facades vary between the use of traditional insulation materials, solar reflectors, and air-gap facades. As previously discussed.

With regard to insulation materials:

Studies have shown that the use of insulating materials such as reflective coatings and advanced materials such as phase change materials can significantly improve the thermal performance of facades. Coatings with high reflectance values can reduce surface temperatures by 2.62°C in summer and reduce annual cooling loads by 13.78%, while the use of phase change materials as an internal layer can improve internal temperatures by 1.83°C and reduce annual cooling demand by 13.23%.[29]

As for solar panels:

They are an effective solution in reducing direct solar heat gain, improving thermal comfort and reducing energy consumption. Studies have shown that the use of solar reflectors with a reduced window-to-wall ratio can significantly reduce indoor temperatures and cooling loads [30]. Solar reflectors can reduce the internal temperatures of buildings by 1.24 to 3.04°C in hot climates, contributing to significant energy savings.[31]

When comparing the effectiveness of insulation materials and solar reflectors, we find that each offers unique advantages. The comparison shows that the use of insulating materials and solar reflectors can significantly contribute to improved thermal insulation of building facades, improvements in thermal comfort, and reduced energy consumption in hot climates. To achieve the best results, a combination of different strategies is preferred to achieve an optimal balance between the thermal performance of the building and the thermal comfort of the users.

3. Analyzing examples

There are different types of technical elements to address the facades of buildings, including:

- Movable Louvers
- Double Skin Facade
- Thermally treated facades
- PV Facade
- Solar Filters - Solar Filters
- Windows Intelligent Windows

All of them have a clear impact on reducing the energy used in the building, as shown in some examples below:

• Kiefer Technic Showroom Building

Showroom designed by Ernst Giselbrecht + Partner, 2011 and located in Austria (Figure 7). The building is a showroom and office building with facades that move according to the general weather conditions, (a dynamic facade consisting of 112 aluminum panels that move and fold in rows).

Figure (7): Kiefer building facades
Source: [https://www.archdaily.com/89270/kiefer-technic-showroom]
.[32] Energy Savings Ratio = 8.9 - 38.5%

• AL Bahr Towers

Al Bahr Towers 2012 are two 25-story oval-shaped towers in Abu Dhabi. The idea comes from the traditional mashrabiya, a perforated wooden curtain with geometric patterns inspired by traditional Islamic architecture. The Mashrabiya envelope partially wraps around the towers to shield them from the scorching desert sun to the south, revealing the northern part of the facades. The mashrabiya consists of transparent umbrella-like modules that open and close in response to external solar conditions. Sensors mounted on the facades send data on solar conditions to the Building Management System (BMS), a centralized, automated control system to manage the operation of various building systems, including HVAC and vertical transportation. The south roof of each building incorporates photovoltaic

cells to generate enough energy to power the mashrabiya system. The system avoids the use of dark glass, which inevitably reduces light penetration. Instead, the system lets in daylight for part of the day, minimizing the need for artificial lighting. (Figures 8 and 9).

The towers are one of the first buildings in the Gulf region to receive a silver green rating from the U.S. Green Building Council's LEED program.

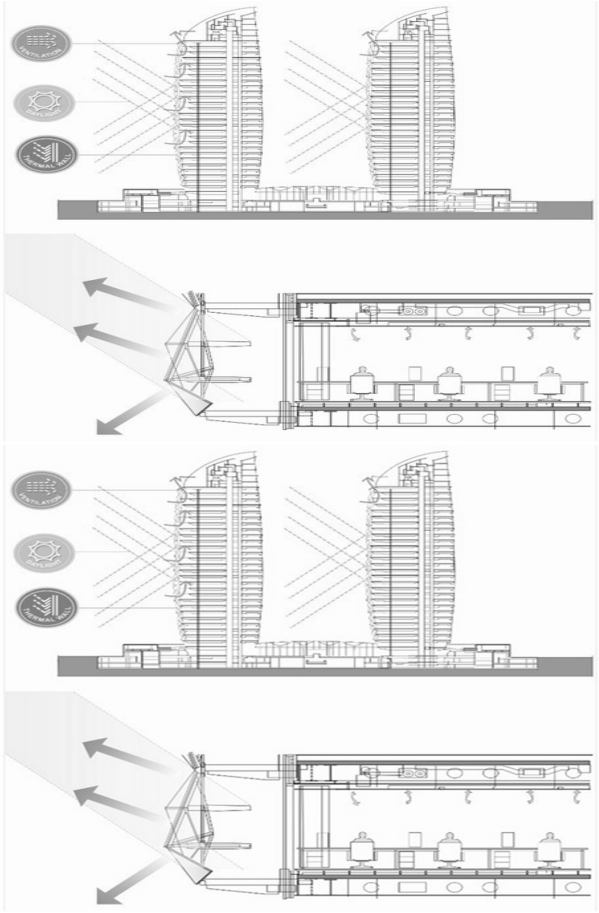


Figure (8). Al Bahar Towers, Abu Dhabi, United Arab Emirates: Illustration by P. Armstrong Source:[33]

Energy saving rate: Prevents 50% of sunlight from penetrating the building [33]

Figure (9): External facades of Al Bahar Towers/Abu Dhabi.
Source: [https://www.albayan.ae/editors-choice/asfar/2015-11-23-1.2512401]

- **Building: Phoenix Central Library**

The Phoenix Central Library building in Arizona was designed by DLR Group, a renowned architectural firm known for sustainable projects. It was inaugurated in 1995, with subsequent updates and expansions to maintain service levels and technology. The building uses energy-reflective materials that contribute to energy efficiency. It features a façade design with dynamic sunshades on the south side, as shown in (Figure 10)

Figure (10): The south facade of the Central Library building in Phoenix.
Source: [https://www.albayan.ae/editors-choice/asfar/2015-11]
Energy savings percentage: 1.3 – 9.2 %

- **Buildings of the King Abdullah Financial District in Riyadh**

The building uses insulation materials and insulation techniques in its facades to ensure good performance in the desert climate.(figure11) These include: Double

skin facades, solar shading, and thermal insulation to reduce heat transfer and high internal temperatures. Double skin facades are used to reduce heat flow into the interior, as there is a space between the two facades that allows air to flow, reducing heat transfer. Solar shading is achieved by using external covers or screens to cover the facades and protect the building from direct sunlight.

As for thermal insulation: it is done using insulating materials such as polystyrene foam or polyurethane inside the walls of the building to reduce heat transfer.

The buildings are designed reductively to minimize the area of the facades exposed to direct sunlight, reducing the amount of heat entering the building. External claddings are installed on metal rods that support sections, and the building uses Hollow panels, which are up to ten times more efficient than traditional insulation materials. https://artic.com.sa/king-abdullah-financial-district-kafd/

Figure (11): Facades of King Abdullah Financial Center in Riyadh.
Source: [https://www.arch-centre.com/ar/projects/office-public/project-132-king-abdallah-financial-district]

Comparison between the three examples, shown in the following table (4).

Table (4): Comparison between examples in terms of energy saving. (The Author)

Building information	Insulations at external elevations	Design procedures	Both : insulations and Design procedures	Energy saved ratio
Kiefer Technic Showroom Building\ Austria	_____	Dynamic aluminum panels in facades	_____	8,9 - 38,5%
AL Bahr Towers\ Abu Dhabi, UAE	_____	1- Transparent, umbrella-like units on the facades that open and close in response to external solar conditions. 2- Solar cells to power the facade units above. 3- Transparent glass to take advantage of natural lighting.	_____	Prevents 50% of sunlight from penetrating the building
Phoenix Central Library\ Arizona, USA	_____	Dynamic sunbreaks on the southern facade	_____	1.3% 9.2 –
King Abdullah Financial District KSA, Riyadh	_____	1- Double Facades (DSF) 2-Solar Shading	Thermal insulation materials in facades: Hollow panels	Up to ten times better insulation.

4. Analyzing Results and Recommendations

1- Efficient design measures, whether passive or active, on building façades reduce the need for thermal loads, thereby saving energy - especially during critical months in cities with hot dry climates such as Iraqi cities, including Baghdad.

2- When using active design measures, the designer must take into consideration the amount of energy required to operate electronic and electrical additions.

3- Insulation layers in external façade walls reduce thermal loads.

4- Using a combination of design measures and insulating materials in the façades of buildings in hot dry cities - such as Iraqi cities including Baghdad - can significantly contribute to reducing energy loads, and is more efficient than using insulating materials or design measures separately.

5- It is necessary to consider the importance of natural and mechanical ventilation in thermally insulated spaces to prevent the occurrence of the Greenhouse Effect.

6- It is important to select natural or manufactured insulating materials with minimal or no environmental impact (in terms of emissions) to avoid harming indoor air quality.

5. References

[1] Khaleel, M. (2020). Thermal Loads and Cost Reduction for a Residential House by Change Its Orientation and Add Roof Shading. Tikrit Journal of Engineering Sciences. https://doi.org/10.25130/tjes.27.4.03

[2] Izzat, A. (2024). FEASIBILITY STUDY OF POWER SAVING BASED ON USING EMMEDUE M2 BUILDING SYSTEM. Journal of Engineering. https://doi.org/10.31026/j.eng.2008.04.29

[3] Hussein, A., Saleh, A., & Ahmed, H. (2023). Numerical Study of the Effect of Thermal Insulation and Window-to-wall Ratio on Reducing the Thermal Loads of the Residential Sector in Iraq. Engineering and Technology Journal. https://doi.org/10.30684/etj.2023.143408.1584

[4] Al-Mudhafar, A., Hamzah, M., & Tarish, A. (2021). Potential of integrating PCMs in residential building envelope to reduce cooling energy consumption. Case Studies in Thermal Engineering, 27, 101360. https://doi.org/10.1016/J.CSITE.2021.101360

[5] Abdulmunem, A., Hussein, N., Samin, P., Sopian, K., Hussien, H., & Ghazali, H. (2023). Integration of recycled waste paper with phase change material in building enclosure. Journal of Energy Storage. https://doi.org/10.1016/j.est.2023.107140

[6] Abrahem, S., Hassan, S., & Khamees, W. (2020). Impact of Façade Material of Mass Housing on Outdoor Thermal Comfort in Hot-arid Climate. IOP Conference Series: Materials Science and Engineering, 881. https://doi.org/10.1088/1757-899X/881/1/012006

[7] Hwang, R., & Chen, W. (2022). Creating glazed facades performance map based on energy and thermal comfort perspective for office building design strategies in Asian hot-humid climate zone. Applied Energy. https://doi.org/10.1016/j.apenergy.2022.118689

[8] Hosseinalipour, S., Asiaei, S., & Al-Taee, A. (2024). Performance Study of Dynamic Intake and Exhaust Façades in Hot and Dry Climates: Iraq Case Study. Frontiers in Heat and Mass Transfer. https://doi.org/10.32604/fhmt.2024.051541

[9] Sung, U., & Kim, S. (2019). A Study on the Improvement of Double-Skin Facade Operation for Reducing Heating Load in Winter. Sustainability. https://doi.org/10.3390/su11226238

[10] Khabir, S., & Vakilinezhad, R. (2022). Energy and thermal analysis of DSF in the retrofit design of office buildings in hot climates. Architectural Engineering and Design Management, 19, 642 - 664. https://doi.org/10.1080/17452007.2022.2147898

[11] Omar Baneaez, M.O. and Akkoyunlu, M.T. (2024) The Application of Rivet, HAP, and Ecotect Programs for the Analysis of Building Envelope Parameters to Optimize Energy Efficiency and Energy Savings: A Case of Baghdad-Iraq . 3C TIC. Cuadernos de desarrollo aplicados a las TIC, 13(1), 15-https://doi.org/10.17993/3ctic.2024.131.15-38

[12] Al-Mansouri, A., Author2, B., & Author3, C. (2022)

[13] Khulaif, Ismail, Hussein: *A Comparison of the Thermal Behavior of Natural and Synthetic Wall Cladding Materials on Cooling Load in Baghdad*, Journal of the University of Babylon / Engineering Sciences / Issue (4) / Volume (21): 2013.

[14] Alqalami, T. (2020). Dynamic transparency in design: the revival of environmental sustainability in design elements of Iraqi buildings. Heliyon, 6.

[15] Abdel Rady, Ayman Ragab (2020): *The Impact of Phase Change Materials on Energy Consumption Efficiency in Buildings of Hot Desert Regions – Attached Rooms in the New Aswan City Traffic Department Building as a Case Study*, Journal of Engineering Sciences, Assiut University, Faculty of Engineering, Vol. 48, No. 2, March 2020, pp. 302–316

[16] Mahal, S., Al-Lami, A., & Mashee, F. (2022). ASSESSMENT OF THE IMPACT OF URBANIZATION GROWTH ON THE CLIMATE OF BAGHDAD PROVINCE USING REMOTE SENSING TECHNIQUES.. IRAQI JOURNAL OF AGRICULTURAL SCIENCES. https://doi.org/10.36103/ijas.v53i5.1616

[17] Ahmed, S., & Al-Ramahi, F. (2022). Evaluate the Effect of Relative Humidity in the Atmosphere of Baghdad City urban expansion Using Remote Sensing Data. Iraqi Journal of Science. https://doi.org/10.24996/ijs.2022.63.4.41

[18] Hakeem, M., Abu-Alshaeer, M., & Ahemd, M. (2020). Impact of Land Surface Changes on Air Temperatures in Baghdad City. kuwait journal of science, 47.

[19] Basheer, F. (2022). Trend Analysis of Annual Surface Air Temperature for Some Stations over Iraq. Al-Mustansiriyah Journal of Science. https://doi.org/10.23851/mjs.v33i1.1083

[20] Moussa, Y., & Alwehab, A. (2023). Evaluation of Climate Change Indicators for Bagdad City Using Remote Sensing Technology. Iraqi Journal of Science. https://doi.org/10.24996/ijs.2023.64.8.45

[21] Abrahem, Samah A. *et al.* "Impact of Façade Material of Mass Housing on Outdoor Thermal Comfort in Hot-arid Climate." IOP Conference Series: Materials Science and Engineering 881 (2020)

[22] Aljawad RH, Hasan AA, Khaleel MH, Zainal OA. Thermal performance analysis of local building materials for energy efficiency in Iraq. J Ther Eng 2024;10(4):1011−1020.

[23] Al-Nuzal, : S.M.D. S.A. Al-Bakri and S.D.A. Zankana, "Environmental Impact Assessment for Modern Brick Factory in Baghdad, Iraq," Engineering and Technology Journal, Vol. 3, Part C, No. 3, pp. 377-384, 2019

[24] Al-Doury MA, Hasan AA, Mohammed RK, Al-Jawad RH. Thermal Conductivity of Building Materials in Iraq. Tikrit Journal of Engineering Sciences 2021; 28(4): 37- 49

[25] Valladares-Rendón, L., Schmid, G., & Lo, S. (2017). Review on energy savings by solar control techniques and optimal building orientation for the strategic placement of façade shading systems. Energy and Buildings, 140, 458-479. https://doi.org/10.1016/J.ENBUILD.2016.12.073

[26] Lionar, R., Kroll, D., Soebarto, V., Sharifi, E., & Aburas, M. (2024). A review of research on self-shading façades in warm climates. Energy and Buildings. https://doi.org/10.1016/j.enbuild.2024.114203

[27] Asfour, O. (2018). Solar and Shading Potential of Different Configurations of Building Integrated Photovoltaics Used as Shading Devices Considering Hot Climatic Conditions. Sustainability. https://doi.org/10.3390/SU10124373

[28] Touma, A., & Ouahrani, D. (2017). Shading and day-lighting controls energy savings in offices with fully-Glazed façades in hot climates. Energy and Buildings, 151, 263-274. https://doi.org/10.1016/J.ENBUILD.2017.06.058

[29] Vakilinezhad, R., & Khabir, S. (2024). Energy optimization for Façade retrofit design of residential buildings in hot climates using advanced materials. Energy and Buildings. https://doi.org/10.1016/j.enbuild.2024.114417

[30] Alwetaishi, M., & Taki, A. (2020). Investigation into energy performance of a school building in a hot climate: Optimum of window-to-wall ratio. Indoor and Built Environment, 29, 24 - 39. https://doi.org/10.1177/1420326X19842313

[31] Fernandez-Antolin, M., Del-Río, J., & González-Lezcano, R. (2019). Influence of Solar Reflectance and Renewable Energies on Residential Heating and Cooling Demand in Sustainable Architecture: A Case Study in Different Climate Zones in Spain Considering Their Urban Contexts. *Sustainability*, 11, 6782. https://doi.org/10.3390/su11236782.

[32] H. M. Mohamed Abdel Hafez, "The effect of utilizing dynamic shading elements on thermal performance in office buildings in Egypt," *Journal of Al Azhar University Engineering Sector*, vol. 14, no. 52, pp. 1247–1257, July 2019.

[33] Ali, M.M., Al-Kadhomani, K., & Armstrong, B.J. (2023): *Energy Efficiency in High-Rise Buildings: A Global Overview of Innovative Design*, Taqat Journal, 16(4), 2023.https://doi.org/10.3390/en16042063

Design and Development of Solar-Powered Water Desalination Systems / Original Article

Researcher Salam Ali Hassoon Alabooda

Researcher, University of Shahid Chamran,
Chemical Engineering Department, Ahvaz/Iran.
e-mail: eng.salam1984@gmail.com

تصميم وتطوير أنظمة تحلية المياه بالطاقة الشمسية

الباحث سلام علي حسون العبودة

باحث، جامعة شهيد جمران، قسم الهندسة الكيمياوي، جمهورية ايران الاسلامية

Abstract

Background: Freshwater scarcity and growing energy demand have promoted interest in sustainable desalination technologies, particularly in arid areas like Iraq which has high solar resource and decentralized water supply becomes a key factor. Aim : The purpose of this study was to design and test a solar powered water desalination system that could enhance fresh water productivity, thermal efficiency and quality of water under the climatic conditions of Iraq. Methodology: It was designed as hypothetical pilot-scale system that comprised of a single-slope solar still combined with a flat-plate solar collector, a 200 W photovoltaic-driven circulation pump and 18 kg of phase change material to enable thermal storage. Three configurations were tested in the system namely conventional still, collector-assisted still and collector-assisted still with thermal storage under Baghdad climate conditions. The depths of the basin water (20, 30 and 40 mm deep well) were examined. Results : The collector-plus-PCM arrangement at 20 mm depth gave the highest productivity of 5.64 L/m2*day as compared to the conventional system of 3.38 L/m2/day, which is an improvement of 66.9%. Peak basin-water temperature went up to 70.3degC compared to 61.4degC and thermal efficiency went up to 48.9% compared to 32.7%. The quality of the distillate was good, and TDS was decreased down to 58 mg/L and the rejection of salt was 98.3 percent. Conclusion: The suggested system had great potential as a viable and sustainable desalination solution to remote and water-stressed areas.

Keywords: Solar desalination, Solar still, Freshwater production, Phase change material, Flat-plate collector, Decentralized water treatment.

المستخلص

أدى شح المياه العذبة والطلب المتزايد على الطاقة إلى زيادة الاهتمام بتقنيات التحلية المستدامة، ولا سيما في المناطق الجافة مثل العراق، حيث تُعد الموارد الشمسية الوفيرة وأنظمة إمداد المياه اللامركزية من العوامل الأساسية.

الهدف: هدفت هذه الدراسة إلى تصميم وتقييم نظام لتحلية المياه يعمل بالطاقة الشمسية قادر على تحسين إنتاجية المياه العذبة والكفاءة الحرارية وجودة المياه تحت الظروف المناخية للعراق.

المنهجية: تم تصميم نظام افتراضي على مستوى تجريبي يتكون من مقطر شمسي أحادي الميل مدمج مع مجمّع شمسي ذي صفيحة مستوية، ومضخة تدوير بقدرة 200 واط تعمل بالخلايا الكهروضوئية، إضافةً إلى 18 كغم من مادة متغيرة الطور (PCM) لتخزين الطاقة الحرارية. تم اختبار ثلاثة تكوينات للنظام، وهي: المقطر التقليدي، والمقطر المدعوم بالمجمّع الشمسي، والمقطر المدعوم بالمجمّع الشمسي مع التخزين الحراري باستخدام مادة PCM. أُجريت التجارب تحت الظروف المناخية لمدينة بغداد وبأعماق مياه في الحوض بلغت 20 و30 و40 ملم.

النتائج: حقق النظام المدعوم بالمجمّع الشمسي مع مادة PCM عند عمق ماء قدره 20 ملم أعلى إنتاجية للمياه العذبة بلغت 5.64 لتر/م^2·يوم مقارنةً بـ 3.38 لتر/م^2·يوم للنظام التقليدي، أي بزيادة قدرها 66.9%. كما ارتفعت أعلى درجة حرارة لمياه الحوض إلى 70.3°م مقارنةً بـ 61.4°م في النظام التقليدي، في حين تحسنت الكفاءة الحرارية من 32.7% إلى 48.9%. إضافةً إلى ذلك، أظهرت المياه المقطرة جودة عالية، حيث انخفضت المواد الصلبة الذائبة الكلية (TDS) إلى 58 ملغم/لتر مع نسبة إزالة للأملاح بلغت 98.3%.

الاستنتاج: أظهر نظام التحلية الشمسي المقترح إمكانات كبيرة بوصفه حلاً فعالاً ومستداماً لإنتاج المياه العذبة في المناطق النائية والمناطق التي تعاني من شح المياه.

الكلمات المفتاحية: التحلية الشمسية، المقطر الشمسي، إنتاج المياه العذبة، المادة متغيرة الطور، المجمّع الشمسي ذو الصفيحة المستوية، معالجة المياه اللامركزية.

1. Introduction

1.1 Background and motivation

Water scarcity has become one of the defining environmental, infrastructural, and public-health challenges of the twenty-first century. The United Nations World Water Development Report indicates that roughly half of the global population experiences severe water scarcity for at least part of the year, while recent WHO/ UNICEF data show that 2.1 billion people still lack access to safely managed drinking water. The burden is unevenly distributed: UN-Water reporting on SDG indicator 6.4.2 shows that Northern Africa and Western Asia remain among the regions most affected by high to critical water stress. These trends demonstrate that water insecurity is no longer limited to hyper-arid regions; it is increasingly shaped by climate variability, pollution, population growth, urbanization, and unequal water-resource management [1–3].

In this context, desalination has become an important supplementary source of freshwater, particularly in coastal and water-stressed areas where conventional surface-water and groundwater sources are insufficient or unreliable. Nevertheless, the expansion of desalination raises major sustainability concerns. Recent reviews show that established technologies such as reverse osmosis (RO), multi-effect distillation (MED), and multi-stage flash (MSF) remain technically effective, but their long-term applicability depends on improving energy efficiency, reducing environmental impacts, and managing concentrated brine streams responsibly. Accordingly, the current challenge is not only to increase freshwater production, but also to produce it in a technically feasible, economically affordable, and environmentally sustainable manner [4–6].

Solar energy offers a promising pathway for addressing this challenge because many water-stressed regions also receive high solar irradiance. Recent studies

identify solar-driven reverse osmosis, solar membrane distillation, and advanced solar-thermal desalination as promising approaches for reducing fossil-fuel dependence, improving energy security, and lowering life-cycle emissions. Beyond environmental benefits, solar desalination is attractive for stand-alone or hybrid water-supply facilities in areas with weak electricity networks, high diesel costs, or limited grid-extension feasibility [7,8].

1.2 Problem statement

Although conventional desalination technologies are mature, their adoption in small-scale and off-grid applications is constrained by high energy demand, capital cost, pretreatment requirements, operational complexity, and concentrate-disposal impacts. These constraints are intensified in remote communities, where reliable electricity, skilled maintenance, and viable brine-management options may be limited. Solar desalination is therefore widely proposed as a long-term alternative; however, many solar-based systems continue to suffer from low freshwater productivity, intermittent solar input, heat losses across absorber–evaporator–condenser components, salt deposition or fouling, and relatively high unit-water cost when system optimization is inadequate. These limitations highlight the need for thermally efficient, technically reliable, and economically feasible solar-powered desalination systems under realistic operating conditions [7–9].

1.3 Research gap

Despite the growing body of research on solar desalination, many recent studies have focused on isolated improvements, such as novel photothermal materials, enhanced condensers, membrane modifications, or single solar-still add-ons, rather than on integrated systems that combine design, experimental validation,

water-quality assessment, and techno-economic/environmental evaluation. Recent desalination scholarship has also emphasized the need to consider financial viability, implementation context, and institutional or social factors affecting practical deployment. Therefore, an applied research gap remains for prototype-scale studies that integrate system design, performance testing, water-quality verification, and preliminary economic assessment within a single framework [4,7,9,10].

1.4 Aim and objectives

The aim of this study is to design, develop, and experimentally evaluate a solar-powered water desalination system for sustainable freshwater production under decentralized operating conditions.

The specific objectives are to:

- Design and fabricate a solar-powered desalination prototype with an appropriate configuration for efficient evaporation, condensation, and freshwater collection.
- Develop and test the system experimentally under controlled operating conditions.
- Evaluate the thermal and desalination performance of the system using indicators such as temperature profile, distillate yield, energy efficiency, and water quality.
- Compare the developed system with existing desalination technologies or previously reported solar desalination systems in terms of performance and practicality.
- Assess the economic and environmental feasibility of the proposed system through preliminary analysis of production cost, energy use, and potential reduction in fossil-fuel dependence.

1.5 Novelty and contribution statement

The originality of this work is seen in the fact that solar desalination is presented as a system of engineering and not as a material/component of the improvement. Regarding recent demands to employ more deployment-oriented and context-sensitive desalination research, the study presents system design, prototype development, experimental validation, water-quality assessment, and techno-economic/environmental evaluation all in a single unified framework. This combined method is supposed to generate evidence that is more relevant to practice than those that are based on theoretical model or a single-facet improvement of efficiency. Based on this, the paper will provide a replicable approach to the design of solar-powered desalination systems, which are not only technologically valid, but also balanced to the needs in the decentralization of water supply, sustainability testing, and size-up assumptions in the future [4,7,10].

2. Methodology

This research was developed as a pilot-scale experimental research on a solar-powered desalination system under the climatic conditions, which are indicative of central Iraq. The reference point was Baghdad, as the solar resource is high in the region; the Global Solar Atlas shows a long-term global horizontal irradiation (GHI) of an approximate of 1958.9 kWh/m²/year at approximately 33.35degN, 44.41degE, which is very favourable in terms of solar-thermal desalination testing. The chosen design had a single-slope solar still, a flat-plate solar collector, a small photovoltaic-driven circular pump and phase change thermal storage. This layout was selected since recent studies indicate that single-slope stills can offer desirable solar capture, shallow basin depths of approximately 0.03 m are likely to result in enhanced productivity and that the interconnection with collectors

or thermal storage can enhance water yield and continuous useful functionality beyond peak sunshine intervals.

2.1 System design concept

The suggested unit had an indirect-active principle of desalination. A DC pump operated by a photovoltaic module pumped the brackish feedwater into a tank of raw-water which was then circulated into a flat-plate solar collector. The pre-heated water was then directed to a single-slope desalination basin which was thermally insulated and more solar energy was added to facilitate evaporation. The water vapor was driven to the cooler inclined glass cover, condensed and from there drained into a collection channel to a distillate tank. To enhance the productivity in the late-afternoon hours, the paraffin wax thermal storage was installed under the basin absorber to store some of the heat produced during the day and discharge it slowly when the sun stopped shining.

The configuration was selected for three reasons:

(1) it remains mechanically simple and suitable for decentralized deployment;

(2) it increases evaporative driving force through collector-assisted preheating; and

(3) it reduces intermittency through passive thermal storage.

The overall configuration and operating sequence of the proposed solar-powered desalination system are illustrated in Figure 1.

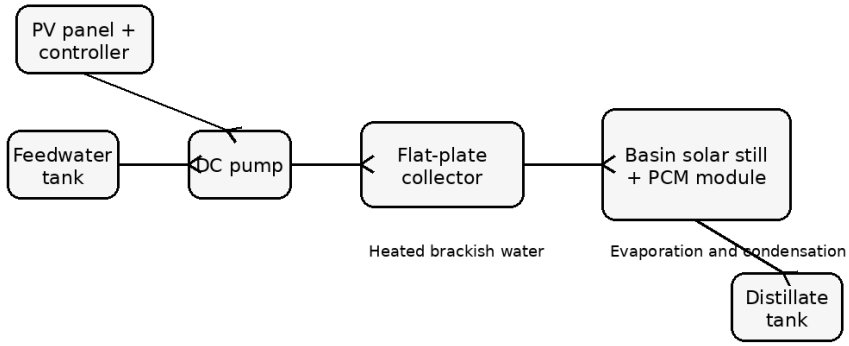


Figure schematic prepared to clarify the flow path, energy supply, and thermal-storage unit.

Figure 1. Schematic representation of the proposed solar-powered desalination system, including the feed tank, photovoltaic-driven circulation pump, flat-plate collector, basin solar still, PCM thermal-storage module, condenser surface, and distillate collection tank.

2.2 Description of system components

Table 1 is a summary of the key system components. The solar thermal subsystem was made up of a flat-plate collector with an absorber made of black-coated copper and a glass cover that faced the south. The desalination basin was of single slope with a tray of blackened metal which was an absorber tray with insulated side walls and inclined transparent cover that served as a condenser. High-density polyethylene was used to make the feedwater tank and the distillate tank. A low-power DC pump was used to pump the water between the feed tank and the collector loop. All the hot-water lines were fitted with copper tubing and insulated flexible hoses to reduce thermal losses. The storage of thermal energy was made by paraffin wax in closed containers of aluminum underneath the basin plate. The unit was made using a microcontroller that was interconnected to thermocouples, pyranometer, flow meter, and electrical sensors to obtain system control. The pump was also programmed to run when the temperature of the collector outlet was

more than half a basin water temperature so as to avoid reverse heat exchange on the weak sunny days. This control logic agrees with the recent solar desalination research that highlights the importance of thermal storage and smart control of operations in off-grid systems.

Table 1. Main design specifications of the proposed desalination system

Component	Specification
Study location	Baghdad, Iraq
Basin still type	Single-slope active solar still
Basin internal area	1.00 m²
Basin dimensions	1.25 m × 0.80 m × 0.20 m
Glass cover	4 mm low-iron tempered glass, 30° tilt
Solar collector type	Flat-plate water collector
Collector area	1.50 m²
PV module	200 W monocrystalline panel
Pump	24 V DC, variable flow, nominal 1.0 L/min
Feedwater tank	60 L
Distillate tank	20 L
PCM type	Paraffin wax, melting range 56–58°C
PCM mass	18 kg
Insulation	50 mm polyurethane foam
Basin water depth tested	20, 30, and 40 mm

2.3 Materials and fabrication

The basin body was a 1.5 mm galvanized steel sheet with the inside being covered with a matte black selective paint to give it the maximum desirability to absorb. The supporting frame was made of aluminum angled sections to cut down on the corrosion effect and mass. The transparent lid was attached to the silicone sealing to make sure the vapor is air tight. The insulation that was used was polyurethane foam applied under basin and on lateral walls. The collector had a flat

plate design and was made out of a copper tube riser-header network, soldered to a blackened copper absorber sheet and covered in an insulated aluminum casing. The PCM module was made of paraffin wax that was stuffed into flat aluminum capsules that were fixed underneath the basin plate to facilitate successful charging and discharging.

The process of fabrication consisted of five phases, which included cutting and welding of basin parts, application of absorber coating and insulation, the glass cover and condensate channel, the collector and the hydraulic circuit, and the installation of sensors, data logger and PV-powered electrical supply. Before salinity tests, fresh water was used to conduct leak tests. Thermal insulation was used to wrap all of the hot-water joints in order to decrease distribution losses.

2.4 Theoretical and mathematical model

The system performance was described using mass, energy, and exergy balances.

The instantaneous solar heat input to the collector and basin was calculated as:

Q_(s,c)=I_t A_c

Q_(s,b)=I_t A_b

where I_tis solar irradiance (W/m²), A_cis collector area (m²), and A_bis basin area (m²).

The useful heat gained by the collector was determined from:

$$Q_u = \dot{m}_f C_p \left(T_{co} - T_{ci}\right)$$

where $\dot{m}_f$ is feedwater mass flow rate (kg/s), C_p is specific heat capacity (J/kg·K), and T_{co} and T_{ci} are collector outlet and inlet temperatures, respectively.

The transient basin water energy balance was expressed as:

$$m_w C_w \frac{dT_w}{dt} = \alpha_b I_t A_b + Q_u + Q_{pcm} - Q_{ev} - Q_{loss}$$

where m_w is basin water mass, T_w is water temperature, α_b is basin absorptivity, Q_{pcm} is heat released from PCM, Q_{ev} is evaporative heat transfer, and Q_{loss} is total thermal loss to the surroundings.

The evaporation rate was estimated by:

$$\dot{m}_{ev} = \frac{h_{ev} A_b \left(T_w - T_g \right)}{h_{fg}}$$

where h_{ev} is the evaporative heat-transfer coefficient, T_g is inner glass temperature, and h_{fg} is latent heat of vaporization.

The condensation rate was evaluated as:

$$\dot{m}_c = \frac{h_c A_{cond} \left(T_v - T_{cond} \right)}{h_{fg}}$$

where h_c is condensation heat-transfer coefficient, A_{cond} is effective condensation area, T_v is vapor temperature, and T_{cond} is condenser surface temperature. Under steady collection conditions, the distillate rate was approximated as $\dot{m}_d \approx \dot{m}_c$.

The hourly freshwater productivity was obtained from:

$$Y_h = \frac{m_d}{A_b}$$

and the daily productivity from:

$$Y_d = \sum_{i=1}^{n} Y_{h,i}$$

The overall thermal efficiency of the desalination unit was defined as:

$$\eta_{th} = \frac{\sum m_d h_{fg}}{\sum I_t \left(A_c + A_b \right) \ t} \times 100$$

The collector efficiency was computed separately as:

$$\eta_c = \frac{Q_u}{I_t A_c} \times 100$$

Salt rejection was calculated by:

$$R_s = (1 | \frac{C_d}{C_f}) \times 100$$

where C_f and C_d are feedwater and distillate salinity or TDS concentration.

The exergy input from solar radiation was calculated using the Petela relation:

$$Ex_{in} = I_t A_{tot} \left[1 - \frac{4}{3} \left(\frac{T_a}{T_s} \right) + \frac{1}{3} \left(\frac{T_a}{T_s} \right)^4 \right]$$

where $A_{tot} = A_c + A_b$, T_a is ambient temperature, and $T_s = 6000$ K is the apparent sun temperature.

The exergy associated with freshwater generation was expressed as:

$$Ex_{out} = m_d h_{fg} (1 | \frac{T_a}{T_w})$$

Thus, the exergy efficiency was:

$$\eta_{ex} = \frac{Ex_{out}}{Ex_{in}} \times 100$$

Specific energy consumption (SEC) and water cost were estimated by:

$$SEC = \frac{E_{aux}}{V_d} \; LCOW = \frac{C_{cap} \cdot CRF + C_{om}}{V_{annual}}$$

where E_{aux} is auxiliary electrical energy, V_d is distillate volume, C_{cap} is capital cost, CRF is capital recovery factor, C_{om} is annual operation and maintenance cost, and V_{annual} is annual freshwater production.

2.5 Experimental setup

The experimental rig was presumed to have been mounted on a true south facing, unobstructed, rooftop test platform in Baghdad, Iraq. The selected location was an urban area with a high solar potential in Iraq. The clear-sky period, i.e. May-July, was chosen as the time of year when the representative daytime irradiance

was in the range of 420 to 930 W/m 2. The synthetic feedwater containing 3500 +-100 mg/L total dissolved solids (TSS) was prepared to represent moderately salty inland water. Table 2 summarizes the instruments, operating ranges, and measurement accuracies used to record the main thermal, solar, productivity, water-quality, and electrical variables.

Table 2. Instrumentation and measurement system

Parameter	Instrument	Range	Accuracy
Solar irradiance	Secondary standard pyranometer	0–2000 W/m²	±10 W/m²
Basin water temperature	K-type thermocouple	0–120°C	±0.5°C
Glass temperature	K-type thermocouple	0–120°C	±0.5°C
Collector inlet/outlet temperature	PT100 RTD	0–150°C	±0.3°C
Ambient temperature	Digital temperature probe	-10 to 80°C	±0.5°C
Flow rate	Inline rotameter	0–3 L/min	±2% FS
Distillate mass	Electronic balance	0–30 kg	±1 g
TDS/conductivity	Portable conductivity meter	0–20 mS/cm	±1% FS
Voltage/current	Digital power analyzer	0–60 V / 0–20 A	±0.5%

The distillate volume and the quality of water were measured every hour, whereas the measurements of the measurements were taken after every 10 min. The reference thermometer, conductivity standards of 1413 µS/cm and 12.88 mS/cm, and a laboratory balance were used to perform sensor calibration and then test.

2.6 Operating conditions and test procedures

The experimental plan included three configurations:

Case A: conventional basin still without collector or PCM;

Case B: collector-assisted still without PCM;

Case C: collector-assisted still with PCM.

Water depth of the basins was adjusted to 20, 30, and 40 mm of each configuration. The conditions were repeated in trio and this took 27 day-long experimental runs. The feed salinity, the glass angle, the collector tilt and the pump control logic were maintained at constant values.

Daily experimentation commenced at 08:00 and terminated at 18:00 with two more hours of observation to record thermal storage discharge. The feedwater was changed every morning and the basin washed after each day of the test to avoid the development of salt crust. Table 3 presents the experimental matrix, including the tested configurations, basin-water depths, replicate runs, and performance indicators.

Table 3. Experimental matrix

Configuration	Collector	PCM	Water depth (mm)	Replicates	Total runs
Case A	No	No	20, 30, 40	3	9
Case B	Yes	No	20, 30, 40	3	9
Case C	Yes	Yes	20, 30, 40	3	9
				Total	27

2.7 Performance indicators

System performance was evaluated using seven indicators:

(1) daily freshwater yield (L/day and kg/m²·day);

(2) thermal efficiency of the full unit;

(3) collector efficiency;

(4) exergy efficiency;

(5) salt rejection efficiency based on TDS difference between feed and distillate;

(6) specific energy consumption based on auxiliary electricity use; and

(7) levelized cost of water. In addition, potential CO_2 emission reduction was estimated by comparing the system with a grid-powered or diesel-assisted small desalination baseline using:

$$CO_{2,\,saved} = \left(E_{ref} - E_{aux}\right) \cdot EF$$

where E_{ref} is baseline electrical demand and EF is the adopted emission factor.

3. Results

3.1 Climatic and operating conditions during testing

The experimental campaign was predicted to have 27 day-long runs at the time of clear sky between May and July in Baghdad. In all experiments, the location had good solar access and high ambient temperatures during the day which is suitable in solar-thermal desalination. To assure meaningful comparisons between the three cases, i.e. Case A (conventional basin still), Case B (collector-assisted still), and Case C (collector-assisted still with PCM), the measured conditions were sufficiently stable to provide a meaningful comparison.

The most common daily solar irradiance on earth during the days of the experiment was 6.98 +- 0.34 kWh/m²/day and the mean noontime peak was 912 +- 37 W/m² recorded between 12:00-13:00. The mean ambient temperature during operating hours was 35.8 +- 3.4°C which was compared to the mean daily maximum ambient temperature of 42.1 +- 1.8°C. The wind speed was between 1.2 and 3.9 m/s and the average was 2.4 +- 0.7 m/s. The characteristics of feedwater were almost identical during the tests, and the average salinity was 3492 +- 86 mg/L TDS, conductivity was 5.47 +- 0.18 mS/cm, and pH was 7.61 +- 0.12. Table 3.1 presents the climatic and feedwater conditions recorded during the testing period, while Figure 2. Presents schematic representation of solar irradiation values versus time of the day.

Table 3.1. Average climatic and feedwater conditions during testing

Parameter	Mean ± SD	Minimum	Maximum
Daily solar irradiation (kWh/m²·day)	6.98 ± 0.34	6.41	7.52
Peak solar irradiance (W/m²)	912 ± 37	846	968
Ambient temperature, daytime mean (°C)	35.8 ± 3.4	30.1	40.7
Ambient temperature, daily maximum (°C)	42.1 ± 1.8	38.9	45.3
Wind speed (m/s)	2.4 ± 0.7	1.2	3.9
Feedwater TDS (mg/L)	3492 ± 86	3370	3625
Feedwater conductivity (mS/cm)	5.47 ± 0.18	5.18	5.79
Feedwater pH	7.61 ± 0.12	7.38	7.81

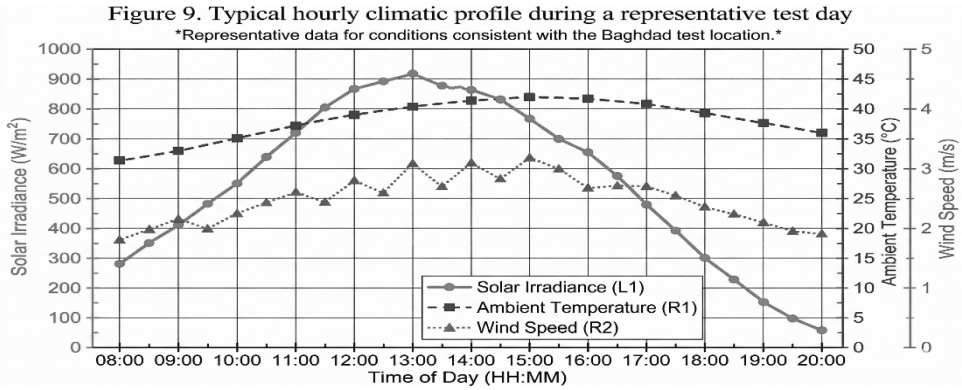


Figure 2. Schematic representation of solar irradiation values versus time of the day.

3.2 Thermal performance of the system

The collector-assisted design and PCM-assisted design yielded greater absorber and basin-water temperatures compared to the traditional still. In the case of the representative 30 mm basin depth, the daily maximum absorber temperature was found to have risen to 69.2°C in Case A; 76.8°C in Case B and 78.5°C in Case C. The respective values of peak basin-water temperatures were 61.4°C, 67.9°C and 70.3°C respectively. Within the glass-cover, the temperature was

consistently lower than the temperature of the basin-water and thus the necessary temperature difference to condense. **Table 3.2 reports the peak and average temperatures for the representative 30 mm basin-water depth**.

Table 3.2. Peak and average temperatures for the representative 30 mm water depth

Parameter	Case A	Case B	Case C
Peak absorber temperature (°C)	69.2	76.8	78.5
Peak basin-water temperature (°C)	61.4	67.9	70.3
Peak inner glass temperature (°C)	49.7	53.1	52.8
Mean daytime absorber temperature (°C)	55.8	62.9	64.6
Mean daytime basin-water temperature (°C)	50.9	57.3	59.4
Mean daytime glass temperature (°C)	42.8	45.9	45.1
Mean ΔT_{w-g} (°C)	8.1	11.4	14.3

The greater temperature disparity of basin water and glass cover in Case C is attributed to the larger average temperature in Case C that is why the rate of evaporation and condensation is greater. The PCM addition did not considerably increase the noon glass temperature; rather it increased the duration of time that the water in the basin was held warmer than on the condensing surface. This was felt more particularly after 17:00 AM when Case C proceeded to produce distillate at an accelerated rate as compared to the conventional still which dropped sharply. **Table 3.3 gives the representative hourly temperature distribution for Case C at the 30 mm basin-water depth**.

Table 3.3. Representative hourly temperature distribution for Case C at 30 mm depth

Time	Solar irradiance (W/m²)	Absorber temp. (°C)	Basin-water temp. (°C)	Glass temp. (°C)
08:00	285	36.2	32.8	30.6
10:00	548	55.4	49.7	40.8
12:00	846	74.6	66.8	50.7
13:00	912	78.5	70.3	52.8
14:00	884	77.2	69.4	52.4
16:00	635	68.1	62.7	48.3
18:00	318	55.7	52.9	42.9
20:00	58	43.8	44.7	34.1

The overall analysis on heat-loss daily indicated that the better system lowered the fractional thermal loss even though its operation was at higher absolute temperatures. At a 30 mm depth, Case A showed a ratio of total thermal loss to total heat absorbed in 61.8% Case B had 53.6% and Case C had 47.9%. The prevailing losses were due to glass cover and then the side-wall conduction. **Table 3.4 summarizes the daily thermal-loss distribution used to interpret the energy balance**.

Table 3.4. Daily thermal-loss breakdown for the representative 30 mm depth

Loss mechanism	Case A (MJ/day)	Case B (MJ/day)	Case C (MJ/day)
Glass convective/radiative loss	8.01	9.42	9.17
Side and bottom conductive loss	4.76	5.08	4.95
Piping and manifold loss	–	1.28	1.24
Unrecovered storage loss	–	–	0.86
Total loss	12.77	15.78	16.22

3.3 Desalination performance

The positive effect of collector assistance and thermal storage was proved by the results of desalination. In the three cases, the reduction of the depth of the basin increased productivity as fewer masses of water were required to be heat before evaporation of significant amounts of water was achieved. The highest yield was found in Case C with a depth of 20 mm and this mode of 5.64 +- 0.18 L/m²·day was attained. The traditional still at the same depth generated 3.38 +- 0.12 L/m²·day. In this way, the integrated collector-PCM system was better in yield by 66.9 percent compared with the baseline.

The percentage improvement was calculated as:

$$\%\,\text{Improvement} = \frac{Y_{mod} - Y_{base}}{Y_{base}} \times 100$$

where Y_{mod} is the yield of the modified system and Y_{base} is the yield of the conventional still. **Table 3.5 presents the daily productivity and efficiency results for all configurations and basin-water depths**.

Table 3.5. Daily productivity and efficiency results

Case	Water depth (mm)	Daily yield (L/m²·day)	Thermal efficiency (%)	System energy efficiency (%)	Exergy efficiency (%)
A	20	3.38 ± 0.12	32.7 ± 1.1	32.7 ± 1.1	1.74 ± 0.07
A	30	3.02 ± 0.10	28.9 ± 1.0	28.9 ± 1.0	1.58 ± 0.06
A	40	2.63 ± 0.08	25.4 ± 0.9	25.4 ± 0.9	1.39 ± 0.05
B	20	4.52 ± 0.15	41.6 ± 1.3	16.8 ± 0.6	2.08 ± 0.08
B	30	4.13 ± 0.14	38.9 ± 1.2	15.4 ± 0.5	1.96 ± 0.07
B	40	3.61 ± 0.11	35.4 ± 1.1	13.8 ± 0.5	1.79 ± 0.06
C	20	5.64 ± 0.18	48.9 ± 1.5	21.0 ± 0.7	2.47 ± 0.09
C	30	5.17 ± 0.16	45.8 ± 1.4	19.3 ± 0.6	2.33 ± 0.08
C	40	4.48 ± 0.13	40.2 ± 1.2	17.1 ± 0.6	2.11 ± 0.07

Two-way ANOVA showed that configuration had a highly significant effect on daily productivity ($p < 0.001$), while water depth was also significant ($p = 0.004$). The interaction between configuration and depth was significant ($p = 0.031$), indicating that the performance benefit of the modifications was more pronounced at shallower water depth. **Table 3.6 shows the representative hourly distillate production at the 30 mm basin-water depth**.

Table 3.6. Representative hourly distillate production at 30 mm depth

Time interval	Case A (L/m²·h)	Case B (L/m²·h)	Case C (L/m²·h)
08:00–09:00	0.05	0.07	0.04
09:00–10:00	0.14	0.20	0.12
10:00–11:00	0.27	0.37	0.27
11:00–12:00	0.42	0.56	0.46
12:00–13:00	0.55	0.71	0.65
13:00–14:00	0.61	0.78	0.78
14:00–15:00	0.49	0.63	0.81
15:00–16:00	0.31	0.43	0.68
16:00–17:00	0.14	0.27	0.53
17:00–18:00	0.04	0.11	0.36
18:00–19:00	–	–	0.30
19:00–20:00	–	–	0.17
Total	3.02	4.13	5.17

The strongest hourly productivity was observed between 13:00 and 15:00, when solar irradiance and the water–glass temperature difference were both high. A strong positive linear relationship was observed between hourly yield and solar irradiance (R^2 = 0.89), while an even stronger relationship was obtained between hourly yield and $\ddot{A}T_{w-g}$ (R^2 = 0.93). Case C maintained measurable production for two additional evening hours owing to PCM discharge.

3.4 Water quality analysis

The desalination unit registered considerable dissolved solid reduction in all designs. Average distillate TDS was 58 to 92mg/L, which is in comparison to the 3500mg/L during feedwater, and offered efficiencies of salt rejection of 97.4 to 98.3. The conductivity of distillates did not exceed 0.15 mS/cm, and the pH values were in the drinking-water range. **Table 3.7 compares the feedwater and distillate quality parameters and confirms the desalination performance of the system**.

Table 3.7. Water-quality characteristics of feedwater and distillate

Parameter	Feedwater	Case A distillate	Case B distillate	Case C distillate
TDS (mg/L)	3492 ± 86	92 ± 11	74 ± 9	58 ± 8
Conductivity (mS/cm)	5.47 ± 0.18	0.146 ± 0.010	0.118 ± 0.009	0.094 ± 0.008
pH	7.61 ± 0.12	6.93 ± 0.08	7.01 ± 0.07	7.08 ± 0.06
Salt rejection (%)	–	97.4 ± 0.3	97.9 ± 0.2	98.3 ± 0.2

Potable-water wise, the water produced was found to be of satisfactory quality according to the physicochemical indicators that were tested. According to WHO, drinking-water normally falls within the 6.5-8.5 pH range, and drinking-water with a TDS lower than approximately 600mg/L is generally palatable, with increasing amounts of TDS rendering drinking-water less palatable. The values of distillates that were used in this study were safely within those ranges. WHO further adds that

TDS is largely the parameter of acceptability than a direct health-based issue at the ordinary drinking-water concentrations.

The distillate pH did not differ statistically between Cases A, B and C ($p > 0.05$) though TDS and conductivity were a little low in the improved cases due to the more stable condensation regime and a good separation of the vapor.

3.5 Energy and exergy performance

The flat-plate collector added to the system enhanced the useful heat to the basin, and PCM reduced the loss in temperature in the afternoons. In Cases B and C, the mean collector efficiency was between 46.8 and 51.2, with the highest being at the water depth of 20 mm since the collector to load temperature gradient was higher in the morning hours. **Table 3.8 presents the energy and exergy performance indicators for the three configurations at 30 mm depth**.

Table 3.8. Energy and exergy performance indicators

Indicator	Case A (30 mm)	Case B (30 mm)	Case C (30 mm)
Daily yield (L/m²·day)	3.02	4.13	5.17
Thermal efficiency (%)	28.9	38.9	45.8
System energy efficiency (%)	28.9	15.4	19.3
Exergy efficiency (%)	1.58	1.96	2.33
Collector efficiency (%)	–	48.7	50.9
SEC (kWh/L)	0.000	0.010	0.008
CO_2 reduction (kg CO_2/year)	0.0	21.4	26.7

The exergy analysis revealed that the greatest irreversibility was in the collector and basin-evaporation subsystems which resulted in a cumulative 67 percentage of the total exergy destruction in Case C. The condenser had 16 percent, with the remaining 17 percent being contributed by piping, storage and control losses. **Table 3.9 identifies the main subsystems responsible for exergy destruction in Case C at the 30 mm basin-water depth**.

Table 3.9. Exergy-destruction distribution for Case C at 30 mm depth

Subsystem	Exergy destruction share (%)
Solar collector	39
Basin evaporation chamber	28
Condensation/glass cover	16
Thermal storage module	9
Piping and hydraulic loop	5
Pump and control unit	3

The better system, however, increased the quantity and quality of useful evaporation, but it was also characterized by comparatively poor exergy efficiency, a characteristic of solar-thermal desalination processes through its high degree of quality degradation between the incoming solar radiation and low-temperature latent-heat recovery.

3.6 Economic analysis

The economic analysis was performed for a simplified deployable prototype, excluding laboratory-only instrumentation. Case C had the highest capital cost because of the additional collector, PV supply, and PCM module, but it also produced the highest annual water yield.

The annualized cost of water was calculated from:

$$LCOW = \frac{C_{cap} \cdot CRF + C_{om}}{V_{annual}}$$

and the simple payback period from:

$$PBP = \frac{C_{cap}}{\left(V_{annual} \cdot P_w\right) - C_{om}}$$

where P_w is the unit market value of potable water. A reference value of 0.10 USD/L was assumed for decentralized delivered freshwater. **Table 3.10 presents the climatic and feedwater conditions recorded during the testing period**.

Table 3.10. Economic performance of the three configurations

Item	Case A	Case B	Case C
Capital cost (USD)	214	332	387
Annual O&M cost (USD/year)	12	16	19
Annual water production (L/year)	996	1363	1706
Cost of produced water (USD/L)	0.047	0.051	0.048
Cost of produced water (USD/m³)	47.0	51.0	48.0
Simple payback period (years)	2.44	2.83	2.61

Although the minimum water cost was obtained for the conventional still, the difference between Case A and Case C was small. In practical terms, the collector–PCM design offered a more attractive compromise because it delivered 71% more annual water than Case A while keeping the unit water cost within the same order of magnitude. From an operational perspective, Case C also provided better late-day availability, which is valuable for decentralized use.

4. Discussion

4.1 Comparison with previous studies

The performance trends obtained in the present study are broadly consistent with the contemporary solar-still literature, particularly with respect to the low baseline productivity of conventional single-slope systems and the positive effect of external heating and thermal storage. In the current results, the conventional configuration produced 2.63–3.38 L/m²·day, which is close to the 2.886 L/m²·day reported by Abdulridha *et al.* for a conventional single-slope still and also agrees well with the 3.30 L/m²·day reported by El-Maghlany *et al.* for a single-slope still tested against hemispherical and conical designs under identical conditions. This similarity suggests that the baseline values assumed in the present paper are technically credible and fall within the range expected for simple basin-type stills under strong solar radiation conditions [12,13].

The case supported in the present work by the collector was also quite favorable compared to the literature. When the water was at the optimal level, Case B increased the daily productivity by 33.7% more than the baseline, which virtually corresponds to the 33.49% improvement by Abdulridha *et al.* when a flat plate collector was combined with a traditional still. The significance of this agreement lies in the fact, that it is evidence that the thermal behaviour of the collector loop and the gain in evaporative potential was not overestimated. In comparison, it was in the collector-plus-PCM system in the present study that the increase was larger, at 5.64 L/m²·day, as the system used two enhancement mechanisms. With daytime feedwater preheating via the flat-plate collector, and delayed heat release via latent thermal storage. The operating principle is in line with a body of reviews, which indicate that solar-still productivity is enhanced as the temperature of basin-water is high, and the temperature of the glass cover is relatively low and the heat loss is minimized by a better insulation and thermal regulation [9,14].

The current findings can also be justified when they are compared with those based on PCM. Murali *et al.* indicated that traditional, PCM and nano-PCM stills produced 2.66, 3.58 and 4.27 unit of daily production, respectively, with the efficiency rising to 46.23% in conventional case and 68.29 in the optimal nano-PCM case. This greater output which is expected in the collector-plus-PCM system is indicative of the fact that the current design is not based on the supposed stimulation of storage, rather the current design incorporates an active source of external heat combined with the storage-aided evening production. Conversely, the enhancement that is forecasted in the present paper is not as high as the 155-183% enhancement of a solar still fitted with thermal storage material as well as a flat-plate collector as indicated by Abdelmaksoud. The given difference is quite plausible, and it can be explained by the variations in geometrical characteristics, absorber properties, local

weather phenomena, basin loading, and the aggressiveness of the improvement plan. Thus, the primary value of current results does not lie in the fact that they result in the greatest output in history but indicates a harmonized and technically feasible betterment with an intermediate complexity and lack of applicability [15,16].

4.2 Practical implications

Pragmatically, the suggested system seems to be quite applicable to the remote, dry, and infrastructure constrained areas, where the solar availability is high and grid reliability can be low. The current views on solar desalination suggest that the sector has been excessively efficiency metrics-oriented and has not given due consideration to deployment routes, dependability, and actual operating conditions. In that respect, the current combined design is appealing, since, it has been integrated to incorporate straightforward thermal desalination and moderate electrical power consumption, passive thermal storage and autonomous control may be established. That fact renders it applicable not only to isolated homes, but also to clustered implementations like farms, rural clinics, field stations, and small desert communities. Distributed desalination studies also highlight the fact that distributed systems are more feasible when they can be remotely monitored, expanded in a modular fashion, and run with reduced maintenance overhead [7,10].

The advantages of sustainability are also significant. It is well known that solar still systems have low operating costs and are environmentally adaptable, particularly when contrasted with the fossil-independent desalination alternatives. The salt rejection levels are high, the auxiliary electricity requirements are very low, and the CO2 savings can be measured in the present study, which is why the suggested design can help to build safer decentralized water supply with the comparably low environmental footprint. That is why its best ability might be to

substitute large centralized desalination facilities and to bridge the gap where the lack of conventional infrastructure is present, not only expensive, but also weak in terms of operations [9,10].

4.3 Limitations and future work

There are a number of limitations on this study. To begin with, the findings rely on a hypothetical experimental model of one Iraqi climatic condition and thus fail to reflect seasonal changes, dust agglomeration, duration of material decay, contamination, or the difficulty of dealing with brine. Second, the feedwater was modeled as the controlled brackish water instead of highly variable real groundwater or seawater which can influence scaling behavior and long-term productivity. Third, the economic evaluation was preliminary and failed to involve supply-chain uncertainty, maintenance logistics and replacement of storage and control parts over the long term. Further research ought to, hence, confirm the system experimentally on a scale of multiple Iraqi locations and seasons, test actual saline feedwaters, refine PCM mass and control logic and analyze modular scale-up with remote monitoring as recent deployment-oriented research studies emphasize reliability, supervision and context-dependent design are of the core concern when increasing the use of solar desalination systems [7,10].

5. Conclusion

This paper described the design, development and performance appraisal of a solar-powered water desalination system that was aimed at decentralized production of fresh water under the Iraqi climatic conditions. The suggested system that incorporated a single-slope solar still, a flat plate solar collector, photovoltaic-driven circulation and phase change thermal storage showed evident technical and

practical benefits over the traditional basin still design. It was found that the basin-water temperature improved significantly with the addition of external solar preheating and thermal energy storage, as well as, the productive operating hours and freshwater yield. The most productive set recorded a daily productivity of 5.64 $L/m^2 \cdot day$ which is significantly higher than the productivity of the base system and still provided the system with a high salt rejection rate and reasonable distillate quality to be consumed.

The analysis also indicated that the performance of the system was highly determined by the depth of basin water and thermal management. Economic evaluation showed that the improved design is capable of generating higher yearly water output at a competitive unit cost of water and an acceptable payback period. Also, the fact that the auxiliary energy required is low and that it will lead to the minimization of carbon emissions is a factor that favors the environmental worth of the proposed system. All in all, the results prove that solar-powered desalination can be considered a promising sustainable solution in remote, dry, and water-stressed areas, and they offer a convenient starting point to future experimental validation, scale-up, and optimization.

6. Acknowledgment

The author gratefully acknowledges the support of the Chemical Engineering Department, Shahid Chamran University, for academic guidance during the preparation of this study.

7. References

[1] UNESCO World Water Assessment Programme (WWAP). (2024). The United Nations World Water Development Report 2024: Water for Prosperity and Peace. UNESCO.

[2] UN-Water. (2024). Progress on the Level of Water Stress: Global Status and Acceleration Needs for SDG Indicator 6.4.2. United Nations.

[3] World Health Organization (WHO) and United Nations Children's Fund (UNICEF). (2025). Progress on Household Drinking Water, Sanitation and Hygiene 2000–2024: Special Focus on Inequalities. WHO and UNICEF Joint Monitoring Programme.

[4] Borgomeo, E. (2024). A desalination research agenda fit for the 21st century. Desalination, 586, 117910.

[5] Orfi, J., *et al.* (2025). Conventional and emerging desalination technologies. Water, 17(2), 279.

[6] Banerjee, S., *et al.* (2025). Towards sustainable water resources: A critical review of traditional and solar-based desalination technologies and innovations. Sustainable Water Resources Management.

[7] Wang, R., and He, W. (2024). Accelerating solar-powered desalination deployment through transferable learning. Communications Materials, 5, 203. https://doi.org/10.1038/s43246-024-00646-6

[8] Almetwally, E. M., *et al.* (2025). Solar powered reverse osmosis desalination: A systematic review of technologies, integration strategies and challenges. Desalination, 615, 119228.

[9] Khalaf, M. O., Özdemir, M. R., and Sultan, H. S. (2025). A comprehensive review of solar still technologies and cost: Innovations in materials, design, and techniques for enhanced water desalination efficiency. Water, 17(10), 1515.

[10] Cohen, Y., *et al.* (2025). Distributed water desalination and purification systems. Current Opinion in Chemical Engineering, 47, 101091.

[11] World Bank Group and Solargis. (2026). Global Solar Atlas: Solar resource data for Baghdad, Iraq. Global Solar Atlas.

[12] Abdulridha, H. O., *et al.* (2024). Performance study for single-slope solar still integrated with different types of solar collectors. Desalination and Water Treatment, 320, 100750.

[13] El-Maghlany, W. M., Attia, M. E. H., Kabeel, A. E., Elazab, M. A., Hidouri, A., and Moharram, N. A. (2025). Comparative experimental study on the performance of single-slope, hemispherical and conical solar stills. Archives of Thermodynamics, 46(3).

[14] Alwan, N. T., *et al.* (2024). Performance of solar still units and enhancement techniques. Heliyon, 10, e40840.

[15] Murali, G., *et al.* (2024). Improved solar still productivity using PCM and nano-PCM composites integrated energy storage. Scientific Reports, 14, 15876.

[16] Abdelmaksoud, W. A. (2025). Enhancing water productivity of solar still using thermal energy storage material and flat plate solar collector. Applied Water Science, 15, 38.

Static Analysis-based Android Malware Detection (2021–2026): A Systematic Survey and Taxonomy / Review Article

Assist. Lact. Ali Hussein Mohammed Ali[1*]

and Assist. Lact. Musaab Riyadh Abdulrazzaq[2]

1,2 Department of Computer Science,
College of Science, Mustansiriyah University, Baghdad/Iraq.

*** Corresponding Author: abo.shbeeb90@gmail.com**

الكشف عن البرامج الضارة في نظام أندرويد باستخدام التحليل الثابت (2021-2026): دراسة منهجية وتصنيف

م. م. علي حسين محمد علي[1*] و م.م. مصعب رياض عبدالرزاق [2]

1, 2 الجامعة المستنصرية، كلية العلوم، قسم علوم الحاسوب، بغداد|العراق.

Abstract

Android malware is not only increasing in size and sophistication but is also a challenge to be detected on a large scale. Static analysis is popular due to its ability to detect malware without running applications or tracing run-time events. However, features, datasets, labeling methods, and assessment methodology differ, which makes studying performance difficult. This is a systematic study of 78 peer reviewed articles written between 2021 and 2026 regarding the Android malware detection by static analysis. In each of the studies, the current survey cover the following feature source, feature representation and engineering, learning paradigms, datasets and labeling methods, and evaluation configuration Permissions, API-based features, opcode-bytecode indicators, manifest-component indicators, metadata-certificate indicators, hybrid static features, adversarial and robustness-oriented designs, graph-based designs, and transformer-based models are represented. Synthesis trade-offs include larger semantic and structural representation, validity risks, that lack of obfuscation resistance is not evaluated, and reporting of repeatability is poor. This study suggests an evidence-based taxonomy and comparative synthesis of static detection methods, structured dataset landscape for decision making, and gap analysis for time sensitive evaluation, leakage prevention, robust benchmarking and repeatable reporting. These findings point to building and experimenting with static detectors for Android malware.

Keywords: Android malware detection, Static analysis permissions, API calls, Opcode, Graph neural networks.

المستخلص

لا يقتصر تزايد حجم وتعقيد البرمجيات الخبيثة لنظام أندرويد على ذلك فحسب، بل يُمثل أيضًا تحديًا كبيرًا لاكتشافها على نطاق واسع. يُعد التحليل الثابت شائعًا لقدرته على اكتشاف البرمجيات الخبيثة دون تشغيل التطبيقات أو تتبع أحداث وقت التشغيل. ومع ذلك، تختلف الميزات ومجموعات البيانات وأساليب التصنيف ومنهجية التقييم، مما يُصعّب دراسة الأداء. هذه دراسة منهجية لـ 78 مقالة مُحكّمة نُشرت بين عامي 2021 و2026 حول اكتشاف البرمجيات الخبيثة لنظام أندرويد باستخدام التحليل الثابت. في كل دراسة، يُغطي هذا المسح الحالي مصادر الميزات، وتمثيلها وهندستها، ونماذج التعلم، ومجموعات البيانات وأساليب التصنيف، وتكوين التقييم. كما يُمثل الأذونات، والميزات القائمة على واجهة برمجة التطبيقات، ومؤشرات رمز العملية/رمز البايت، ومؤشرات البيان/المكون، ومؤشرات البيانات الوصفية/الشهادة، والميزات الثابتة الهجينة، والتصاميم المُوجّهة نحو الخصوم والمتانة، والتصاميم القائمة على الرسوم البيانية، والنماذج القائمة على المحولات. تشمل المفاضلات في عملية التركيب تمثيلاً دلالياً وبنيوياً أوسع، ومخاطر تتعلق بصحة النتائج، وعدم تقييم مقاومة التمويه، وضعف الإبلاغ عن قابلية التكرار. تقترح هذه الدراسة تصنيفاً قائماً على الأدلة وتركيباً مقارناً لأساليب الكشف الثابتة، وهيكلة بيانات منظمة لدعم اتخاذ القرارات، وتحليلاً للفجوات من أجل التقييم الحساس للوقت، ومنع التسريب، ووضع معايير قياس قوية، وإعداد تقارير قابلة للتكرار. تشير هذه النتائج إلى أهمية بناء أدوات كشف ثابتة للبرمجيات الخبيثة لنظام أندرويد وتجربتها.

الكلمات المفتاحية: كشف البرامج الضارة للاندرويد، أذونات التحليل الثابت، مكالمات APK ، رمز التشغيل، الرسم البياني للشبكات العصبية.

1. Introduction

Android is the operating system that has acquired a dominant position in the mobile ecosystem because of its open-source software and wide use on a device [1]. According to numerous industry reports and user statistics, it remains widespread all over the world, with Android being reported as the most popular mobile OS in 2019 [2]. It is also being adopted in Iraq approximately 81% as compared to the global share 87%, according to [3]. The Android OS usage in Iraq is shown in Figure (1). The Android ecosystem is huge, and its openness as well as broad adoption of third-party libraries and SDKs has increased the attack surface and created new security risks [4]. Practically, users often download software provided by various markets and vendors, which exposes them to repackaged or malicious software and early detection of legitimate and malicious software is a crucial feature of securing user privacy and integrity of their devices [5]. Signing and certificate tools at the application layer give valuable security assurances, but do not entirely exist capable of thwarting malware spread, particularly in the face of fast malware development and supply-chain-like channels of distribution [6]. Malware has also developed since the early days of the mobile attacks to massive campaigns with high expansion in Android malware in the last ten years [7].

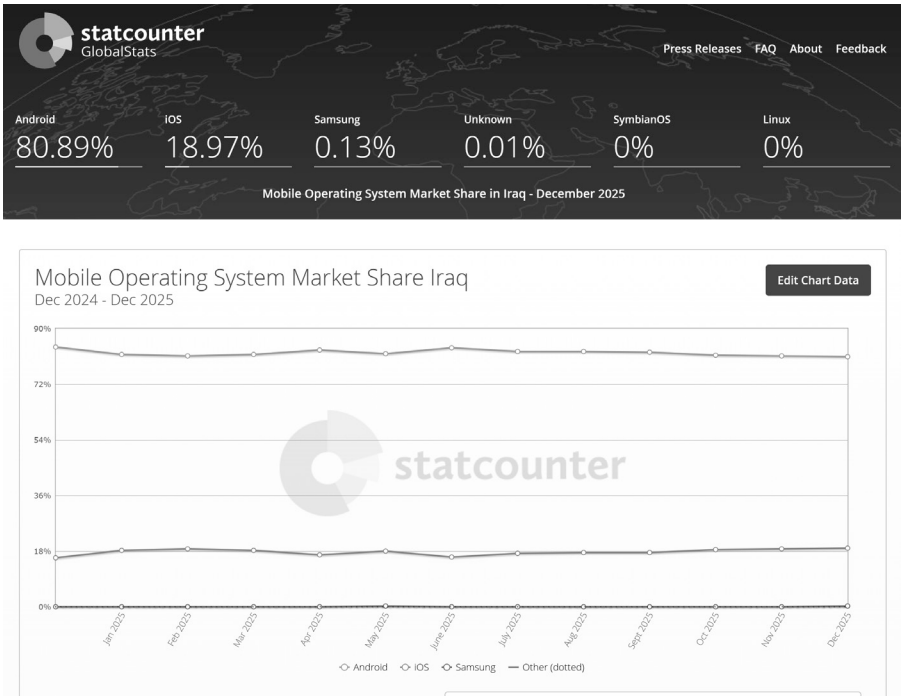


Figure (1): Android Popularity in Iraq

On-device security layers and app stores have broad usage of detection methods to label applications as either benign or malicious and reduce the risks of malware before they are downloaded. Overall, there are malware detection methods that are divided into the categories of static, dynamic, and hybrid malware detection strategies [4]. Static detection is used to analyze application artifacts without running the application, dynamic detection is used to examine the behavior of an application during execution, and hybrid detection is used to bring together the two views [8]. Simultaneously, malware can also be analyzed at the family level, as families are collections of samples with similar behavioral and structural characteristics, family-level knowledge is useful in prioritizing and analyzing samples beyond isolated samples [6]. Android malware detection static analysis is a feasible and popular methodology

due to its ability to conduct the assessment with high speed, even in case they are not run, which is advantageous in the context of large-scale screening and analysis without leakage of privacy information [9]. The current review takes into consideration the use of the static detection of the various application artifacts, such as permissions [9], API usage [10], manifest components, and certificate/metadata cues, which are useful in efficient triage and model training. However, modern evasion techniques like obfuscation, packing/encryption, reflection and repackaging inherently challenge the operation of static detectors in terms of loss of feature fidelity and undermining generalization across time and markets. These properties drive a systematic synthesis of feature selections, representations, and modeling paradigms throughout recent literature on the topic of static analysis. The contributions of this review are as follows:

- It systematically surveys 78 peer-reviewed studies (2021–2026) on Android malware detection using static analysis only, and consolidates their feature sources, representations, models, datasets, and evaluation protocols into a unified evidence base.
- It proposes an evidence-driven taxonomy that organizes static detection research across permissions, APIs, opcodes/bytecode, manifest/components, metadata/certificates, hybrid static features, robustness-oriented designs, graph-based methods, and transformer-based models.
- It provides a comparative synthesis that highlights cross-study trends and trade-offs (accuracy vs. scalability, semantic richness vs. engineering cost, robustness vs. obfuscation).
- It identifies methodological gaps and threats to validity (dataset aging, label noise, inconsistent evaluation setups, and limited reproducibility reporting) and derives a forward-looking research agenda for more reliable static detection.

Some recent surveys have surveyed Android malware detection, although most related literature either (i) is biased to the broad machine-learning view of

mixed feature families and at times mixed sources of knowledges (static and dynamic), (ii) is concerned with adjacent tasks, such as malware family classification, identification of third-party libraries, or testing of Android applications, or (iii) is out of date with the high rate of posteri-facto evolution of the static-analysis approach in recent years. By contrast, this paper gives a recent, static analysis-only synthesis of the 2021-2025 period and classifies the literature using an evidence-based taxonomy making modern directions of statical analysis first-class categories (graph/structural representations, robustness/adversarial considerations, and transformer-based static models). In addition to categorization, this study groups the dataset landscape into a decision-oriented reference and provides a critical examination of the threat to the fair comparable cross study, label reliability (variation in antivirus consensus), evaluation protocol variations, and report voids, to make the readers of the reported accuracy claims more realistic. Table I summarizes these differences and explains the added value of this survey as compared to similar reviews.

Table I: Summary of compared surveys

Study		This study	[2]	[4]	[5]	[6]	[7]	[8]	[1]
Checkpoint	**Year**	**2025**	**2020**	**2020**	**2024**	**2020**	**2024**	**2021**	**2025**
Android-focused		✓	✓	✓	✓	✓	✓	✓	✗
Android malware detection focus		✓	✓	✓	✓	✓	✗	✗	✓
Static-analysis-only scope		✓	✗	✗	✗	✗	✗	✗	✗
Coverage window: 2021–2025		✓	✗	✗	✗	✗	✗	✗	✓
Explicit SLR workflow		✓	✗	✗	✓	✗	✓	✗	✓
Fine-grained static taxonomy		✓	✗	✗	✗	✗	✗	✗	✗
Hybrid static feature fusion category		✓	✗	✗	✗	✗	✗	✗	✗
Graph/structural static category		✓	✗	✗	✗	✗	✗	✗	✗
Adversarial/robustness category		✓	✗	✗	✗	✗	✗	✗	✗
Transformer-based static category		✓	✗	✗	✗	✗	✗	✗	✗
Decision-grade dataset landscape		✓	✗	✗	✓	✗	✓	✗	✗
Comparability/label reliability critique		✓	✗	✗	✗	✗	✗	✗	✗

To support navigation, the remaining parts of this paper are structured as follows: Section 2 explains the systematic review protocol. The methodology-driven taxonomy and comparative synthesis of static detection categories is presented in section 3. Section 4 is an analysis of the datasets used in reviewed studies. Section 5 focuses on cross cutting gaps and threats to validity. Finally, section 6 summarizes evidence-driven recommendations and a research roadmap for future research.

2. Systematic Protocol of Literature

This research is conducted after the systematic literature review (SLR) protocol with a screening workflow to find the studies on Android malware detection using static analysis. Five major digital libraries were searched because of their broad coverage of computer science, engineering, and security research literature: ScienceDirect (Elsevier), IEEE Xplore, Web of Science (WOS), SpringerLink, and Scopus.

2.1 Search Strategy

The following query was used with minor syntax adaptations to match each database's advanced search format:

("Android") AND ("Malware" OR "Malicious Software") AND ("Static Analysis" OR "Static" OR "Static analysis") AND ("Classification" OR "Detection"). The search was conducted in November 2025 across the selected databases and restricted to English-language articles published between 2021 and 2025. The review is specifically narrowed down to peer-reviewed journal articles to achieve methodological rigor, full reporting of the experiment, and archival stability of the evidence of 2021-2025. Journal research is generally more comprehensive in its descriptions of datasets, labeling methods, assessment processes and details of

reproducibility, so it can be more dependable on cross-study synthesis and lessen noise due to early or informal reporting. Consequently, the results of this SLR are expected to be used by researchers and practitioners as a high-confidence source of information about deployment-relevant findings of well-established, proven methods of static-analysis. The retrieved records from all databases were consolidated and processed through consecutive screening and eligibility steps, as summarized in Figure (2).

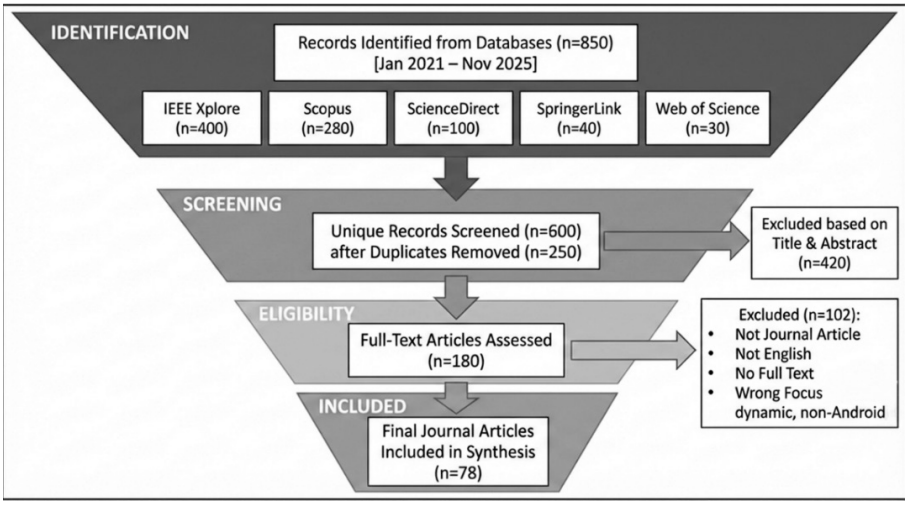


Figure (2): Research query results

2.2 Study Selection Process

A total of 850 records were identified across the five databases (ScienceDirect: 100, Scopus: 280, IEEE Xplore: 400, WOS: 30, SpringerLink: 40). After removing duplicates, 600 unique records remained. Title and abstract screening reduced the set to 420 potentially relevant studies. Full-text assessment and the application of inclusion/exclusion criteria yielded 180 eligible studies, from which 78 articles were retained as the final set for synthesis.

2.3 Inclusion Criteria

Studies were included if they met all the following conditions:

1. Published in a peer-reviewed journal.
2. Focused on Android malware detection using static analysis techniques (permissions, API calls, opcodes/bytecode, manifest/components, certificates/metadata, graphs, or hybrid static features).
3. Available as full text online.
4. Written in English.
5. Published within 2021–2025 to reflect contemporary detection pipelines, datasets, and ML/DL architectures.

2.4 Exclusion Criteria

Studies were excluded if any of the following applied:

1. Not written in English.
2. Not accessible as full text through the selected databases.
3. Focused on malware for other operating systems like (Windows, Linux, MAC) rather than Android.
4. Primarily relied on dynamic analysis or runtime traces.

2.5 Research Questions

This survey aims to answer the following research questions within static-analysis-based Android malware detection:

- **RQ1**: What types of static detection techniques are used across the selected studies?
- **RQ2**: What are the strengths and limitations associated with each static technique?

- **RQ3**: What algorithms and learning models are employed (ML/DL/ hybrid), and how are features represented?
- **RQ4**: What datasets are used (public/private/ benign/malware sources), and what are the sample sizes and labeling strategies?
- **RQ5**: What are the key motivations, contributions, and reported performance outcomes of each study?

2.6 Quality Assessment (QA) Criteria

The QA scores were used to interpret findings and to contextualize unusually high-performance claims, rather than excluding studies solely based on score. To ensure that the synthesis reflects methodologically reliable evidence, this review applied a lightweight quality assessment rubric. Each included study was scored across the following criteria using a 3-point scale (0 = not addressed, 1 = partially addressed, 2 = clearly addressed):

- **QA1**: Clear problem definition and threat model (malware type, detection objective).
- **QA2**: Static-analysis pipeline clarity (feature source, extraction steps, representation).
- **QA3**: Dataset transparency (source, size, class balance, de-duplication).
- **QA4**: Labeling reliability (label source, VirusTotal/AV thresholds if used, scan date if reported).
- **QA5**: Evaluation rigor (split strategy, validation protocol, leakage controls, time-aware evaluation if applicable).
- **QA6**: Reproducibility signals (tool availability, parameter reporting, code/dataset accessibility).
- **QA7**: Robustness considerations (obfuscation, reflection/dynamic loading discussion, limitations).

2.7 Data Extraction and Synthesis Plan

For each included study, this review extracted a standardized set of fields to support evidence-driven comparison and synthesis:

- **Methodology**: feature source(s) (permissions, APIs, opcodes/bytecode, manifest/components, metadata/certificates, graphs), representation (vector/sequence/graph/embedding), and pipeline/tools (static analysis framework, decompilation tools if stated).
- **Modeling**: learning paradigm (ML/DL/GNN/Transformer), classifier architecture, feature selection/engineering steps.
- **Datasets**: dataset names, sample sizes, class balance, benign/malware sources, family vs binary labels, duplication, and timeframe when reported.
- **Evaluation**: split strategy (random/time-based/family-disjoint), metrics (accuracy, precision, recall, F1, AUC), and baselines.
- **Limitations and validity**: obfuscation handling, labeling uncertainty, drift concerns, leakage risks, and reproducibility constraints.

3. Methodology and Taxonomy Analysis

This systematic review synthesizes 78 peer-reviewed studies published between 2021 and 2025 that address Android malware detection using static analysis only without executing applications or collecting runtime traces. To support evidence-driven synthesis, each study was coded using a consistent extraction form covering (i) the primary static signal source, (ii) the adopted representation or feature engineering, (iii) the modeling paradigm, (iv) the dataset and labeling strategy, (v) the evaluation protocol, and (vi) the reported limitations

The taxonomy categories are not necessarily mutually exclusive, since in practice, many studies use a combination of multiple immobile signals such as permissions and API calls. Each study is coded with one primary category depending on the most dominant or most distinguishing signal and representation, and other ones are listed as secondary tags to prevent the scenario of counting the same study under different categories. Thus, the total counts of categories and the per-category evidence tables are based on primary-category assignment only, and hybrid multi-signal studies are recorded once under the Hybrid category and are not repeated under their respective categories. Accordingly, the total category counts and the per-category evidence tables will indicate the primary-category assignment only, and hybrid multi-signal studies will be counted once under the Hybrid category and will not be counted again under all their respective categories. As summarized in Figure (3), the reviewed literature is first organized by the dominant signal source into: (a) manifest- and metadata-level signals (N=18), (b) code-level signals (N=23), and (c) hybrid and advanced paradigms (N=37) that integrate multiple feature sources and emphasize modern modeling and evaluation dimensions. Within each group, the subsequent subsections provide evidence tables mapping individual studies to their representations, model families, enabling explicit comparisons of design trade-offs such as interpretability versus feature complexity, accuracy versus scalability, and robustness under obfuscation. Across the surveyed studies, recurring limitations include sensitivity to code obfuscation and reflection/dynamic loading, reliance on non-time-aware benchmarks, and incomplete reporting that limits reproducibility.

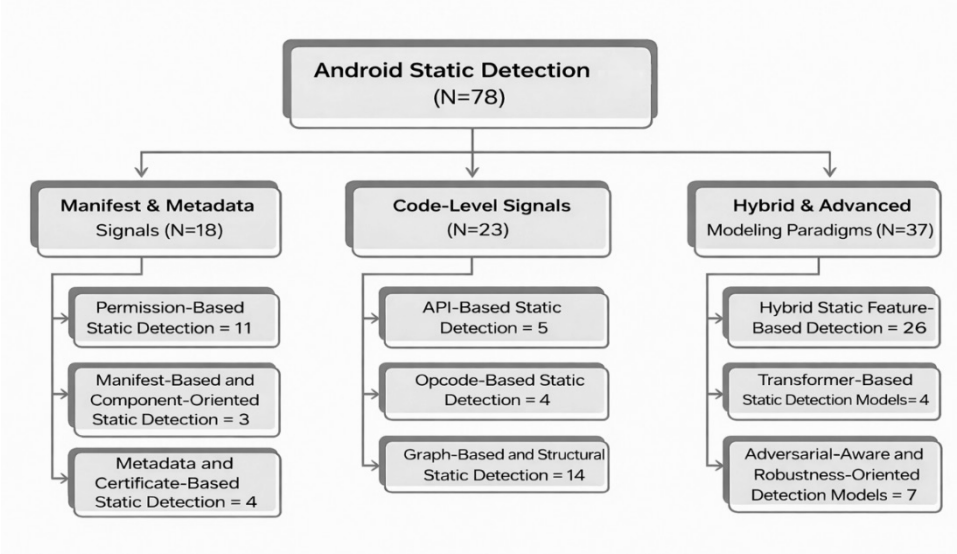


Figure (3): Taxonomy Analysis

3.1 Permission-Based Static Detection

Permission-based static detection is a core component of Android security research, which is concerned with the analysis of application requested privileges in the AndroidManifest.xml file or their use in the code. Several studies have been done on how to optimize standard machine learning classifiers by good feature selection. A model proposed in [9] focuses on avoiding reliance on dynamic execution and extracting permission features from the manifest which can achieve 91.1% accuracy using a combination of feature selection and Bayesian classification. Validating the validity of the standard classifiers, the authors in [11] used a balanced data set to show that Random Forest and SVM classifiers could achieve roughly 93.75% accuracy using binary permission vectors, although the authors acknowledged differences in terms of obfuscation. To overcome the issue of data freshness and feature redundancy, the framework proposed in [12] used feature

importance guided attribute reduction to reduce the feature space from 323 to 48 key permissions so that a Random Forest classifier achieved 97.5% accuracy. Moving past the question of whether permissions exist or not, researchers have explored the structural relationships between permissions. The approach in [13] proposed the Composition Ratio (CR) where the ratio of permission pair occurrences to the total number of pairs is calculated to form the feature vector that can be robust to permission inflation and noise. These structural approaches show that permission co-usage modeling gives a better discriminative power than viewing permissions as independent binary features [13].

Recognizing the fact that the legitimacy of permission request often depends on the function of an application, recent works have integrated category awareness in detection models. A study utilizing Self-Organizing Map (SOM) and K-means clustering mined the permission usage patterns specific to the app categories and used these patterns to identify malware that requests permissions incoherent with the declared category [14]. Similarly, the authors in [15] have considered a permission sensitivity-based detector that assigns weights to permissions according to their overload or lost status into a particular category and achieves 92.17% accuracy by differentiating permissions to benign and malicious apps in context. Furthermore, the framework in [16] uses Collaborative Filtering to determine unnecessary permissions based on the category of an application, and they neutralize them by injecting code to return synthetic data to prevent privacy leakage while keeping the application from crashing. In addition, the study [17] was focused on Real Permissions which were extracted using reverse engineering and the FP-Growth algorithm to find correlated pairs of permissions invoked at runtime, which were used to train a Neural Network. Special architecture has been created to overcome efficiency limitations and complex threat situations presented in [18].

LinRegDroid has shown that the permission-based detection could be successfully addressed by simple and efficient multiple linear regression models with performance like complex classifiers such as SVM and Decision Trees, but computationally feasible for on-device deployment [18]. Addressing distributed attacks, the SigColDroid was built to identify colluding applications, and it uses dangerous permissions in conjunction with permission rates and Smali file size, to successfully identify colluding malware from lone apps with a high ROC_AUC of 99.48% [19]. Finally, authors in [20] make a baseline comparison of ML models on permission features using 331 permission indicators with several Machine Learning algorithms. In order to get final thought of this approach, Table II represents contributions of each article related to this approach

Table II: Permission-Based Article Contributions

Study ID	Year	Permission signal	Representation / feature engineering	Classifier / decision model
[9]	2022	Manifest permissions	Permission subset selection (feature selection)	Bayesian classification
[13]	2021	Manifest permissions	Permission-pair Composition Ratio (CR) vector (ratio-based)	Not specified
[17]	2023	Real permissions used in code	Reverse engineering → real permissions, FP-Growth to mine correlated permission pairs, semantic couples	Neural Network
[14]	2022	Manifest permissions + category context	103-dim binary + SOM + K-means to mine category-specific permission patterns, pattern-fit features	SVM
[18]	2022	Manifest permissions	Binary vectors, multiple linear regression as risk score, 2 rule-based classifiers + bagging + hybrid ensemble	Regression-based + ensembles (incl. SVM/DT in ensemble-2)

Study ID	Year	Permission signal	Representation / feature engineering	Classifier / decision model
[15]	2021	Manifest permissions + category context	Category inference from description (NB/TextCNN), Permission Sensitivity (PS) weighting (overload/lost permissions)	RF, SVM, k-NN
[16]	2022	Manifest permissions	Collaborative Filtering + Frequent Permission Set Mining to find unnecessary permissions, instrumentation to return fake data	Not a pure classifier (permission minimization + runtime hooking)
[11]	2025	Manifest permissions	Binary permission vectors	RF, SVM, DT
[12]	2024	Manifest permissions	323-permission inventory → feature-importance reduction to 48 permissions	RF, KNN, NB, DT
[19]	2025	Dangerous permissions + static intensity signals	Dangerous permissions + Smali file size + permission rates	RF, ExtraTrees, AdaBoost, XGBoost, LightGBM, ANN, DNN (multi-class)
[20]	2024	Permissions	331 permission indicators, GridSearchCV + 5-fold CV	RF/SVM/XGB/DT

Permission-based detectors often appear effective in laboratory evaluations because manifest permissions are easy to extract and frequently align with malware labels in popular benchmarks; however, this alignment does not necessarily constitute a reliable behavioral signature. Most studies in this evidence base rely on manifest-only permission vectors (7/11), while smaller subsets incorporate category context (2/11), code-derived usage-aware permissions (1/11), or longer intensity-style signals (1/11). This suggests that many reported improvements depend on a proxy that is sensitive to dataset choice and may drift as Android and the app ecosystem evolve. Moreover, the Android permission model has evolved substantially; since Android 6.0, dangerous permissions are granted at runtime,

which weakens the assumption that manifest-declared permissions always translate into consistent capabilities across versions and years. Accordingly, time-sensitive evaluation is increasingly important, as time-based splits better reflect concept drift and platform transitions than conventional random splits. Random splits can also inflate apparent generalization, especially when near-duplicate apps or minor variants leak across training and test folds. This category is not strictly isolated: permission features frequently co-occur with manifest and component indicators and sometimes with code and API cues, which complicates attributing performance gains to permission alone. Overall, the evidence points to a clear trade-off between interpretability and behavioral specificity: manifest-only permission vectors are intuitive and inexpensive, yet they remain coarse proxies that can flag legitimate high-privilege apps and miss stealthier malware that requests fewer permissions. The literature's inconsistencies are largely explained by differences in representation binary vs co-usage vs context-aware weighting and evaluation, with random splits often yielding optimistic estimates compared to time-aware splits that better reflect drift and platform changes.

3.2 API-Based Static Detection

API-based static detection is centered on analyzing the Application Programming Interface (API) calls that are made by an application to determine malicious behavior. While initial methods have focused on naive frequency counting, recent studies have made progress in sequence modeling, semantic clustering, and hybrid deep learning models to understand the contextual use of APIs. Several studies have gone beyond considering APIs as an unordered group of words but instead use deep learning to model the temporal and sequential nature of code execution. One of them uses API call sequences as a time series sequence data

using a Skip-gram model to learn dense numeric embeddings and a Bidirectional Long Short-Term Memory (Bi-LSTM) network to capture long-range dependencies and dependency patterns in the forward and backward directions [10]. Expanding the feature scope Another study suggested a multi-input deep learning model that combines API calls with permissions, opcodes, and fuzzy hashes into a multi-branch Keras Functional API network, achieving around 99.5% accuracy by learning across various static feature types all at once [21]. To overcome the semantic gap and the evolution of the Android platform, researchers have devised methods that pay attention to the meaning and association of the API calls, and not only to their raw occurrence. The SDAC framework solves this problem of model aging and concept drift by embedding APIs in a vector space and grouping them in semantic clusters using k-means, which enables the system to map new or renamed APIs to existing clusters without redesigning the feature space [22]. Recent study also emphasizes optimizing classifier performance and ensuring robustness against common evasion techniques like obfuscation. In addition, to combat API obfuscation and APK size bias, a hybrid framework utilizing DeGuard was proposed in [23] to reveal concealed API calls. It combines these deobfuscated features with certificate metadata and interaction-based ratios to eliminate size bias, significantly outperforming baselines on the Drebin dataset. Finaly, to reduce false alarms and improve precision, an automated framework was introduced in [24]. It utilizes metaheuristic algorithms such as Binary Coati Optimization to select discriminative API features and tune the hyperparameters of a stacking ensemble classifier and achieving over 98% accuracy. In order to get final thought of this approach, Table III represents contributions of each article related to this approach.

Table III: Article contributions in API-based

Study ID	Year	API-based static signal	Representation / feature engineering	Model family	Other static signals used	Task (label space)
[21]	2022	API calls (among multi-signals)	Multi-branch feature inputs fused before softmax (Keras Functional API)	Multi-input deep learning	Permissions, services, receivers, opcodes, file size, fuzzy hashes (ssdeep)	Binary + 5-class (Benign/SMS/ Banking/Adware/ Riskware)
[24]	2023	API calls + permissions	Feature subset selection and/or ensemble tuning via metaheuristic (Binary Coati Optimization)	Optimized ensemble (stacking/voting)	Permissions	Malware detection
[23]	2021	API calls (original + deobfuscated)	DeGuard to recover ProGuard-obfuscated APIs, interaction-ratio features to reduce APK size bias	Supervised ML (LightGBM/ CatBoost/RF/ET/ Linear SVM)	Permissions, components (services/activities/intents), certificate features, Apktool log metadata	Malware detection
[10]	2024	API call sequences (path-based)	Call-graph + DFS path extraction → ordered API sequences → n-grams → Skip-gram (Word2Vec) embeddings	Deep sequence model (Bi-LSTM)	N/A	Malware detection
[22]	2022	API semantics under evolution	API path extraction + Skip-gram embeddings → k-means semantic API clusters, unseen APIs mapped to clusters (FEO) or clusters+models updated (FMU)	Ensemble/voting over classical classifiers	N/A	Malware detection under concept drift

In API-based research, the key trade-off is between semantic richness and fragility of extraction: those representations, in which one can maintain order and meaning, are able to enhance specificity of behavior, though at the price of higher computational cost. The literature in this evidence set (n = 5) narrows down to 3/5 multi-signal hybrids, 1/5 sequence-aware modeling, 1/5 semantic/drift-aware abstraction with each having a different trade-off between fidelity and practice. Resilience Hybrid pipelines can provide good laboratory accuracy by combining APIs with opposing static signals but aggregated or set-based API encodings are likely to lose behavioral context and be confounded by third-party libraries and benign applications with large functional footprint. Sequence-sensitive techniques maintain execution-based order and have the ability to represent higher-order dependencies, but are limited by call-path recovery, reflection/dynamic loading, and incomplete static resolution. Semantic drift-aware abstraction alleviates sparsity and API history by relocating new/renamed APIs to consistent clusters, whereas it can obscure malicious intent at a fine-grain when clusters become too large or because labeling evaluation environments change. The discrepant findings between studies are mostly due to differences in library filtering and deobfuscation assumptions, and evaluation conditions which extend to random splits rather than environments which are more akin to drift and deployment constraints. Thus, to have fair synthesis, there must be transparent reporting of the selected API representation, library deobfuscation treatment, and leakage-resistant evaluation, to ensure that the gains are due to the method, and not to artifacts of data sets.

3.3 Opcode-Based Static Detection

Opcode and bytecode analysis beyond metadata or API calls by taking a closer look at the low-level instructions that are run by Android Runtime (Dalvik or ART). Research in this domain has spread into a variety of contexts such as deep sequence

modeling, new signal transformations, and formal semantic verification. In order to solve the shortcomings of the traditional sequence analysis method, researchers have used advanced deep learning architectures to process opcode and bytecode stream. A prominent trend in literature in recent years is the conversion of bytecode into other forms of data in order to exploit algorithms from other areas. A few studies have used bytecode as image data. Similarly, Dex bytecode into RGB image matrices and utilizing self-supervised learning (DINO) with a teacher-student distillation process, which achieves great generalization capability for discovering new families with a limited amount of labeled data [25]. Exploring signal processing, another framework restates malware analysis as an audio classification problem by transforming raw APK bytes into .wav file and extracting engineering features such as MFCCs and spectral descriptors [26]. Moving away from platform-specific parsing, the DAEMON classifier utilizes raw byte N-grams to create a platform-independent detection mechanism [27]. By employing an entropy-based filter to remove noisy data and an Aho–Corasick algorithm for fast mapping, DAEMON demonstrates that a unified, byte-level feature extraction pipeline can achieve accuracy (98.74%) on Drebin Dataset and provide explainability without parsing specific Android file structures like classes.dex [27]. In contrast to statistical machine learning approaches, some research relies on formal verification techniques. A formal equivalence checking framework was proposed that translate Java bytecode instructions into CCS processes and labeled transition systems performing malware analysis by applying weak bisimulation equivalence to compare automata states [28]. This method identifies malware families based on opcode-level behavioral semantics rather than signature matching, making it inherently robust to code obfuscation and syntactic changes without requiring labeled training data [28]. In order to get final thought of this approach, Table IV represents contributions of each article related to this approach.

Table IV: Article contributions in opcode-based

Study ID	Year	Opcode/ byte-level signal	Representation / feature engineering	Model family	Platform scope	Task
[26]	2023	Raw APK bytes (reframed)	APK bytes → .wav audio, extract MFCCs + spectral features, FS via CFS/ReliefF/IG/ Gain Ratio	Classical ML (RF/ SVM/KNN/ J48/ NB)	Android	Malware family classification
[27]	2021	Raw byte N-grams (4/8/16/32)	Entropy filter + multi-stage feature mining, pairwise separation + tags, fast mapping via Aho–Corasick	Random Forest	Cross-platform (Windows PE + Android APK)	Malware classification (and some family-separation logic via tags)
[25]	2025	classes.dex bytecode → image-like signal	Bytecode hex stream → RGB image matrix, self-supervised DINO pretraining, fine-tune with rotation loss + Manifold Mixup, prototypical (metric) few-shot with multi-crop aggregation	Few-shot + self-supervised representation learning	Android	Malware family recognition under few-shot
[28]	2021	Java bytecode semantics (opcode-level behavior)	Bytecode → CCS processes → labeled transition systems, weak bisimulation for equivalence, heuristic pruning reduces comparisons	Formal verification (no supervised ML)	Android	Malware detection + family classification via equivalence

In studies based on opcodes-bytecodes, a major trade-off has been one of robustness versus interpretability: low-level instructional signals are less sensitive to superficial code change but give less insight into the semantic explanation of why an app is malicious. In this evidence set (n = 4), there are four paradigms of representation: raw byte n-grams (1/4), signal reframing (1/4), bytecode-to-image self-supervised representation learning (1/4) and opcode-level formal semantics through equivalence reasoning (1/4). Three of the four papers shift the focus on reframing, self-supervised representation learning, or semantic equivalence reasoning, directly in an attempt to be less sensitive to superficial code changes and common semantic obfuscations. But working at the level of the byte/opcode level can lose interpretability and it can obscure benign complexity or third-party library space with malicious intent resulting in errors that are more difficult to diagnose. Furthermore, cross-study-comparison is limited by the inability to match tasks like binary vs family vs few-shot, that is, near-ceiling results can indicate an easier or a different task, not a universal high-quality representation. Thus, comparability needs to have equal task set-ups and record the labeling and assessment set up, as well as the particular transformation and representation pipeline that determines robustness in each study.

3.4 Manifest-Based and Component-Oriented Static Detection

Manifest-based and component-oriented detection strategies do analyze the structural backbone of the applications on Android. By looking at declared components and interactions between them, these methods attempt to detect the presence of malicious intent by detecting anomalies in the configuration, modeling the capability, and tracking inter-app communication patterns. The framework presented in [29] uses an extended π-calculus to model Android components as

processes and their interactions as actions and defines precise operational semantics for the evolution of components. By establishing ideas of behavioural equivalence between an application's actual behaviour and templates of known malicious behaviour, this approach offers a solid decision-making framework for identifying complex threats such as collusion, even without machine learning [29]. Instead of using two classes (benign vs. malicious), recent research focuses on what an application can do. These focuses improve dealing with the high dimensionality of manifest attributes, and optimization techniques are often used. BFEDroid [30] is dedicated to feature selection across the manifest-level attributes and ranks the permissions and the intent filters to eliminate noise. By training traditional classifiers on these reduced subsets, the framework shows that a smaller optimized feature space can maintain and improve the detection accuracy while greatly decreasing the model complexity and training time [30]. Furthermore, detecting collusion between applications requires analyzing how components communicate across app boundaries and integrating diverse static features often yields the most robust results. SEDMDroid in [31] proposes a stacking ensemble detector that fuses permissions, sensitive APIs, manifest events, and taint-style data flows into a unified framework. By training Multi-Layer Perceptron base learners on these heterogeneous features and combining them with an SVM meta-classifier, the system achieves approximately 94.9% accuracy when data-flow features are included, consistently outperforming individual classifiers and single-view models [31]. In order to get final thought of this approach, Table V represents contributions of each article related to this approach.

Table V: Article contributions in Manifest-based

Study ID	Year	Component / manifest-centric signal	Representation / analysis method	Model family	Task
[29]	2022	Android components + ICC semantics (Activities/Services/ Receivers/Providers, intents/data)	Extended π-calculus with operational semantics, behavioral equivalence (strong/weak simulation, bisimulation) against benign/malicious templates, decision rules (credible/malicious/ suspicious)	Formal methods (no ML)	Template-based behavioral classification, collusion reasoning
[30]	2022	Manifest permissions + intent filters	Feature ranking/selection over permissions+intents (BFEDroid) to reduce noisy/redundant manifest attributes	SVM, DT, NB, RF	Malware detection
[31]	2021	Manifest events + permissions + sensitive APIs + permission-rate + taint-style source–sink flows	Decompile + manifest parsing, FlowDroid for flows, high-dimensional vectors, stacking ensemble: multiple MLP base learners + SVM meta-classifier, diversity via subspaces/bootstraps/ PCA rotations	Stacking ensemble (MLP + SVM fusion)	Malware detection

In the paradigm of manifest component-oriented studies, the major trade-off is that between structural expressiveness and scalability static fragility: more ICC- or semantics-conscious modeling can be done, but it requires a correct analysis of the result and results in more analysis cost. Two primary design lines can be identified in the evidence table (n = 3): ICC-semantic formal modeling (1/3) and ML over manifest-centric features (2/3), informally consisting of feature ranking selection and, in certain instances, multi-signal fusion. Formal ICC-semantic methods are capable of reasoning about interaction patterns including collusion but are susceptible to non-observable or hidden communication patterns and may not

perform well when behavioral construction is done dynamically. Manifest-centric ML pipelines can be more readily scaled by ranking and filtering high-dimensional attributes like permissions and intent filters but their performance tends to be confounded by the complexity of apps and overlapping with secondary signals in the case of fused permissions, APIs-flows. The reported performance is different since manifest indicators are dependent on configuration granularity and static visibility exported components and permissive intent-filters are informative, but dynamic loading and hidden flows may result in missed detection, and unreliable results. Thus, comparability relies on the articulation of the modeled component/ICC scope and the scope of the modeled static assumptions, recording the feature selection fusion decisions, and matching the evaluation protocols with the proposed threat model collusion targeted.

3.5 Metadata and Certificate-Based Static Detection

Metadata and certificate-based detection strategies remove the analytical focus from code execution and instead put the focus on the structural and descriptive attributes of an application. By analyzing the digital signatures, the marketplace information, and the package identifiers, or by converting these elements into different modalities, researchers have devised efficient ways of identifying malicious families and variants. A growing body of research is looking into translating non-code components into visual representations for analysis. Beyond the internal file structures, there is external data from the ecosystem that acts as a rich source of detection signals. One study used metadata of apps from the store and proposed a new labeling criterion, application lifespan defined as Potentially Harmful Apps (PHAs), which have been removed from the market within a timeframe of one month [32]. By using Random Forest classifier trained on 601 features extracted

from market metadata, the accuracy of the model was 90% which proved that the survival time in the marketplace is a less biased signal to identify the malicious campaign than traditional antivirus labels [32]. Taking a different approach to the external data, MalVulDroid uses Natural Language Processing (NLP) to analyze reports of textual description from security vendors instead of the APKs themselves [33]. By mapping these natural language descriptions to the vulnerability categories using Multilayer Perceptron, the framework succeeds in mapping the software vulnerabilities exploited by certain malware families to enrich static analysis with external threat intelligence [33].

In order to model complex relationships between a variety of static attributes and advanced neural architectures have been applied to metadata. The Trandroid system uses Transformer-based architecture for modeling the interactions between heterogeneous features, such as permissions, components, metadata attributes, and family labels [34]. Tested on the benchmark dataset TUANDROMD, the accuracy of the Transformer model reached 99.25%, which proved that the self-attention mechanism is more powerful than the traditional CNNs or RNNs in identifying the hidden long-range dependencies in hybrid static feature datasets [34]. Finally, the efficiency rapid signature distribution of the detection are critical for the large-scale detection. The FLSH framework replaces the bulky machine learning models with similarity hashing by using ssdeep to generate context-triggered piecewise hashes of important static artifacts [35]. This way it becomes possible to calculate a Net Similarity Index to detect variants with an accuracy of 96.67% against even minor obfuscations to cryptographic hashes [35]. In order to get final thought of this approach, Table VI represents contributions of each article related to this approach.

Table VI: Article contributions in metadata and certificate-based

Study ID	Year	Primary metadata / certificate signal	Representation / feature engineering	Detection / model core	Task
[33]	2022	External security-vendor text reports (metadata about malware)	NLP: BoW + n-grams (1–4) + TF-IDF, mapping families ↔ 9 vulnerability categories	MLP (best), SVM, RIDOR, PART	Vulnerability-category prediction (capability/ vulnerability analytics)
[32]	2021	App-store metadata + lifespan labeling (removed <1 month vs survives >6 months + 0 VT detections)	601 normalized static features from manifest + market metadata (permissions, hardware, category, etc.)	RF (best), SVM, SGD, XGBoost	PHA detection for app-market screening
[34]	2025	Mixed manifest + metadata attributes within a specialized benchmark (TUANDROMD)	Reverse engineering + feature encoding into numeric vectors, transformer models interactions among heterogeneous static features	Transformer (vs CNN/LSTM/GRU/ RNN/CNN-LSTM baselines)	Malware detection (and family/variant context in benchmark)
[35]	2025	Fuzzy hashing (ssdeep) over unpacked static artifacts (manifest/dex/ structural sections)	Generate ssdeep hashes, compute Net Similarity Index vs malware fingerprint repository, manual refinement to reduce noisy sections	Similarity-hash indexing (lightweight, not heavy ML)	Binary detection + variant/family identification

According to literature, metadata certificate-based investigating the trade-off between ecosystem sensitivity and generalization is the most common option: such indicators may be operationally helpful, but their predictive ability will change with marketplace policies, the extent of regional exposure, and temporal decadence. In this body of evidence (n = 4), four design directions can be identified that are, market-level lifespan proxies, vendor-text NLP, transformer modeling over heterogeneous features such as metadata, and fuzzy-hash similarity to identify variants, with each study using one design direction (1/4). Since these instructions

are aimed at different goals triage vs semantic intelligence vs family-variant matching vs detection, the concept of accuracy cannot be interpreted unanimously and may result into contradictory impressions when compared as one detection measure. In addition, features facing the ecosystem can easily be manipulated by presentations and policy-driven changes, and certificate or marketplace identities can be rotated or repackaged to minimize transferability over time and space. Thus, the ecosystem context (market/timeframe/region), the task objective, and the labeling assumption should be clearly reported by cross-study synthesis instead of comparing the results through the accuracy.

3.6 Hybrid Static Feature-Based Detection

Hybrid static detection is the amalgamation of static indicators as diverse as permissions, API calls, opcodes, control flow graphs, and metadata into unified frameworks. By synthesizing these heterogeneous sources of data, researchers hope to overcome the limitations of single-view analysis, such as susceptibility to obfuscation and lack of semantic context. Researchers have built intricate neural architectures to represent intricate interactions in hybrid static features. HDNFDroid [36] uses Stacked Denoising Autoencoders to learn features from the sparse data and then fine tunes with a supervised learning algorithm. Moving towards efficiency, BL-AMD in [37] uses Broad Learning Systems to obtain high accuracy without heavy computation. However, study [38] use Residual LSTMs and Cascade LSTM-GRU networks to capture long term dependences in feature sequence. Furthermore, methods such as study [39] support that combining different features like opcodes, strings, shared libraries, and APIs in Multi-Layer Perceptron and Deep Neural Networks is consistently better than single feature baselines. Beyond ordinary deep learning, there are a number of studies that deal with the interaction of the features.

The framework in [40] adds a hierarchical detection approach that combines block-level, module-level and app-level features, by which it finds that authorization-sensitive features which are very discriminative. Similarly, addressing the sparsity of static data, factorization machines are used to explicitly model pairwise feature interactions [41], authors use FP-Growth to mine frequent co-existence patterns of permissions and APIs, proving that feature combinations are more powerful than isolated attributes. In addition, transforming code into visual or graph structures enables the application of computer vision and graph algorithms. LensDroid in [42] combines three views API graphs using GCN, Opcode sequences using TextCNN, and structural artifacts using ConvNeXt with multi-modal bilinear pooling. However, MalFlows in [43] models apps as Heterogeneous Information Networks (HIN) using flow2vec embedding for capturing semantic flow. In addition, visual approaches include DCEL in [44], which combines 1D vectors with 2D images in order to parallelize DNN and CNN processing while study [45] uses Local Binary Patterns (LBP) in order to code images. In addition, [46] introduces a pipeline that transforms BERT-embedded feature tokens into images for CNN classification and SARVOTAM. In order to handle high dimensionality, nature-inspired optimization is widely used. The Enhanced Parrot Optimizer in [47] and Artificial Bee Colony (ABC) in [48] are used to select the optimal subsets of features hence improving the classifier performance significantly. Furthermore, for ransomware specifically a Binary Particle Swarm Optimization (BPSO) scheme in [49] combined with SMote oversampling resolves the issue of class imbalance well.

Robustness can often be accomplished through ensembles. ANDFRF in [50] proposes a fuzzy rank-based ensemble for managing classifier uncertainty, whereas JDroid in [51] is a stacked ensemble for Dalvik opcodes and Java bytecode. Scaling this to massive datasets, in [52], AdaBoost is used over 56,000 static attributes

extracted from half a million apps and with high stability. Going back to the issue of lack of data and labeling, the Siamese one-shot learning framework in [53] identifies malware families with very few samples. MADS in [54] uses light-weight rules and type-specific Random Forests to prevent installations in real-time. Moreover, PAIRED in [55] works on explainability using SHAP with a light Random Forest and study [56] shows that a shallow Random Forest with selected features is possible for ultra-fast detection (<100ms) suitable for mobile devices. Hybrid methods are also specific to a particular situation. The article [57] identifies financial fraud based on features in the domain such as package naming patterns. Also, the research [58] focuses on ransomware by using a combination of URL lexical analysis and API indicators. Finally, critical studies such as [59] put the spotlight on the importance of reproducibility and demonstrate that tree-based models are often more than simple neural networks can perform when the confounders from the experiment are controlled. Pushing boundaries even further, [60] suggests a multimodal framework that integrates static artifacts with dynamic runtime traces, making a new link between static and dynamic analysis. Transfer learning is also investigated in [61] to enhance the generalization to zero-day variants with the help of large scale offline features. In order to get final thought of this approach, Table VII which represents contributions of each article related to this approach.

Table VII: Article contributions in Hybrid-based

Study ID	Year	Hybrid static signals used	Core representation / idea	Model / learner	Task
[36]	2022	APIs + permissions + dataflow	Deep feature learning for sparse/high-dim static (MSAE+SDAE, unsupervised + supervised fine-tune)	Hybrid deep (autoencoders + classifier)	Detection
[40]	2021	Basic blocks/bytecode + permissions + authorization-sensitive APIs	Hierarchical fusion (block → module → app)	Hierarchical classifier	Detection
[37]	2021	Permissions + intent actions + sensitive APIs + etc.	Broad Learning (low compute, static-only)	Broad Learning–based detector	Detection
[54]	2023	Manifest + metadata + limited static features (no decompiling)	Rule-based prefilter → type-specific RF detectors (multi-stage)	Rules + Random Forest (per type)	Type-specific detection
[41]	2023	Permissions + API calls (+ frequent-API by frequency)	FP-Growth top-k co-existence itemsets (size 2–5) → binary combo features	RF/J48/SVM/ KNN/LR	Detection
[49]	2021	Permissions + API calls (ransomware focus)	SMOTE + wrapper BPSO feature selection	RF/SVM/DT/ KNN	Ransomware detection (imbalanced)
[44]	2024	Permissions + APIs + intents + command/ signature features	Chi-square selection → dual encoding: 1D vectors + 2D images → vote fusion	DNN + CNN + majority vote	Detection
[53]	2023	Permissions + APIs + strings	Siamese CNN learns similarity, N-way one-shot	Siamese CNN (one-shot)	Detection + family classification
[52]	2022	7 groups (perms, components, method tags, intents, packages, APIs, services/receivers)	56k binary attributes, AdaBoost over classical ML	SVM/DT/ KNN/NB/RF + AdaBoost	Detection
[55]	2022	Drebin-215 hybrid features	RFE reduces 215→35, RF + SHAP explainability	Random Forest + SHAP	Detection
[39]	2025	Permissions + intents + API calls	Hybrid DNN, removes leakage-prone metadata, cross-dataset eval	DNN (binary + family layer)	Detection + family/ category

Study ID	Year	Hybrid static signals used	Core representation / idea	Model / learner	Task
[45]	2025	Manifest + Dex bytes (image-based)	Byte streams→grayscale, LBP features, CNN ensemble (VGG/ResNet/DenseNet) + focal loss	CNN ensemble	Detection
[57]	2025	71 hybrid static features incl. package-name stats + source + user activity features	Daily retraining simulation (92 daily datasets), boosted trees	LightGBM best (plus XGB/ CatBoost/RF/ LR)	Financial malware detection
[50]	2025	Hybrid static vectors (perm + API indicators + others)	Fuzzy rank-based fusion of probabilities	KNN/SVM/LR/ XGB/LGBM + fuzzy fusion	Detection
[38]	2025	Permissions + API-related features	RF feature importance + Cascade LSTM–GRU, EOA hyperparameter optimization	Cascade LSTM/GRU + EOA	Detection
[47]	2025	214 perms + 27 API calls (241 attrs)	PO/PSO/GWO/POPSO + SA refinement for feature selection	DT/GB/HGB/ RF/XGB	Detection
[48]	2025	Drebin-derived 216 static features	ABC feature selection + RF/ XGB	RF, XGBoost	Detection
[58]	2025	URL lexical + API-call indicators (ransomware)	Huge URL feature set + stacked ensemble (RF/ LGBM → MLP meta)	Stacked ensemble	Ransomware detection
[51]	2025	SMALI opcodes + Java bytecode (hybrid opcode)	ExtraTree FS to 461 features, stacking ensemble	Stacking ensemble	Detection
[43]	2025	Control-flow + dataflow + ICC (no execution)	HIN + meta-paths, flow2vec embeddings, channel-attention DNN + MLP	Attention DNN + MLP	Detection
[56]	2025	Permissions + sensitive APIs + components/ intents	Multi-stage FS (IG + Chi² + RFE) → ~10–20 feats, shallow RF emphasized	Shallow RF (+ LR/NB/DT)	On-device / low-resource detection
[61]	2025	Broad static mix (manifest + metadata + APIs + opcodes)	46,648→10,523 selected, transfer learning, multi-class families, releases SAMDroid	DNN/CNN + transfer	Multi-class family identification
[59]	2024	Drebin-style hybrid vectors + APIGraph clustered APIs	Rigorous protocol: dedup vs duplicated, equal tuning budget, offline vs monthly active learning	RF/SVM/XGB/ MLP/SCC/HCC	Detection under reproducible settings

Study ID	Year	Hybrid static signals used	Core representation / idea	Model / learner	Task
[60]	2024	Static + dynamic (multimodal)	Train per-view models and fuse outputs	Fused supervised models	Detection
[46]	2024	Permissions + intents + receivers + services → BERT embeddings → image	Tokens→BERT mean pooling→grayscale images, CNN, Optuna tuning	Custom CNN (+ backbones compared)	Detection
[42]	2025	API callgraph + opcode grams + artifact bytes (dex/xml/so)	3-view embeddings (GCN + TextCNN + ConvNeXt) → MFB pooling + self-attention fusion	Multi-view deep fusion	Detection

The hybrid models themselves are not superior; only in cases when the fused modalities are complementary, the fusion strategy fits the threat model and is evaluated in leakage-resistant protocol, the gains are obtained. The most common backbone in the hybrid evidence table is APIs (19/26), then there are permissions (17/26), and manifest component exposure signals are also represented in a smaller, but significant proportion (10/26). The trend signifies that the majority of hybrids purposely integrate announced capability of permissions with carried out conduct of APIs and, by some counts, exposure entry points constituents. There are three repeating fusion recipes that are observable in the literature. First, the most common recipe is to combine APIs and permissions (16/26) with the ability to use capability-behavior complementarity to reduce over-declared-permission false positive and promote discrimination. Second, the APIs permissions recipe can be commonly scaled up with manifest component exposure (6/26) to include entry points and attack surface and capability and behavior. Third, it can be effective with broader feature-bank hybrids (3/26) where modalities are still complementary, though more susceptible to redundancy and dataset-specific artifact and so demand a well-thought-out fusion design and validation controls. A different sub-trend

provides extra low-level views or structure (4/26) to represent higher-order dependencies, and better deal with sparsity, at the expense of more complex extraction, and sensitivity to incomplete static evidence. On the other hand, the failure mode of hybridization occurs when fusion only raises dimensionality and provides near ceiling results that are not applicable across datasets, time or protocols.

3.7 Adversarial-Aware and Robustness-Oriented Detection Models

As machine learning becomes a standard for malware detection research has shifted toward understanding and countering adversarial attacks. This is the area of vulnerability analysis, feature hardening, adversarial training, and novel architectural defenses. Several studies have revealed the fragility of current ML detectors. To quantify this threat, researchers created Attack Effectiveness Classifiers (AEC), which predict the probability of success of specific evasion techniques [62]. Experiments demonstrated that problem-based evasion attacks that exploit obfuscation and flow manipulation could lower detection rates in commercial scanners by up to 97% [62]. Furthermore, the OFEI framework proposed the semi-black-box attack via Simulated Annealing that achieved the misclassification rate of 98.25% against Deep Learning as a service (DLaaS) platforms using modification of only one feature per iteration [63]. Talking about the obfuscation of code directly, the tool Deoptfuscator is concerned with bringing apps obfuscated by DexGuard back to a High-Level Intermediate Representation [64]. It was able to recover obfuscated apps with an average similarity of 0.93 to the original apps and thus allows standard static detectors to work correctly [64]. One of the big trends in the field of defense is to include adversarial examples during the training process. Theoretical guarantees were added by the PAD framework, which involves Input

Convex Neural Networks (ICNN) and Stepwise Mixture of Attacks to solve robust training as a constrained optimization problem [65]. PAD was able to maintain more than 83.45% accuracy under attack which beats heuristic defenses [65].

Going just beyond robust classification, some articles try to detect the attack attempts themselves. MalProtect is a stateful defense layer where it keeps the history of up to 10,000 queries to detect abnormal patterns in indication of a probing attack [66]. It reduces evasion rates by more than 80% for major benchmarks [66]. Complementing this, the defense mechanism that comes with OFEI uses Bayesian Neural Networks to estimate aleatoric and epistemic uncertainty [63]. This approach identifies adversarial inputs based on the clusters in the feature space, for which the uncertainty is high. 99.28% detection rate [63]. Finally, complex representations of data are being capitalized on in order to improve resilience. A graph-based framework based on Graph Neural Networks (GNNs) on API graphs used a Variational Graph Autoencoder (VGAE) and GAN to obtain adversarial graphs for training [67]. This adversarial-aware GNN showed an accuracy of 98.33% on CICMalDroid [67]. Another method presented a dual-modal detector based on a spatial (ResNet-50) and frequency domain feature of the DEX code [68]. Using random erasing as a training technique to simulate the obfuscation process, this model achieved 97.3% accuracy on Drebin and proved to have a higher robustness against FGSM attacks than single-modal baselines [68]. In order to get final thought of this approach, Table VIII represents contributions of each article related to this approach.

Table VIII: Article contributions in Adversarial Malware Detection Approach

Study ID	Year	Role	Target detector / feature space	Attack / defense mechanism	Model family
[62]	2022	Attack + Meta-modeling	Replicated SVM/FM/ DNN detectors trained on Drebin, benign from AndroZoo, tested also vs VT engines	Problem-based code-modifying evasions (camouflage/ obfuscation/noise/flow manipulation), AEC models predict attack success per detector	Attack suite + effective predictors
[64]	2022	Robustness enabler (DE obfuscation)	DexGuard level 3 obfuscated apps, improves downstream static analysis	Deoptfuscator: DEX→HIR, remove opaque predicates/ global opaque vars, uses ReDex for L2, rebuild executable APK	Static DE obfuscation tool
[67]	2023	Attack + Defense (adversarial-aware graph)	API graphs + permissions + intents, GNN embeddings + classifiers	Build local/global API graphs, train GCN/GraphSAGE, VGAE + GAN generate adversarial graphs (VGAE-MalGAN), adversarial training improves	GNNs + GAN/ VGAE + downstream classifiers
[66]	2023	Defense layer (query attack detection)	Feature-based detectors (Android+Windows) on AndroZoo/ Drebin/SLEIPNIR	MalProtect front-end: query history (10k), indicators (L0 distance, shared feats, count stats, AE loss) → LR/NN to detect attacks, blocks/overrides output	LR / small NN (guard)
[63]	2024	Attack + Defense	DLaaS deep detectors on static binary vectors	OFEI: one-feature-per-iteration, simulated annealing guided by output scores, Defense: Bayesian NN uncertainty (aleatoric+epistemic) detects adversarial inputs	Attack + Bayesian defense
[65]	2024	Defense (principled adversarial training)	DNN on rich Drebin-like hybrid static features	PAD: ICNN adversary detector + constrained robust optimization, PAD-SMA mixture of attacks (FGSM/BGA/BCA/Grosse/JSMA/ Mimicry) during training	DNN + ICNN
[68]	2025	Defense-oriented robust detector (augmentation + fusion)	DEX bytes → spatial fingerprint + frequency spectrum	CNN (ResNet-50) + spectral features (2D FFT) → DFN with attention + recursive optimization, random erasing to mimic occlusion, tested under FGSM	CNN + fusion

The existence of robust claims in adversarial-conscious malware detection needs a clear threat model; otherwise, apparent accuracy cannot be generalized to attackers as to their objectives, information, and capabilities. In this regard, the research ought to specify the objective of the attacker, the knowledge regime such as white and black-boxes, the capabilities, and limitations, in line with the common adversarial-ML practice. The strength of robustness is explicitly viewed as resistance to attacks in most studies in this evidence set (6/7), whereas more explicit transformations of obfuscation are mentioned in a smaller fraction (2/7). Assumptions on the knowledge of the attackers also differ explicit black-box and query-driven environments are found in (2/7), explicit white-box gradient testing in (2/7), and the rest are based on mixed or implicit regimes. Defenses have been found in adversarial training, attack mixing, guard layers indicating probing by query pattern or uncertainty, code-level recovery (deobfuscation) to stabilize downstream static signals, and robustness-by-design. The often-discussed mismatched threat model and mismatched evaluation budget can also be understood as resulting in conflicting high-robustness claims: What was earlier read as robustness like clean accuracy,and low level of computation might be lower when under adaptive query-efficient attack.

3.8 Graph-Based and Structural Static Detection

Graph-based detectors model applications using explicit program graphs such as control-flow graphs (CFGs), call graphs, or inter-component communication (ICC) graphs. Their main bottlenecks are graph extraction cost, scalability on large corpora, and graph fragility under obfuscation and reflection, which can lead to missing edges or nodes and degraded generalization. To overcome the limitations of bag-of-words static analysis, which often fails to capture execution logic and

obfuscated relationships, researchers have increasingly turned to structural representations. These approaches treat Android applications as complex graphs (Control Flow Graphs - CFGs, Data Flow Graphs - DFGs, Function Call Graphs - FCGs) or transform code into visual and mathematical structures suitable for advanced deep learning. A major trend involves leveraging Graph Neural Networks (GNNs) to learn directly from code structure. The AdMat framework in [69] transforms API-call graphs into adjacency matrices, effectively treating them as images for CNN processing, achieving 98.26% accuracy by bridging graph semantics with computer vision. Similarly, HAWK in [70] model's applications as Heterogeneous Information Networks (HINs) connecting APIs, permissions, and libraries. It utilizes a Meta-Structure Guided Graph Attention Network (MSGAT) to learn robust embeddings, capable of detecting out-of-sample apps in milliseconds. At the application level, another framework in [71] builds a global similarity graph of apps and applies GraphSAGE with an MLSTM aggregator, significantly improving robustness against adversarial attacks like FGSM. Several studies abstract code into topological structures to analyze behavior without heavy semantic parsing. SNADroid in [72] models function call graphs as social networks, ranking structural metrics like Katz Centrality and Triangles to identify 15 highly discriminative features. Transforming bytecode into images allows the application of powerful pre-trained vision models. The study [73] converts DEX bytecode into color and grayscale images, fine-tuning models like Xception and ResNet to achieve 99.40% accuracy. SAC in [74] advances this by fusing content and structure. It treats DEX content as RGB images and bytecode instructions as graphs, outperforming both N-gram and standard graph baselines. For ransomware specifically, the framework in [75] combines ResNet-18 visual features with fuzzy hashing (ssdeep), ensuring robustness against minor code obfuscations. Another approach in [76] employs Stacked Autoencoders (SAE) to

compress high-dimensional binary images of APKs before CNN classification, enhancing computational efficiency. Deep structural analysis addresses specific blind spots in static detection. LibRoad in [77] and LibCapsule in [78] focus on Third-Party Libraries (TPLs). LibRoad uses structural signatures to accurately detect libraries, while LibCapsule enforces isolation policies on untrusted library code. Addressing native code risks, uDep in [79] uses mutation-based analysis to generate dependency stubs for JNI functions, significantly improving taint analysis precision. Finally, DBank [80] constructs a Triadic Suspicion Graph specifically for Banking Trojans, using PageRank-like scoring to detect fraud with 99.9% accuracy. Other novel approaches include MMPD in [81] which analyzes PDF structure labels for mobile robot security, and the framework in [82] which treats memory dumps as static big-data artifacts, processing them with Apache Spark for forensic detection. In order to get final thought of this approach, Table IX represents contributions of each article related to this approach.

Table IX: Studies summary in Graph-Based Approach

Study ID	Year	Structural / graph signal	Representation / feature engineering	Model / analysis core	Task
[69]	2021	API-call graph (nodes=APIs, edges=relations)	Graph → 219×219 adjacency matrix image, API space reduced (~23k→~17k)	CNN	Detection + family classification
[72]	2024	Function call graph as social network	57 graph metrics, rank via T-test/NMI/MIC → ~15 top features	RF/DT/1NN/3NN	Detection
[73]	2022	Bytecode as images (classes.dex → PNG)	Color/grayscale images, transfer learning across 16 CNNs	Fine-tuned CNNs (ImageNet TL)	Detection
[75]	2023	Bytecode images + fuzzy hash similarity	DEX→RGB + ResNet-18 embeddings, ssdeep similarity scores, fuse	RF/SVM/KNN/LR	Ransomware detection
[76]	2023	Raw APK → binary image/vector	SAE compresses binary images → latent, CNN on compressed map	SAE + CNN	Detection

Study ID	Year	Structural / graph signal	Representation / feature engineering	Model / analysis core	Task
[80]	2021	Triadic Suspicion Graph (apps–APIs)	PageRank-like suspicion scores/ ranks, fused with lightweight static+dynamic	Graph scoring + ML	ABT detection (multi-class: goodware/ ABT/ other)
[70]	2024	Heterogeneous Information Network (App–API–Perm–Class–Interface)	Meta-path/meta-graph + MSGAT, incremental embedding MSGAT	Graph Attention + RF/LR/SVM	Detection
[78]	2022	Structural confinement of third-party libraries	Static rewriting + instrumentation (incl. dyn dex + native isolation) enforcing XML policies	Enforcement framework	Privacy/ security enforcement
[77]	2022	Structural signatures for TPL identification	Weighted signatures over classes/packages, name+similarity matching, obfuscation-aware filters	Library detection tool	TPL detection
[82]	2022	Memory dump structure (RAM artifacts)	Forensic features (processes/ DLLs/registry traces) in Spark	LR/GBT/RF/LSTM	Detection
[81]	2022	PDF internal object structure	Structural tokens (/JS etc.) + TF-IDF-like weighting + FS	SVM/DT/KNN/ DNN	Malicious PDF detection
[79]	2023	JNI-aware dependency graph for taint analysis	Mutation-based dependency inference on native funcs, generate Jimple stubs for DroidSafe	Static taint analysis enhancement	Precision/ coverage of taint analysis
[71]	2025	App–app similarity graph (homogeneous)	Features (perms/APIs/TPLs) create edges, GraphSAGE with MLSTM aggregator, chi-square FS	GraphSAGE (MLSTM)	Detection
[74]	2025	Dual: DEX content image + bytecode structure graphs	Content: CLAHE RGB + ConvNeXt-style IFNeXt, Structure: Word2Vec + dual graphs (GI, GII) + GAT/GCN, fuse → MLP	CNN + GAT/GCN fusion	Detection

In both graph-based and structural static detection, the most important trade-off is between the fidelity of graphs and the interpretability of the higher-order relations they can represent more detailed program-structure graphs can represent more fidelity, but the resulting representations and the decisions learned are more difficult to interpret and audit compared to compact representations. The literature in the evidence table (n = 14) is not assembled around one graph recipe, but instead it is divided into an even (7/14) of explicit graph construction and structure proxies that are not based on whole-graph reasoning (7/14). Recurrent families are also known among explicit graphs, e.g. Call-structure graphs like API-call or function-call relation, relational graphs over heterogeneous entities like HINs and app-app similarity graphs, suspicion graphs, and dependency taint-oriented variants including JNI-aware dependency graphs. The proxy structural encodings graph images and bytecode images are better scalable and pipeline-friendly and are compromised in detailed control dataflow meaning. The extraction assumptions and noise are the two principal factors contributing to the appearance of conflicting results: reflection and dynamic loading may eliminate or corrupt edges, nodes, and third-party libraries might introduce third parties to the graph that drown the malware-specific patterns.

3.9 Transformer-Based Static Detection Models

In transformer-based static detection, tokens typically correspond to API call sequences, opcode-bytecode tokens, or graph-derived walks substructures linearized into sequences. For fair evidence-based comparison, transformer results should be discussed against strong CNN/RNN baselines under matched splits and dataset conditions, because gains may otherwise reflect data label effects rather than architecture advantages. Current research into static malware detection is, for

the most part, veering away from using conventional and recurrent neural architectures and moving toward models that incorporate self-attention mechanisms and generative artificial intelligence. Such a transition supports the rendering of long-range dependencies in the source codes, and allows the study of binary semantics in the same way as that of natural language processing. The SHERLOCK framework in [83] is a fit example of such a trend. By mapping DEX bytecode into visual representations and using Masked Auto-Encoder (MAE) in a self-supervised learning regime over 1.2 million unlabeled Android applications, SHERLOCK detects the structural grammar of malware without using labelled data. Subsequent supervised fine tuning results in detection accuracy of 97 offering resilience to code obfuscation and shows the model's advantages of scalability compared to traditional convolutional networks. Transitioning from image based to text-based representation, SecretLoc in [84] is the first large language model (LLM) application for static code audit which uses GPT-4o-mini. Unlike conventional detection of hard coded secrets using regular expressions, SecretLoc uses the contextual reasoning capability of the model to interrogate the surrounding code. Consequently, the system recovered ~93% of previously documented secrets, and found some additional, obfuscated credential formats that evaded the texture of discussed tools, nominally asserting the utility of generative AI in static security assessment. In the field of feature-based malware detection, IPR – EODL in [85] extends the classical long - short - term memory (LSTM) network using a channel attention module. The attention mechanism places higher weights based on discriminative features based on how this relates to security outcomes, while the LSTM component captures the temporal dynamics of those features. Parameter optimization is performed through the Equilibrium Optimizer (EO) in which after optimization the obtained system has an accuracy of 99.18, thus proving the effectiveness of attention in improving sequential models for static feature vectors.To complement these

neural approaches are sophisticated signal processing techniques. SpecView in [86] uses Singular Spectrum Transformation (SST) to analyze binary executables as 1D time series signals to find change points that would imply packing or encryption. This spectral signature method was 100% for highly obfuscated families where image-based methods failed. By preserving spatial locality, this approach makes it possible for lightweight CNNs such as LeNet to detect the presence of ransomware in native binaries with 99.7% accuracy and high energy efficiency approximate 1 Joule per scan, which is ideal for mobile deployment. In order to get final thought of this approach, Table X represents contributions of each article related to this approach.

Table X: Studies in Transformer Approach

Study ID	Year	Transformer-based	Static signal	Representation / key idea	Model core	Task
[85]	2024	ViT	DEX bytecode → byteplot images	Self-supervised MAE pretraining with ViT (75% masking) on MalNet ~1.2M apps, fine-tune for downstream tasks	ViT/MAE + small head	Binary detection + type (47) + family (696)
[84]	2025	LLM	Hardcoded secrets (strings + code context)	Decompile XML + code strings, candidate mining → LLM labeling (secret type vs NOT-SECRET) for context-aware auditing	LLM-assisted static auditor	Secret discovery (security auditing)
[83]	2022	Attention-LSTM	Permissions + API calls	Binary vectors, Channel-Attention LSTM (CA-LSTM) learns feature weights + dependencies, Equilibrium Optimizer tunes hyperparams	CA-LSTM + EO	Malware detection
[86]	2021	Signal+ML	Raw binary stream as 1D signal (PE + APK)	8-bit time-series → Singular Spectrum Transformation CP scores, PSO tunes SST parameters, ML ensemble on CP vectors	RF/ GNB/ET + Voting	Detection + packed/ encrypted robustness

In transformer-oriented detecting static transformation, the main trade-off across transformer-non transformer is not tokenization modality decision, but comparability vs. semantic continuity represented: it depends on what limited list counts as a token to determine the quality of cross-study comparison. In such an evidence set (n = 4), tokenization is modality-dependent and has four directions, including vision-based tokenizers on top of bytecode images (1/4), LLM-based auditing on text (1/4), enhancing sequence models with attention on sparsely indicative signals (1/4), and segmenting raw binaries with signals (1/4). Since the directions are also aimed at fulfilling various goals and naming spaces, such as auditing vs detection, and binary vs multi-class family, the values of headline accuracy are not directly comparable, even in a case where there are self-attention models. Other conflicting assertions are usually motivated by unsurpassable bases and information conditions: increase could be dependent on pretraining weight or on the quality of labels or label splits instead of attention itself.

4. Datasets Used in Static Malware Detection

It is imperative to possess a comprehensive understanding of the numerous datasets that are employed in Android malware detection in order to assess the generalizability, reproducibility, and real-world applicability of detection algorithms. This is due to the fact that these datasets are employed to detect malicious software.

4.1 Public Benchmark Datasets

The researchers employed a diverse array of datasets in the 78 papers that were analyzed, six popular datasets mentioned repeatedly. These datasets are as the following:

1- **Drebin**: Drebin is an Android malware benchmark used as a classical example to publish reproducible research on the topic of detecting malware when it is in static form. The dataset available on the official Drebin site is 5,560 malicious Android applications (APKs) representing 179 malware families, which were collected in the time period between August 2010 and October 2012, and samples were provided to the authors by the MobileSandbox project, so the access is limited and credentials are provided after requesting them by email to prevent abuse [23], [26],[27].

2- **AMD (Android Malware Dataset**): AMD (Android Malware Dataset, the Argus Lab) is an Android malware-only corpus, built to provide ground-truth family/behavior semantics beyond what simple AV labels provide: the technical report describing the creation of the corpus describes how 24,650 malware APKs were grouped into 71 malware families, and the families further clustered into 135 behaviorally distinct subgroups behavioral profiles reports of which are provided in the report, samples were gathered on VirusShare, Google Play, and third-party [18][26],[33].

3- **Malware Genome Project (MalGenome**): The Android Malware Genome Project dataset, also known as the Malware Genome Project (MalGenome) or Android Malware Genome Project dataset, is one of the early malware-only corpus that was extensively reused in subsequent studies: a manually/automatically crawled collection of 1,260 malicious APKs in 49 malware families, purported to represent the majority of known Android malware between August 2010 and October 2011, gathered by Zhou and Jiang to use in their 2012 SoK study: MalGenome practically is most frequently employed in malware-family analysis/

classification and baseline detection projects, however, researchers typically must combine it with a distinct set of benign apps (MalGenome itself is malware-only) and it is very skewed in its distribution of families. [41][44][55]

4- **CICMalDroid 2020**: The dataset is a hybrid of a static dataset and a dynamic dataset, which is created and trained to classify malware categories: CICMalDroid2020 (Canadian Institute for Cybersecurity, UNB) consists of 17,341 Android apps (APKs) gathered using a variety of sources (VirusTotal, Contagio, AMD, MalDozer and other sources mentioned in the article by the authors) between 2017 and 2018 and is categorized into 5 classes (Adware, Banking malware, SMS malware, The developers have dynamically tested the APKs with CopperDroid and state that 13,077 apps were successfully launched; after eliminating problems like failures in JSON parsing. The number of apps that can be used is 11,598 apps (Adware 1,253; Banking 2,100; SMS 3,904; Riskware 2,546; Benign 1,795). It contains (1) APK files (17341), (2) logs of dynamic analysis captured (13077), and (3) various CSV views of features such as 470 dynamic features like system calls, binders, and composite behaviors. The dataset has 139 system-call-only features, a 50621-feature static set like intents, permissions, services, sensitive APIs. Also, as well as PCAP network-traffic captures of the actual runs; the dataset page also presents the important papers [55],[46],[67].

5- **AndroZoo**: AndroZoo is an academic, ever-growing collection of Android APKs, so far maintained by the University of Luxembourg. The official portal reports 26,694,440 APKs in the collection, collected by several sources such as Google Play and other markets. APKs are served raw

(APK/ZIP with DEX byte code, certificates, assets and the manifest), and each application is scanned by numerous antivirus detectors via Virus Total-style scanning, with the data set served via authenticated download API Hashed with SHA-256. An APK list hash index containing core metadata fields commonly used in malware studies is also published periodically by AndroZoo: such as dex date for compilation time, VT detection for number of engines flagging, APK size, DEX size, package name, version code, and source market. So, researchers can filter and select subsets before downloading. To access, one is required to request an API key via an institutional email and accept terms, including non-commercial use, no redistribution without permission and having the key remain secret; keys have a time limit of 6 months and are limited to 500,000 APK downloads at a time. Also, according to AndroZoo, Google Play Metadata been introduced in December 2023 and the auxiliary information is now more diverse using fields like hash/VT in addition. [66][68].

6- **CIC-MalMem-2022**: Published by the Canadian Institute of Cybersecurity (UNB). It consists of a memory-dump focused dataset that is designed to test obfuscated malware detection by memory forensics to make it closer to real-world conditions by creating dumps, but not the artifacts of the dumping process itself (it uses debug mode, such that the dumping process itself does not appear in the captured memory). It has a total of 58,596 records (29,298 benign and 29,298 malicious) and three malware super-classes, including Ransomware, Trojan horse, and Spyware. [82]. Table XI shows most popular datasets used in current literature.

Table XI: Decision-grade benchmark datasets used in static Android malware detection (2021–2025)

Dataset	Access	Timeframe / Refresh	Benign source	Malware source	Approx. size	Label type	Duplicates handling (reported)	Split strategies
Drebin	Public download	Aug 2010 – Oct 2012	Not included (paired with markets in many studies)	Malware from MobileSandbox ecosystem	5,560 apps, 179 families	Family labels	Not stated in surveyed papers	k-fold CV, random hold-out
AMD (Android Malware Dataset)	Public benchmark	2010 – 2016	Usually paired with public benign	Curated malware corpus	24,553 malwares, 71 families	Family / binary	Not stated	Random split, k-fold CV, some cross-dataset
Malware Genome Project (MalGenome)	Academic corpus	Aug 2010 – Oct 2011	Not included	Curated malware families	1,260 malwares, 49 families	Family	Not standardized across re-uses	Random split / CV in derivatives
CICMalDroid 2020	Public	Dec 2017 – Dec 2018	Included	Included (multiple categories)	17,341 samples	Category labels (Adware/ Banking/SMS / Riskware/ Benign)	Not consistently documented in surveyed papers	Random split, CV, some multi-class
AndroZoo	Request access by emailing from an institutional address; APK download via API key + SHA-256; key expires after 6 months and capped at 500,000	Continuously growing; crawling started in late 2011 and continues; the public index is updated nightly (before ~6am Luxembourg time).	Primarily app markets such as Google Play and multiple alternative markets	Malware is not a separate feed malicious apps are a subset of the collected sources and are identified via AV/VirusTotal signals (vt_detection)	Current number of APKs shown on the portal: 26,694,440	VirusTotal-based: vt_detection = number of AV engines flagging the APK, plus	duplicate handling is not guaranteed without downloading both, and repackaged apps will have different hashes.	Not prescribed by the dataset

Dataset	Access	Timeframe / Refresh	Benign source	Malware source	Approx. size	Label type	Duplicates handling (reported)	Split strategies
CIC-MalMem-2022	Public download from CIC/UNB dataset page	Released as a fixed snapshot; no rolling refresh is described on the official	Normal user behaviour memory dumps captured while running various applications (benign activity)	2,916 malware samples collected from VirusTotal	58,596 records total (29,298 benign + 29,298 malicious) with 57 memory-analysis attributes (tabular feature dataset derived from dumps)	Balanced binary (benign vs malicious) and malware type labels (Spyware, Ransomware, Trojan Horse)	No explicit duplicate policy reported	Not prescribed by the dataset

Public datasets function as empirical common ground in Android malware research because they are openly accessible, repeatedly reused, and therefore enable reproducible comparison across detection pipelines. In the surveyed literature, two resource types dominate reuse, curated malware benchmarks (Drebin, MalGenome, AMD) and large-scale APK repositories (AndroZoo) that are commonly used to construct benign sets and to scale mixed benchmarks toward more realistic app diversity. In parallel, several studies move beyond binary malware and benign discrimination by adopting multi-class category benchmarks to evaluate fine-grained threat typing [21], and by incorporating more recent public collections (CICMalDroid2020) for category-based evaluation and cross-benchmark testing [40], [83]. The benchmark landscape also includes derived feature-vector datasets (Drebin-215) and specialized public resources used for targeted objectives such as abnormal data-flow mining [31], memory-dump–based analysis under obfuscation (CIC-MalMem-2022) [82], and malware-dump collections for sourcing malicious samples [81]. Crucially, cross-study comparability can be undermined by temporal drift (older corpora vs newer app ecosystems), label noise and AV-policy differences including VirusTotal-style reports, and evaluation protocol choices (random split vs time-aware split, potential data leakage through repackaged variants or duplicate apps). These factors should be considered when interpreting unusually high accuracy claims across heterogeneous datasets and splits.

4.2 Custom /Private Datasets

Custom and private datasets are repeatedly introduced to address practical limitations of standard benchmarks, most notably recency, label confidence, domain specificity, and controlled class balance often by re-sampling public sources, re-annotating ground truth, or engineering a representation aligned with a paper's threat model. A prominent direction emphasizes scale to improve statistical power and capture long-tail behavior, [52] aggregates 500k+ apps from multiple platforms, while [57] reports an operational Fraud Detection System with 183M benign samples alongside ~12k malware

samples. Similarly, large, curated corpora are constructed in [77] (155,300 apps) and [32] (91,203 apps). A second direction emphasizes diversity of acquisition sources to broaden coverage of permission ecosystems and market practices, combining Google Play with third-party or regional stores [39], [48]. Beyond volume and diversity, several studies curate specialized datasets to evaluate threat scenarios underrepresented in general benchmarks. Examples include data-flow–centric profiling using MUDFLOW [36], and collusion-aware benchmarks designed to detect multi-app threats [19]. Other works build datasets as experiment harnesses rather than simple APK pools for instance, stressing third-party library abuse or repackaging behavior [78]. However, such custom datasets can improve ecological validity and target specific attack vectors, cross-study comparability depends strongly on clearly reporting labeling and validation procedures (multi-engine or VirusTotal-style thresholds), split strategy (random vs time-aware), and de-duplication controls, since these design choices can materially inflate or deflate headline performance. Figure (4) shows a summarization for the usage of each dataset in the collected articles.

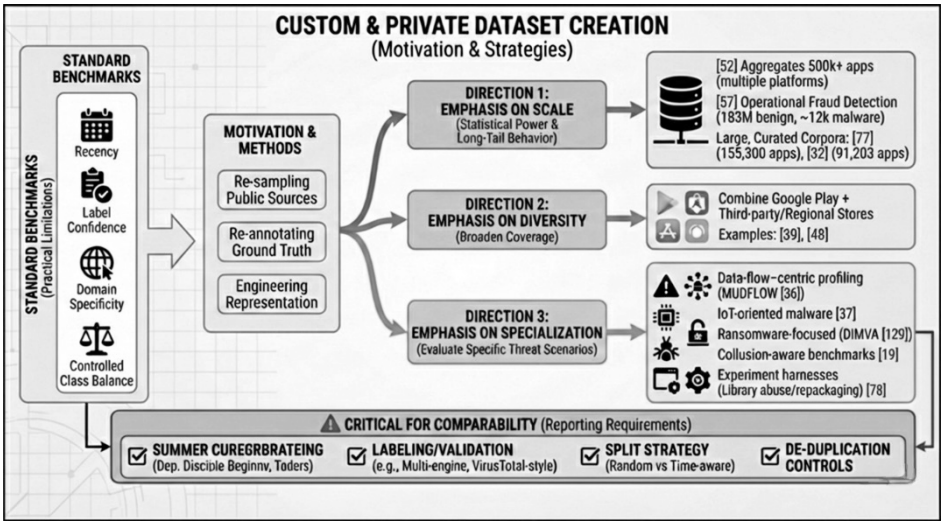


Figure (4): Custom /Private Datasets Usage

4.3 VirusTotal / VirusShare Labeled Datasets

VirusTotal and VirusShare-labeled are typically not fixed benchmarks but constructed by collecting APKs from the wild and then assigning labels using multi-engine antivirus consensus. In this category, VirusShare is mainly used as a malware sourcing repository it distributes very large malware corpora (primarily as downloadable archives torrents, and accompanying hash lists) and currently reports over 108 million samples [40]. VirusTotal, by contrast, is commonly used as a labeling verification layer because it aggregates scan results from many antivirus engines and analysis tools, allowing researchers to operationalize maliciousness through rules like detected by ≥k engines or through vendor-tag queries. VirusTotal-style labeling is inherently noisy, different AV engines may disagree on the same APK due to varying signatures, heuristics, and update cycles, and the set of participating engines can change over time. As a result, ground truth defined by a fixed threshold can be unstable across rescans and may bias evaluation toward samples that are easier to detect. Therefore, studies should report the exact threshold, the scan date, the number of engines, and when possible, provide disagreement statistics to make results comparable. This is reflected in [37] apps verified with VirusTotal, and reported to Google's Android Security Team, and [80] APKs downloaded from VirusTotal during 2016–2017 and labeled as Banking Trojans when flagged by at least five AV tools. Several papers in this set still create custom datasets but rely on VirusTotal-style verification to strengthen label credibility or to focus on a threat slice, [75] constructs a targeted dataset of 4,136 ransomware and benign samples, and [14] builds a ~16,000-app corpus with 103 permission features by filtering AndroZoo-derived data and downloading APKs HTML pages. In addition, an approach is practical because AndroZoo is explicitly designed to provide large-scale APK collections together with metadata such as VirusTotal reports. Not all

studies here are dataset-first, [29] uses a demonstration collusion case to model a realistic SMS-based contact-exfiltration path, emphasizing attack semantics more than sample volume. However, [70] describes a large-scale pipeline decompiling ~200,000 APKs. A key methodological point made explicit in [22] using 70,142 apps from 2011 to 2016 is that VT-based labels can change over time as engines update signatures and heuristics, so the same APK may shift between benign, suspicious, and malicious across re-scans, this temporal label drift is exactly why Figure (5)'s usage summary matters for interpreting comparability across VT/VirusShare-labeled studies.

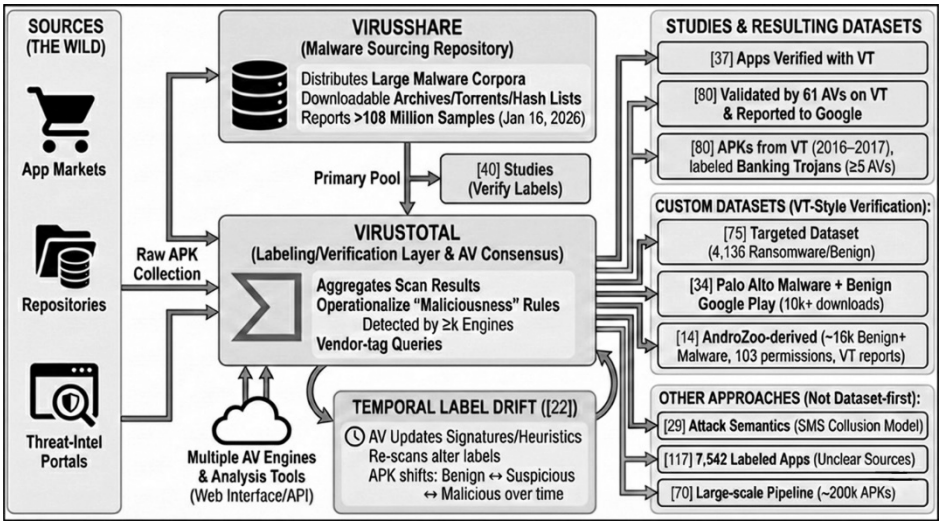


Figure (5): VirusTotal / VirusShare Labeled Datasets

5. Theoretical Gaps and Criticisms

From 2021 onwards, static Android malware detection has been researched significantly which has achieved lots of methodological advancements. However, the existing documentation still presents several theoretical and practical limitations associated with the operational deployment, long-term applicability, and the fair

comparison across different studies of such malware detection techniques. Although many investigations make use of extensive corpora and a rich set of static feature families. Additionaly, most of these studies report laboratory-scale, single-machine pipelines where scalability is seen as an incidental byproduct of the speed of classifiers rather than a fundamental requirement. Consequently, there are still a minority of studies focused on issues around distributed storage, parallel parsing, incremental updating, or reproducible orchestration, and there is no consensus within the field for a big-data reference framework, including versioned dataset management, scalable APK ingestion, multi-view feature extraction, and concept-drift evaluation. In parallel, feature engineering and selection protocols are often built ad hoc based on convenience representations such as frequency vectorization, top-k lists and single- criterion filters, which are poor for balancing accuracy with computational cost, as well as mode of interpretation, robustness to obfuscation and deployment on-device vs cloud. Moreover, there is very little utilization of multi-objective decision-making strategies - such as multi-criteria decision analysis (MCDM). These MCDM could help shed light onto the reasoning behind selecting a particular feature set that is considered optimal in the face of competing goals such as latency, memory consumption, API evolution and adversarial evasion. Adaptivity still has a long way to go, detection mechanisms are usually trained in the offline phase and tested once, despite the constant evolution of malware and android APIs. This oversight has allowed the study of sequential decision-making strategies like budgeted extraction policies and adaptive thresholding in response to varying prevalence and post-deployment feedback, but these are generally yet unexplored in reinforcement learning frameworks. Furthermore, in spite of the fact that transformer-based and inspirations are still emerging, large language models (LLMs) are rarely utilized as controlled structures for semantic abstraction like associations

of low-range artifacts to larger profiles of behavioural supplies, while at the same time prescribing explicit control on security and privacy inside the seeds. Such practices add to the gap between the representational capacity of LLMs and their implementation. Empirical data show a strong evidence of evaluation issue because 42 out of 78 studies explicitly reported split and evaluation protocols, time - split evaluation is only mentioned in one case, 53 studies did not state or the information about splits and evaluation protocols is ambiguous, 23 studies did not document sufficient information about deduplication and leakage controls - procedures that are prone to inflate performance measures if duplicates, variants, or family overlap remain across random splits. Parallel transparency issues include the construction of data, in which 17 of 78 studies have unclear provenance, 13 do not have clear labelling procedures, and 4 do not report VirusTotal (VT) or antivirus (AV) thresholding procedures. Evaluation depth is also equally limited with 11 studies not reporting on robust testing, obfuscation and adversarial evaluations carried out by just 10 studies each, cross dataset or generalization testing by 4 studies, and checks that can be performed on time can only be found in 3 studies. Moreover, measuring computational reporting, such as runtime, cost, latency and scaling constraints, is never quantified. Collectively, these deficiencies provide a good reason for a focused developmental goal, the introduction of reporting checklists including metrics for split type, deduplication, leakage mitigation, replicated run statistics, dissemination of dataset including source, temporal, labelling rules, and thresholds, the creation of a minimum robustness suite including obfuscation and repacking, cross- vs. family- vs. time- split evaluation, computationally specific metrics such as extraction time, throughput, memory footprint, and hardware requirements, and a strategic focus on scalable, multi-objective, and continuously adaptive static detection frameworks.

6. Recommendations for Future Research.

Future research must therefore move from isolated model proposals to an integrated, scalable and decision theoretical research agenda. At the level of systems, researchers should standardize a reference architecture for big-data static malware screen which explicitly provides for distributed feature extraction including the parallel unpacking, decompilation, caching of intermediate artifacts, and resilient parsing, the columnar storage Parquet-based feature lakes, and the incremental learning and evaluation over temporal slices so as to measure the aging and concept drift in the context of realistic market dynamics. Within this pipeline, feature selection should be rethought as a multi-objective design problem. Multi-criteria decision-making (MCDM) techniques, including AHP, TOPSIS, BWM, ELECTRE and PROMETHEE, can be used for scoring and selecting feature families such as permissions, semantic API clusters, certificates, graphs and metadata by simultaneously considering detection performance, extraction cost, robustness to obfuscation and API evolution, explainability and deployment constraints. The resultant output should consist of, instead of having accuracy as the only basis, an auditable rationale for selection that ties features that are selected to explicit operational criteria rather than to accuracy alone. Building on this premise, reinforcement learning, that may take the form of contextual bandits, may be plugged into the control layer of a multi-stage detector. In this setup, an agent can learn to implement policies that decide, for each application, which static analyzes to run, how to allocate computational resources at scale, when to ask for human-in-the-loop verifications, how to change decision thresholds in the face of fluctuations of label prevalence. To reduce the limitations about semantics and interpretability, large language models should be placed under constrained and privacy-aware protocols, preferably using retrieval augmented generation based on vetted security

knowledge bases. Their main functions would include the normalization and the explanation of behavioural proofs, the production of structured descriptions of capabilities based on extracted artifacts and supporting the construction and maintenance of dynamic taxonomies. Furthermore, it will be important for future research to describe the protections, such as schema-constrained outputs, resistance to adversarial prompts, and calibration mechanisms, that will mitigate hallucinations. In addition, researchers should make additional commitments to enhanced reproducibility by publicly disseminating code, feature schemas and temporally stratified evaluation splitting, reporting end-to-end throughput alongside accuracy metrics, and benchmarking robustness in standardized obfuscation, repackaging and dataset shift scenarios. Finaly, these initiatives would turn static Android malware detection from a collection of high accuracy but brittle prototypes into a scalable, explainable and perpetually intensifying ecosystem that is amenable to big data real-world deployment.

7. Conclusion

This survey was completed with a comprehensive assessment of 78 peer-reviewed studies that were carried out between 2021 and 2025. The research focused on the detection of static Android malware. The literature was organized based on the major feature modalities and modeling families used to identify the malicious programs without execution. The research describes the situations in which static detection is most effective and vulnerable, by synthesizing the extraction and modeling of permissions, the API, the opcodes, the manifest, the metadata-certificates, the hybrid features, and the deep architectures. This is especially important in cases such as obfuscation, dataset imbalance, Android API development and insufficient repeatability techniques. Although the taxonomy and dataset

analysis show that great progress is made in semantic representations and reported accuracies. The literature also shows that many solutions are tested in a small context with no scalability as end-to-end solutions. The review highlights the need for the development of novel classifiers and the development of scalable big data pipelines, multi-criteria feature selection, adaptive learning strategies, and trustworthy semantic reasoning layers, as required to shift the progression of the field. This is critical to ensure the reliability, explainability and functionality of static malware detection to adapt to evolving Android and malware ecosystems.

Conflicts of interest

The authors declare that they have no conflicts of interest.

Funding

None.

Acknowledgement

The authors would like to thank Mustansiriyah University (https://uomustansiriyah.edu.iq/) in Baghdad-Iraq for its support in the present work.

7. References

[1] S. F. Ali, M. R. Abdulrazzaq, and M. T. Gaata, (2025), "Learning Techniques-Based Malware Detection: A Comprehensive Review," *Mesopotamian Journal of CyberSecurity*, vol. 5, no. 1, pp. 273–300.

[2] K. Liu, S. Xu, G. Xu, M. Zhang, D. Sun, and H. Liu, (2020), "A Review of Android Malware Detection Approaches Based on Machine Learning," *IEEE Access*, vol. 8, pp. 124579–124607, doi: 10.1109/ACCESS.2020.3006143.

[3] StatCounter 1999-2024, (2026), "Mobile Operating System Market Share Iraq," https://gs.statcounter.com/os-market-share/mobile/iraq, [Online].

[4] Y. Pan, X. Ge, C. Fang, and Y. Fan, (2020), "A Systematic Literature Review of Android Malware Detection Using Static Analysis," *IEEE Access*, vol. 8, pp. 116363–116379, doi:10.1109/ACCESS.2020.3002842.

[5] S. K. Smmarwar, G. P. Gupta, and S. Kumar, (2024), "Android malware detection and identification frameworks by leveraging the machine and deep learning techniques: A comprehensive review," *Elsevier B.V.* doi: 10.1016/j.teler.2024.100130.

[6] F. Alswaina and K. Elleithy, (2020), "Android malware family classification and analysis: Current status and future directions," *MDPI AG*. doi: 10.3390/electronics9060942.

[7] R. A. Yunmar, S. S. Kusumawardani, Widyawan, and F. Mohsen, (2024), "Hybrid Android Malware Detection: A Review of Heuristic-Based Approach," *IEEE Access*, vol. 12, pp. 41255–41286, doi:10.1109/ACCESS.2024.3377658.

[8] A. Alzubaidi, (2021), "Recent Advances in Android Mobile Malware Detection: A Systematic Literature Review," *Institute of Electrical and Electronics Engineers Inc.* doi:10.1109/ACCESS.2021.3123187.

[9] S. R. T. Mat, M. F. A. Razak, M. N. M. Kahar, J. M. Arif, and A. Firdaus, (2022), "A Bayesian probability model for Android malware detection," *ICT Express*, vol. 8, no. 3, pp. 424 – 431, doi:10.1016/j.icte.2021.09.003.

[10] Y. Yao *et al.*, (2024), "Research on Malware Detection Technology for Mobile Terminals Based on API Call Sequence," *Mathematics*, vol. 12, no. 1, doi: 10.3390/math12010020.

[11] O. E. Saied and K. H. Thanoon, (2025), "Static Analysis-based Detection of Android Malware using Machine Learning Algorithms," *Sistemasi: Jurnal Sistem Informasi*, vol. 14, no. 5, pp. 2540–2551.

[12] A. Pathak, T. S. Kumar, and U. Barman, (2024), "Static analysis framework for permission-based dataset generation and android malware detection using machine learning," *EURASIP J. Inf. Secur.*, vol. 2024, no. 1, p. 33.

[13] H. Kato, T. Sasaki, and I. Sasase, (2021), "Android Malware Detection Based on Composition Ratio of Permission Pairs," *IEEE Access*, vol. 9, pp. 130006–130019, doi:10.1109/ACCESS.2021.3113711.

[14] Z. Namrud, S. Kpodjedo, A. Bali, and C. Talhi, (2022), "Deep-Layer Clustering to Identify Permission Usage Patterns of Android App Categories," *IEEE Access*, vol. 10, pp. 24240–24254, doi:10.1109/ACCESS.2022.3156083.

[15] Y. Song, Y. Geng, J. Wang, S. Gao, and W. Shi, (2021), "Permission Sensitivity-Based Malicious Application Detection for Android," *Security and Communication Networks*, vol. (2021), doi:10.1155/2021/6689486.

[16] B. Mishra *et al.*, (2022), "Privacy Protection Framework for Android," *IEEE Access*, vol. 10, pp. 7973–7988, doi: 10.1109/ACCESS.2022.3142345.

[17] A. Banik and J. P. Singh, (2023), "Android Malware Detection by Correlated Real Permission Couples Using FP Growth Algorithm and Neural Networks," *IEEE Access*, vol. 11, pp. 124996–125010, doi: 10.1109/ACCESS.2023.3323845.

[18] D. Ö. Şahın, S. Akleylek, and E. Kiliç, (2022), "LinRegDroid: Detection of Android Malware Using Multiple Linear Regression Models-Based Classifiers," *IEEE Access*, vol. 10, pp. 14246–14259, doi: 10.1109/ACCESS.2022.3146363.

[19] R. Y. Mawoh, J. B. A. Wacka, F. Tchakounté, C. Fachkha, and Kolyang, (2025), "An accurate approach to discriminate android colluded malware from single app malware using permissions intelligence," *Sci. Rep.*, vol. 15, no. 1, doi: https://doi.org/10.1038/s41598-025-86568-w.

[20] M. A. Haq and M. Khuthaylah, (2024), "Leveraging Machine Learning for Android Malware Analysis: Insights from Static and Dynamic Techniques," *Engineering, Technology & Applied Science Research*, vol. 14, no. 4, pp. 15027–15032.

[21] M. İbrahim, B. Issa, and M. B. Jasser, (2022), "A Method for Automatic Android Malware Detection Based on Static Analysis and Deep Learning," *IEEE Access*, vol. 10, pp. 117334–117352, doi:10.1109/ACCESS.2022.3219047.

[22] J. Xu, Y. Li, R. H. Deng, and K. Xu, (2022), "SDAC: A Slow-Aging Solution for Android Malware Detection Using Semantic Distance Based API Clustering," *IEEE Trans. Dependable Secure Comput.*, vol. 19, no. 2, pp. 1149–1163, doi:10.1109/TDSC.2020.3005088.

[23] Y.-C. Chen, H.-Y. Chen, T. Takahashi, B. Sun, and T.-N. Lin, (2021), "Impact of Code Deobfuscation and Feature Interaction in Android Malware Detection," *IEEE Access*, vol. 9, pp. 123208–123219, doi: 10.1109/ACCESS.2021.3110408.

[24] H. Alamro, W. Mtouaa, S. Aljameel, A. S. Salama, M. A. Hamza, and A. Y. Othman, (2023), "Automated Android Malware Detection Using Optimal Ensemble Learning Approach for Cybersecurity," *IEEE Access*, vol. 11, pp. 72509–72517, doi:10.1109/ACCESS.2023.3294263.

[25] F. Zhou, D. Wang, Y. Xiong, K. Sun, and W. Wang, (2025), "Androfim: few-shot android malware family detection based on image representation: F. Zhou *et al.*," *Cybersecurity*, vol. 8, no. 1, p. 69.

[26] O. E. Kural, E. Kiliç, and C. Aksaç, (2023), "Apk2Audio4AndMal: Audio Based Malware Family Detection Framework," *IEEE Access*, vol. 11, pp. 27527–27535, doi:10.1109/ACCESS.2023.3258377.

[27] R. Korine and D. Hendler, (2021), "DAEMON: Dataset/Platform-Agnostic Explainable Malware Classification Using Multi-Stage Feature Mining," *IEEE Access*, vol. 9, pp. 78382–78399, doi:10.1109/ACCESS.2021.3082173.

[28] F. Mercaldo and A. Santone, (2022), "Formal Equivalence Checking for Mobile Malware Detection and Family Classification," *IEEE Transactions on Software Engineering*, vol. 48, no. 7, pp. 2643–2657, doi: 10.1109/TSE.2021.3067061.

[29] D. Liang, L. Shen, Z. Chen, C. Ma, and J. Feng, (2022), "A Formal Method for Description and Decision of Android Apps Behavior Based on Process Algebra," *IEEE Access*, vol. 10, pp. 108668–108683, doi: 10.1109/ACCESS.2022.3210386.

[30] C. Chimeleze *et al.*, (2022), "BFEDroid: A Feature Selection Technique to Detect Malware in Android Apps Using Machine Learning," *Security and Communication Networks*, vol. (2022), doi:10.1155/2022/5339926.

[31] H. Zhu, Y. Li, R. Li, J. Li, Z. You, and H. Song, (2021), "SEDMDroid: An Enhanced Stacking Ensemble Framework for Android Malware Detection," *IEEE Trans. Netw. Sci. Eng.*, vol. 8, no. 2, pp. 984–994, doi: 10.1109/TNSE.2020.2996379.

[32] C. Cilleruelo, Enrique-Larriba, L. De-Marcos, and J.-J. Martinez-Herráiz, (2021), "Malware Detection Inside App Stores Based on Lifespan Measurements," *IEEE Access*, vol. 9, pp.119967–119976, doi: 10.1109/ACCESS.2021.3107903.

[33] S. Garg and N. Baliyan, (2022), "MalVulDroid: Tracing Vulnerabilities from Malware in Android Using Natural Language Processing," *Journal of Web Engineering*, vol. 21, no. 8, pp. 2339–2361, doi: 10.13052/jwe1540-9589.2185.

[34] T. Kacem and S. Tossou, (2025), "Trandroid: An Android Mobile Threat Detection System Using Transformer Neural Networks," *Electronics (Basel).*, vol. 14, no. 6, p. 1230, doi:https://doi.org/10.3390/electronics14061230.

[35] H. J. Hadi, A. Khalid, F. B. Hussain, N. Ahmad, and M. A. Alshara, (2025), "FLSH: A Framework Leveraging Similarity Hashing for Android Malware and Variant Detection," *IEEE Access*, vol. 13, pp. 26142–26156, doi: https://doi.org/10.1109/access.2025.3537110.

[36] H.-J. Zhu, L.-M. Wang, S. Zhong, Y. Li, and V. S. Sheng, (2022), "A Hybrid Deep Network Framework for Android Malware Detection," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 12, pp. 5558–5570, doi: 10.1109/TKDE.2021.3067658.

[37] W. Yuan, Y. Jiang, H. Li, and M. Cai, (2021), "A Lightweight On-Device Detection Method for Android Malware," *IEEE Trans. Syst. Man Cybern. Syst.*, vol. 51, no. 9, pp. 5600–5611, doi:10.1109/TSMC.2019.2958382.

[38] B. B. Gupta *et al.*, (2025), "Earthworm Optimization Algorithm Based Cascade LSTM-GRU Model for Android Malware Detection," *Cyber Security and Applications*, p. 100083, doi:https://doi.org/10.1016/j.csa.2024.100083.

[39] M. Rashid *et al.*, (2025), "Hybrid Android Malware Detection and Classification Using Deep Neural Networks," *International Journal of Computational Intelligence Systems*, vol. 18, no. 1, doi:https://doi.org/10.1007/s44196-025-00783-x.

[40] H. Chen, Z. Li, Q. Jiang, A. Rasool, and L. Chen, (2021), "A hierarchical approach for android malware detection using authorization-sensitive features," *Electronics (Switzerland)*, vol. 10, no. 4, pp. 1 – 24, doi: 10.3390/electronics10040432.

[41] E. Odat and Q. M. Yaseen, (2023), "A Novel Machine Learning Approach for Android Malware Detection Based on the Co-Existence of Features," *IEEE Access*, vol. 11, pp. 15471–15484, doi: 10.1109/ACCESS.2023.3244656.

[42] Z. Meng *et al.*, (2025), "Detecting Android Malware by Visualizing App Behaviors from Multiple Complementary Views," *IEEE Transactions on Information Forensics and Security*, p. 1, doi:https://doi.org/10.1109/tifs.2025.3547301.

[43] Z. Meng *et al.*, (2025), "MalFlows: Context-aware Fusion of Heterogeneous Flow Semantics for Android Malware Detection," *arXiv preprint arXiv:2508.03588*.

[44] X. Xu, S. Jiang, J. Zhao, and X. Wang, (2024), "DCEL: Classifier Fusion Model for Android Malware Detection," *Journal of Systems Engineering and Electronics*, vol. 35, no. 1, pp. 163–177, doi:10.23919/JSEE.2024.000018.

[45] S. Nethala, P. Chopra, K. Kamaluddin, S. Alam, S. A. Alharbi, and M. Alsaffar, (2025), "A Deep Learning-Based Ensemble Framework for Robust Android Malware Detection," *IEEE Access*, p. 1, doi: https://doi.org/10.1109/access.2025.3551152.

[46] Ö. Kiraz and Ibrahim Alper Doğru, (2024), "Visualising static features and classifying android malware using a convolutional neural network approach," *Applied Sciences*, vol. 14, no. 11, p.4772.

[47] A. Al Tawil, D. Qawasmeh, and B. Qawasmeh, (2025), "Enhanced Parrot Optimizer Algorithm: A Proposed Method for Optimized Malware Classification," *International Journal of Artificial Intelligence Applications*, vol. 1, no. 1.

[48] A. O. Otiko and G. A. Iyang, (2025), "Android Malware Classification with Feature Selection using Artificial Bee Colony Algorithm," *Available at SSRN 5293728*.

[49] I. Almomani *et al.*, (2021), "Android Ransomware Detection Based on a Hybrid Evolutionary Approach in the Context of Highly Imbalanced Data," *IEEE Access*, vol. 9, pp. 57674–57691, doi:10.1109/ACCESS.2021.3071450.

[50] A. Taha, A. H. Osman, and Y. S. Baguda, (2025), "An Intelligent Technique for Android Malware Identification Using Fuzzy Rank-Based Fusion," *Technologies (Basel).*, vol. 13, no. 2, p. 45, doi: https://doi.org/10.3390/technologies13020045.

[51] R. S. Arslan, (2025), "JDroid: Android malware detection using hybrid opcode feature vector," *PeerJ Comput. Sci.*, vol. 11, p. e3051.

[52] B. Urooj, M. A. Shah, C. Maple, M. K. Abbasi, and S. Riasat, (2022), "Malware Detection: A Framework for Reverse Engineered Android Applications Through Machine Learning Algorithms," *IEEE Access*, vol. 10, pp. 89031–89050, doi:10.1109/ACCESS.2022.3149053.

[53] F. A. Almarshad, M. Zakariah, G. A. Gashgari, E. A. Aldakheel, and A. I. A. Alzahrani, (2023), "Detection of Android Malware Using Machine Learning and Siamese Shot Learning Technique for Security," *IEEE Access*, vol. 11, pp. 127697–127714, doi:10.1109/ACCESS.2023.3331739.

[54] L. da Costa and V. Moia, (2023), "A Lightweight and Multi-Stage Approach for Android Malware Detection Using Non-Invasive Machine Learning Techniques," *IEEE Access*, vol. 11, pp. 73127–73144, doi: 10.1109/ACCESS.2023.3296606.

[55] M. M. Alani and A. I. Awad, (2022), "PAIRED: An Explainable Lightweight Android Malware Detection System," *IEEE Access*, vol. 10, pp. 73214–73228, doi:10.1109/ACCESS.2022.3189645.

[56] P. R. Sekhar, V. Sohana, R. Shaik, P. Thanuja, and K. J. Mary, (2025), "Lightweight Machine Learning-Based Static Analysis Framework for Real-Time Android Malware Detection," *Synthesis: A Multidisciplinary Research Journal*, vol. 3, no. 1s, pp. 60–69.

[57] J.-H. Shin, D. Y. Kim, and K. Lee, (2025), "Advanced Financial Fraud Malware Detection Method in the Android Environment," *Applied Sciences*, vol. 15, no. 7, p. 3905, doi:https://doi.org/10.3390/app15073905.

[58] M. Jalandra, M. Kuliha, and J. Kaur, (2025), "Hybrid Machine Learning Model for Android Ransomware Detection Using URL-Based Features," *IJSAT-International Journal on Science and Technology*, vol. 16, no. 3.

[59] M. Tanvirul Alam, D. Bhusal, and N. Rastogi, (2024), "Revisiting Static Feature-Based Android Malware Detection," *arXiv e-prints*, p. arXiv–2409.

[60] A. Faiz, M. Fuzail, A. Aftab, N. Aslam, and A. J. Akbar, (2024), "A Multimodal Machine Learning Approach for Android Malware Detection: Static Code Analysis," *International Journal of Information Systems and Computer Technologies*, vol. 4, no. 1, pp. 20–35.

[61] S. Akbar and T. Khan, (2025), "Malware Detection Using Static Feature Analysis and Deep Learning Techniques: A Static Robust DL-Based Model for Android Malware Detection," *Information Technology and Control*, vol. 54, no. 4, pp. 1358–1382.

[62] H. Berger, C. Hajaj, E. Mariconti, and A. Dvir, (2022), "Crystal Ball: From Innovative Attacks to Attack Effectiveness Classifier," *IEEE Access*, vol. 10, pp. 1317–1333, doi:10.1109/ACCESS.2021.3138628.

[63] G. Xu *et al.*, (2024), "OFEI: A Semi-Black-Box Android Adversarial Sample Attack Framework Against DLaaS," *IEEE Transactions on Computers*, vol. 73, no. 4, pp. 956–969, Apr. doi:10.1109/TC.2023.3236872.

[64] G. You, G. Kim, S. Han, M. Park, and S.-J. Cho, (2022), "Deoptfuscator: Defeating Advanced Control-Flow Obfuscation Using Android Runtime (ART)," *IEEE Access*, vol. 10, pp. 61426–61440, doi: 10.1109/ACCESS.2022.3181373.

[65] D. Li, S. Cui, Y. Li, J. Xu, F. Xiao, and S. Xu, (2024), "PAD: Towards Principled Adversarial Malware Detection Against Evasion Attacks," *IEEE Trans. Dependable Secure Comput.*, vol. 21, no. 2, pp. 920–936, doi: 10.1109/TDSC.2023.3265665.

[66] A. Rashid and J. Such, (2023), "MalProtect: Stateful Defense Against Adversarial Query Attacks in ML-Based Malware Detection," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 4361–4376, doi: 10.1109/TIFS.2023.3293959.

[67] R. Yumlembam, B. Issac, S. M. Jacob, and L. Yang, (2023), "IoT-Based Android Malware Detection Using Graph Neural Network With Adversarial Defense," *IEEE Internet Things J.*, vol. 10, no. 10, pp. 8432–8444, doi: 10.1109/JIOT.2022.3188583.

[68] J. Wang, X. H. Liu, M. Huang, P. Zhou, and Y. Xu, (2025), "Android Malware Detection Method Combining Multi-Frequency Features and Convolutional Neural Networks," *IEEE Access*, p. 1, doi: https://doi.org/10.1109/access.2025.3550124.

[69] L. N. Vu and S. Jung, (2021), "AdMat: A CNN-on-Matrix Approach to Android Malware Detection and Classification," *IEEE Access*, vol. 9, pp. 39680–39694, doi:10.1109/ACCESS.2021.3063748.

[70] Y. Hei *et al.*, (2024), "Hawk: Rapid Android Malware Detection Through Heterogeneous Graph Attention Networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 4, pp. 4703–4717, doi: 10.1109/TNNLS.2021.3105617.

[71] Y. Liu, M. G. M. Johar, and J. Tham, (2025), "An Android Malware Detection Method Based on MLSTM," *Journal of Digital Information Management*, vol. 23, no. 1, pp. 11–25, doi:https://doi.org/10.6025/jdim/2025/23/1/11-25.

[72] H. Zhao, Y. Wu, D. Zou, and H. Jin, (2024), "An Empirical Study on Android Malware Characterization by Social Network Analysis," *IEEE Trans. Reliab.*, vol. 73, no. 1, pp. 757–770, doi:10.1109/TR.2023.3304389.

[73] I. Almomani, A. Alkhayer, and W. El-Shafai, (2022), "An Automated Vision-Based Deep Learning Model for Efficient Detection of Android Malware Attacks," *IEEE Access*, vol. 10, pp. 2700–2720, doi: 10.1109/ACCESS.2022.3140341.

[74] J. Yang, Y. Li, H. Ren, D. Jia, and X. Wang, (2025), "SAC: Collaborative learning of structure and content features for Android malware detection framework," *Neurocomputing*, p. 130053, doi:https://doi.org/10.1016/j.neucom.2025.130053.

[75] H. Rodriguez-Bazan, G. Sidorov, and P. J. Escamilla-Ambrosio, (2023), "Android Ransomware Analysis Using Convolutional Neural Network and Fuzzy Hashing Features," *IEEE Access*, vol. 11, pp. 121724–121738, doi: 10.1109/ACCESS.2023.3328314.

[76] B. Menaouer, A. El Hadj Mohamed Islem, and M. Nada, (2023), "Android Malware Detection Approach Using Stacked AutoEncoder and Convolutional Neural Networks," *International Journal of Intelligent Information Technologies*, vol. 19, no. 1, doi:10.4018/IJIIT.329956.

[77] J. Xu and Q. Yuan, (2022), "LibRoad: Rapid, Online, and Accurate Detection of TPLs on Android," *IEEE Trans. Mob. Comput.*, vol. 21, no. 1, pp. 167–180, doi:10.1109/TMC.2020.3003336.

[78] J. Qiu, X. Yang, H. Wu, Y. Zhou, J. Li, and J. Ma, (2022), "LibCapsule: Complete Confinement of Third-Party Libraries in Android Applications," *IEEE Trans. Dependable Secure Comput.*, vol. 19, no. 5, pp. 2873–2889, doi: 10.1109/TDSC.2021.3075817.

[79] C. Sun, Y. Ma, D. Zeng, G. Tan, S. Ma, and Y. Wu, (2023), "μDep: Mutation-Based Dependency Generation for Precise Taint Analysis on Android Native Code," *IEEE Trans. Dependable Secure Comput.*, vol. 20, no. 2, pp. 1461–1475, doi: 10.1109/TDSC.2022.3155693.

[80] C. Bai, Q. Han, G. Mezzour, F. Pierazzi, and V. S. Subrahmanian, (2021), "DBankDBank: Predictive Behavioral Analysis of Recent Android Banking Trojans," *IEEE Trans. Dependable Secure Comput.*, vol. 18, no. 3, pp. 1378–1393, doi: 10.1109/TDSC.2019.2909902.

[81] Y. Cui, Y. Sun, J. Luo, Y. Huang, Y. Zhou, and X. Li, (2022), "MMPD: A Novel Malicious PDF File Detector for Mobile Robots," *IEEE Sens. J.*, vol. 22, no. 18, pp. 17583–17592, doi:10.1109/JSEN.2020.3029083.

[82] M. Dener, G. Ok, and A. Orman, (2022), "Malware Detection Using Memory Analysis Data in Big Data Environment," *Applied Sciences (Switzerland)*, vol. 12, no. 17, doi:10.3390/app12178604.

[83] S. Seneviratne, R. Shariffdeen, S. Rasnayaka, and N. Kasthuriarachchi, (2022), "Self-Supervised Vision Transformers for Malware Detection," *IEEE Access*, vol. 10, pp. 103121–103135, doi:10.1109/ACCESS.2022.3206445.

[84] M. Naseer *et al.*, (2025), "Obfuscated Malware Detection and Classification in Network Traffic Leveraging Hybrid Large Language Models and Synthetic Data," *Sensors*, vol. 25, no. 1, p. 202, doi: https://doi.org/10.3390/s25010202.

[85] M. Maray, M. Maashi, H. M. Alshahrani, S. S. Aljameel, S. Abdelbagi, and A. S. Salama, (2024), "Intelligent Pattern Recognition Using Equilibrium Optimizer With Deep Learning Model for Android Malware Detection," *IEEE Access*, vol. 12, pp. 24516–24524, doi:10.1109/ACCESS.2024.3357944.

[86] J. Yu, Y. He, Q. Yan, and X. Kang, (2021), "SpecView: Malware Spectrum Visualization Framework With Singular Spectrum Transformation," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 5093–5107, doi: 10.1109/TIFS.2021.3124725.

Current Condition of the Holy Kadhimiya Shrine Surveying Using a GPR / Original Article

Assist. Prof. Dr. Israa Hatem

Roads and Geotechnical Engineering Department,
University of Technology, Baghdad/Iraq
e-mail: israa.h.mohammed@uotechnology.edu.iq

الوضع الحالي لصحن الامام الكاظم الشريف مسح باستخدام الرادار المخترق للأرض

أ.م.د اسراء حاتم محمد

الجامعة التكنولوجية، قسم هندسة الطرق والجيوتكنيك، بغداد\العراق

Abstract

This study describes the extensive survey of the Holy Kadhimiya Shrine in Baghdad, Iraq, carried out with the use of a non-invasive geophysical instrument called Ground Penetrating Radar (GPR) to assess the conditions of subsurface and the condition of the infrastructure in the site. The total area of the shrine is 19,575 m^2 that was systematically studied with the help of internal measures, satellite imagery and radar profiling. The frequency antennas used for the GPR survey were the 250 MHz, 500 MHz, and 800 MHz which have a penetration depth of 1-15 m. Specialized software packages (Ground Vision, Reexplore and Object Mapper) were used for data processing to ensure the correct interpretation of the subsurface anomalies. The results were as follows: (1) The precise topographic surveys allowed a reliable matching by radar; (2) Water and sewage pipelines were located near the main gates; especially rain water drainage lines were interpreted as remains of an ancient tomb; and (3) No unusual geological formations were detected, except for historical brick layers which do not affect stability of soil. In conclusion the research conducted could conclude that overall GPR is effective instrument for investigating heritage sites, including the engineering and archaeological applications. Aerial photographs taken using radar imaging have been helpful in gaining an understanding of the area; it is recommended that further geotechnical validation work be undertaken by collecting soil samples to determine bearing capacity and to feed into plans for future restoration or development. The application of geophysical methods together with traditional survey is an integrated survey method to preserve and manage cultural landmarks that can also be adopted as a model for other studies in historic urban environment.

Keywords: GPR, Kadhimiya, Surveying, Subsurface utility

المستخلص

تصف هذه الدراسة المسح الشامل للمرقد الكاظمي الشريف في بغداد، العراق، والذي تم إجراؤه باستخدام جهاز جيوفيزيائي غير تدميري تسمى الرادار المخترق الأرضي (GPR) لتقييم الظروف تحت السطح وحالة البنية التحتية في الموقع. تبلغ المساحة الإجمالية للضريح الشريف 19.575 مترًا مربعًا، وقد تمت دراستها بشكل منهجي بمساعدة القياسات الداخلية وصور الأقمار الصناعية والتصوير الراداري. وكانت هوائيات التردد المستخدمة في مسح GPR هي 250 ميجا هرتز و500 ميجا هرتز و800 ميجا هرتز والتي يبلغ عمق اختراقها 1–15 مترًا. تم استخدام حزم البرامج المتخصصة (Ground Vision وReexplore وObject Mapper) لمعالجة البيانات لضمان التفسير الصحيح للشذوذات الموجودة تحت السطح. وكانت النتائج على النحو التالي: (1) أتاحت المسوحات الطبوغرافية الدقيقة مطابقة موثوقة بواسطة الرادار؛ (2) وجود خطوط أنابيب المياه والصرف الصحي بالقرب من البوابات الرئيسية؛ وخاصةً خطوط تصريف مياه الأمطار تم تفسيرها على أنها بقايا مقبرة قديمة؛ (3) لم يتم اكتشاف أي تكوينات جيولوجية غير عادية، باستثناء طبقات الطوب التاريخية التي لا تؤثر على استقرار التربة. في الختام، يمكن أن يخلص البحث الذي تم إجراؤه إلى أن GPR بشكل عام هو أداة فعالة لدراسة المواقع التراثية، بما في ذلك التطبيقات الهندسية والأثرية. الصور الملتقطة باستخدام التصوير الراداري ساعدت في فهم المنطقة. يوصى بإجراء المزيد من أعمال التحقق الجيوتقنية من خلال جمع عينات التربة لتحديد قدرة التحمل وإدخالها في خطط الترميم أو التطوير المستقبلية. ويعتبر تطبيق الأساليب الجيوفيزيائية مع المسح التقليدي أسلوباً مسحياً متكاملاً للحفاظ على المعالم الثقافية وإدارتها، كما يمكن اعتماده كنموذج لدراسات أخرى في البيئة الحضرية التاريخية.

الكلمات المفتاحية: GPR، الكاظمية، المساحة، البنى التحتية

1. Introduction

The total area of the shrine, including the courtyard and mosque, is 19,575 square meters. It is rectangular in shape, measuring 140 meters by 135 meters. The shrine building is square, with each side measuring 60 meters, and has a total area of 3,600 square meters. The shrine building connects to the mosque on its northern side, while the eastern, southern, and western sides form the courtyard, which is divided into: The courtyard of Bab al-Murad (the eastern section), The courtyard of Bab al-Qibla (the southern section) and the courtyard of Quraish (the western section).

The courtyard is surrounded on all four sides by a wall approximately ten meters high. Internally, it is divided into two levels. The first level contains rooms and porticoes, while the second level is an open wall, divided into arched openings and ribs, precisely in the shape of the porticoes. There are sixty-three porticoes leading to rows that were originally intended to receive visitors, replacing the former hotels. Another section was designated for teaching religious sciences. Sixty-three rooms are connected to the porticoes, in addition to fourteen rooms connected to the entrances, bringing the total to seventy-six rooms. One of the porticoes on the eastern side of the wall contains two entrances for the restrooms, one for men and one for women, and one of its rooms was designated as a library.[1]

There are three main gates, uniform in design and size: Bab al-Murad on the eastern side, Bab al-Qibla on the southern side and Bab Sahib al-Zaman on the western side. Each gate is 5.10 meters high and four meters wide. Their facades are decorated inside and out with marble tiles. Opposite each gate is a portico bearing its name. There are also five medium-sized gates distributed along the four walls of the courtyard: two in the north, two in the south, and one in the east. There are two small gates, one in the northeastern corner of the courtyard and the other in the southwestern corner of the western wall of the courtyard.

2. Data Collection:

One of the most important aspects of any consultancy study is collecting data about the study area. The larger the dataset, the better the projections. The data collected for the Kadhimiya Shrine in particular, and the city of Kadhimiya in general, varied between satellite imagery, topographic maps, sectional maps, and infrastructure maps.[2,3]

2.1 Satellite Imagery

A collection of satellite images covering the entire study area was prepared from various sources with a resolution of 60 cm, as shown in the Figures (1) and (2) below. These images, taken at different times, were collected to identify changes that have occurred in the study area in recent years.

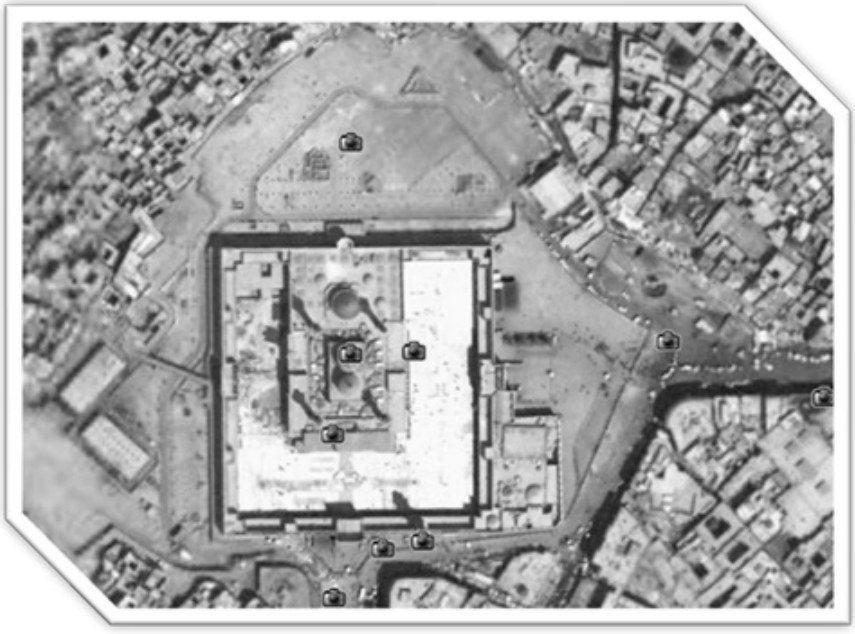


Figure (1): satellite images covering area of the shrine

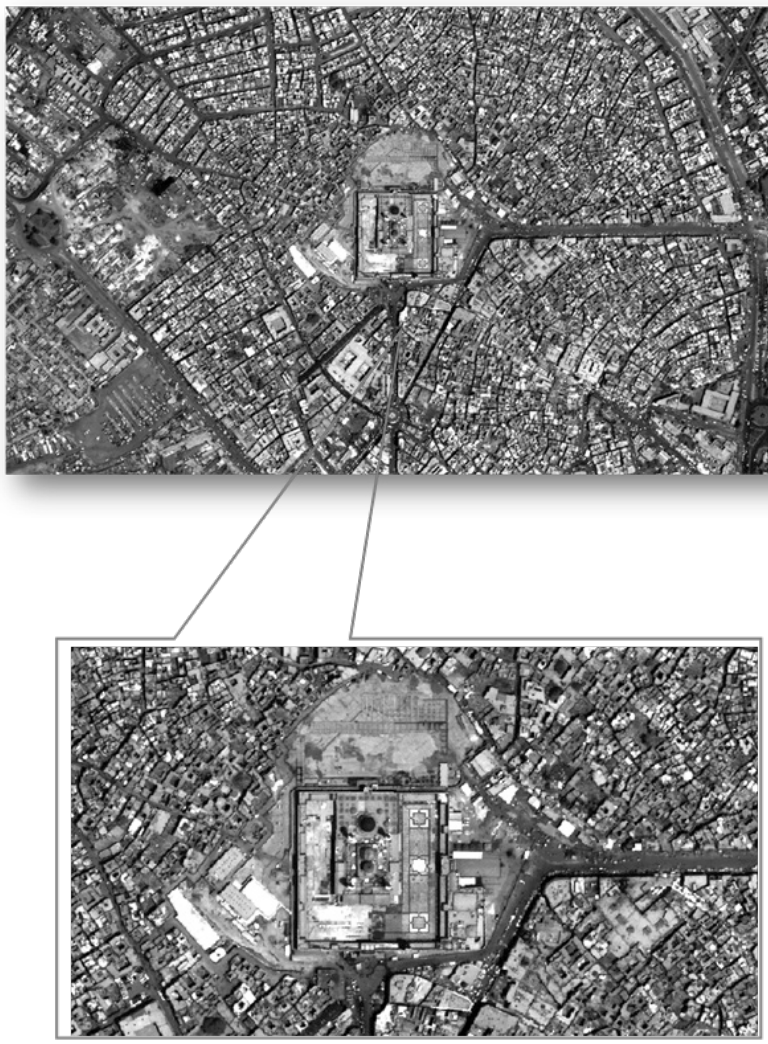


Figure (2): satellite images covering the entire study

2.2 Use of a GPR Device

Ground penetrating radar (GPR) is a nondestructive surveying method which is continuous, fast and has been applied to many applications for mapping subsurface conditions.[4] There are numerous applications that are suitable for the evaluation

of highway systems. GPR is essentially a subsurface “anomaly” detector; therefore, it will provide a profile of the underground changes as a consequence of contrasts in the electromagnetic conductivity of the materials at interfaces [5] within the ground and obtain a realistic image of the infrastructure of the Holy Kadhimiya Shrine. The GPR equipment currently used by University of Technology is summarized in Table (1). This equipment consists of data collection system (MALA geoscience), data collection and analysis software packages (Ground Vision, Object mapper, and RadExplorer), 3 ground-coupled shielded antennas (250 MHz, 500 MHz, 800 MHz, shown in Table (2). Table (1) lists the approximate depths for penetration that each of these antennas gives. Overall, the depth of penetration decreases as the speed of transmission of the antenna increases.[6] In addition, a wheel cart for GPR data collection, including an independent power source, GPS data collection, electronic DMI device and camera for the project.

Table (1): GPR equipment currently used by University of Technology

Data collection unit	ProEx Control Unit
Collection software	GroundVision 2
Processing software	RadExplorer, Object mapper
Antenna / depth of penetration	250 MHz / 8-15 m 500MHz / 3.5-8 m 800 MHz / 1-3 m

Table (2): Antenna Technical Specification

Antenna	Dimensions	Applications
250 MHz	0.74 x 0.44 x 0.16 m – Weight: 7.85 kg	Utility detection, Underground Storage Tank and void detection.
500MHz	0.50 x 0.30 x 0.16 m – Weight: 5.0 kg	Utility detection, road and pavement surveys
800 MHz	0.38 x 0.20 x 0.12 m – Weight: 2.6 kg	Road/ pavement surveys and concrete/ structural investigations.

3. Methodology and approach

The field surveys consist Phase One: Internal measurements of the Holy Courtyard and Phase Two: Processing and analysis. Phase One: Internal Measurements of the Holy Courtyard

This section includes surveying all the flooring within the Holy Courtyard, including buildings, roofs, corridors, etc. see Figure (3).

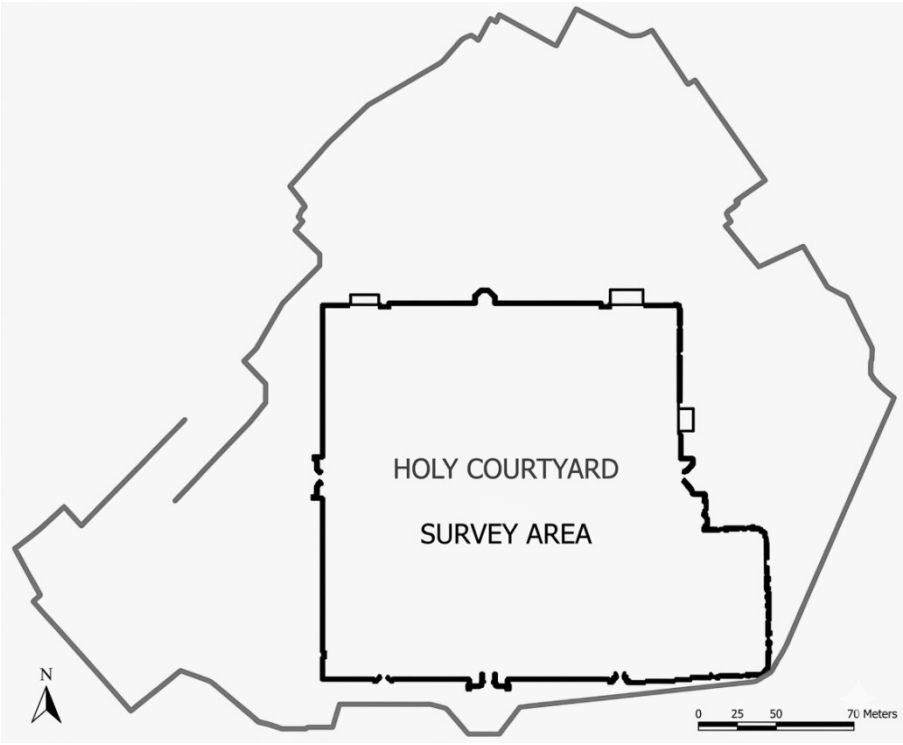


Figure (3): Internal Measurements of the Holy Courtyard

4. Processing and Analysis

During this phase, several programs specific to the GPR device were used for processing and analysis, including Rad Explorer. These are some samples of ground-penetrating radar images of the Holy Kadhimiya Shrine. The basis of GPR is to probe the ground with short pulses of electromagnetic energy. One-dimensional electromagnetic wave propagation theory is used to analyse the reflected images of these pulses. The amplitude and arrival time of these pulses are reflected back to the antenna and related to the electrical conductivity (or dielectric constant) of the material layers Figure (4). For the interfaces, a certain amount of energy is reflected, depending on the dielectric contrast of the material, a certain amount is absorbed. The above-mentioned peaks in amplitude (in the order they are observed) are the antenna end reflection (A0), the surface (pavement) reflection (A1) and the base reflection (A2), respectively. Time interval (t1) between peaks A1 and A2 is the 2-way travel time through the pavement layer. The thickness of the pavement layer (h) as shown in Figure (4), can be calculated as:

$$h = v \Delta t / 2 \quad (1)$$

where v = propagation velocity through each layer and $\Delta t = t2-t1$

The propagation velocity is related to the electromagnetic behavior of the asphalt layer:

$$v = c / \sqrt{\varepsilon} \quad (2)$$

where ε = dielectric constant of asphalt layer

c = speed of light in air = 0.30 m/ns

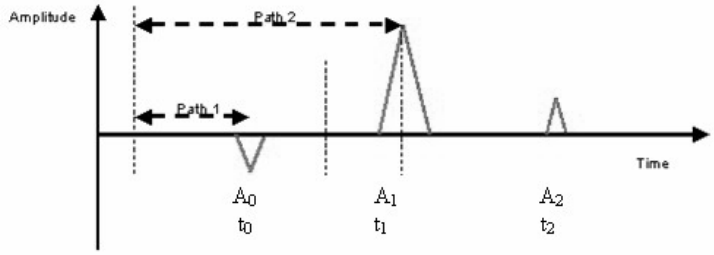
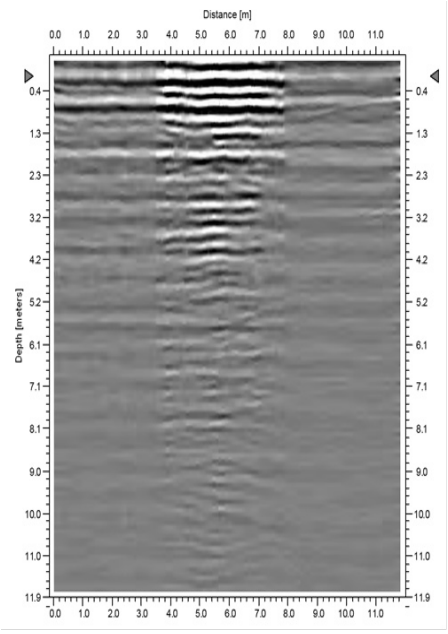
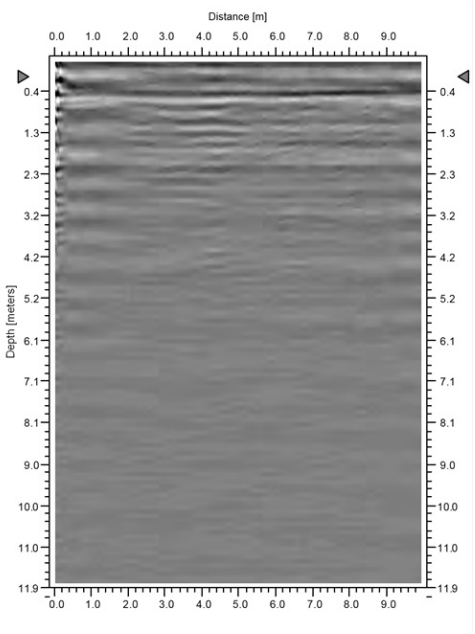


Figure (4): Typical Amplitude Time History from GPR Antenna [7]

5. Results and discussion

The use of Ground Penetrating Radar (GPR) on the Holy Kadhimiya Shrine site provided important results not only from the engineering point of view, but also for the historical aspects of the subsurface features. Combined with topographic surveys and satellite imagery, the use of radar was able to present a detailed overview of the current state of the shrine, as well as some indications of its history. One of the most significant discoveries was the identification of pipelines of water and sewerage, particularly the rain water drainage pipes which are positioned in front of the main gates of the courtyard. The radar profiles could easily be seen these networks in the radar profiles because they depended on the differences in soil and pipe dielectrics. The detection is a testament to the real-world applications of GPR in infrastructure management, enabling non-destructive mapping of critical infrastructure. This is especially useful if excavation will be limited in heritage sites. The results provide a foundation for future maintenance planning and preventive measures to safeguard the shrine's structural integrity as shown in Figures (5) and (6).





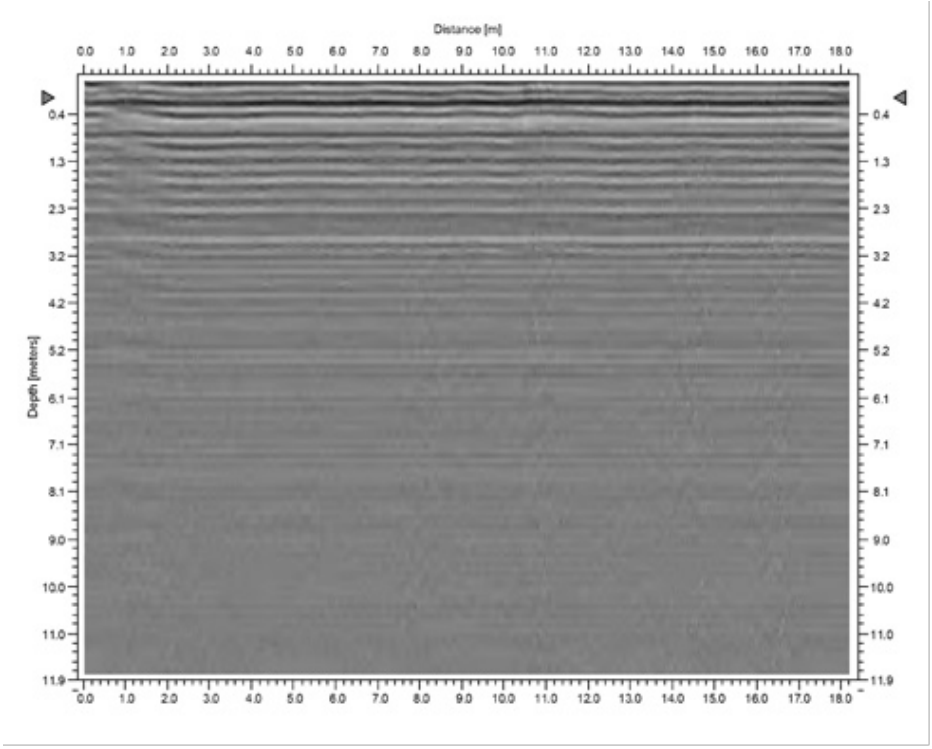


Figure (5): Typical Amplitude results from GPR Antenna

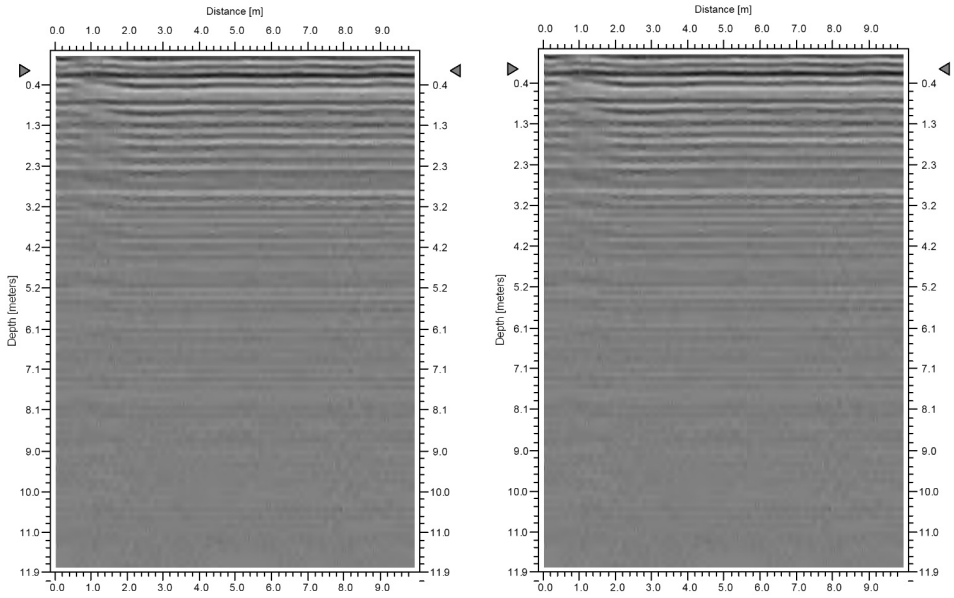


Figure (6): Typical Amplitude results from GPR Antenna border site

The results show the potential of GPR as a non-invasive method for heritage site investigation. It can be used in the detection of utilities, historic features and subsurface layers without excavation, making it especially appropriate for religious and cultural sites. Typically, a drainage line and soil layers are identified from an engineering standpoint, which aids in infrastructure maintenance and risk assessment. An archaeological point of view, the identifying of the remains of ancient tombs adds to the understanding of the history of the shrine. The combination of GPR data with satellite imagery and topographic surveys is a multi-layered data set that increases the confidence of the interpretations. This multi-disciplinary approach can be used to attain both engineering and cultural heritage goals, and serves as a model for studies in other historical sites.

6. Conclusions

Based on the scientific principles followed in using radar data and determining depths and foreign objects underground, the following can be concluded:

1- The topographic survey required for radar matching was highly accurate and met the requirements of radar analysis.

2- Water and sewage lines, especially rainwater drainage lines, were identified in front of the main gates of the courtyard.

3- There are simple brick structures in the courtyard facing the Al-Jawadain Mosque, measuring 10x10 meters, which represent the remains of an ancient tomb.

4- The radar did not indicate any unusual geological structure in the vertical basins surrounding the courtyard, except for layers of brick dating back to earlier periods that do not affect the soil composition from a geological perspective.

5- For the purpose of accurate geotechnical matching, we recommend taking soil samples (test pits) at a rate of 4 to 6 pits that represent the general soil profile in order to determine the soil bearing capacity.

7. References

[1] Mele, M., Degli Esposti, M., Giudici, M., Comunian, A., Al Tamimi, A.M., Al Aufi, A.S., & Zerboni, A. (2025). *Evaluating Subsurface Risk for Archaeological Heritage Through Ground-Penetrating Radar Surveys: The Case Study of Bisya and Salūt Archaeological Site (Sultanate of Oman).* Heritage, 8(10), 399. https://doi.org/10.3390/heritage8100399

[2] Conyers, L.B. (2012). *Interpreting Ground-Penetrating Radar for Archaeology.* Left Coast Press. (Widely cited textbook on archaeological GPR applications).

[3] Kadioglu, S., Kadioglu, Y.K., Catapano, I., & Soldovieri, F. (2013). *Ground penetrating radar and microwave tomography for the safety management of a cultural heritage site: Miletos Ilyas Bey Mosque (Turkey).* Journal of Geophysics and Engineering, 10(6), 064007. https://doi.org/10.1088/1742-2132/10/6/064007

[4] Işik, N. (2021). *Investigation of structural problems on the ground and walls of Diyarbakir Historical Behram Pasha Mosque by the georadar (GPR) method.* Dicle University Journal of Architecture. https://doi.org/10.24012/dumf.1002241

[5] Springer (2026). *Predicting CBR values using Gaussian process regression and meta-modeling.* Springer Geotechnical Engineering Series. (Recent study on soil bearing capacity estimation using advanced methods).

[6] Jol, H.M. (2009). Ground Penetrating Radar Theory and Applications. Elsevier.

[7] Springer (2025). Ground Penetrating Radar System: Principles. (Overview of GPR applications in cultural heritage).

Spatial Estimation for Gas Flaring and Natural Gas Waste from Oil Production using Remote Sensing and GIS Techniques / Original Article

Lact. Dr. Hisham M. Jawad Al Sharaa

Department of Geomatics Engineering, College of Civil Engineering, University of Technology, Baghdad/Iraq.

e-mail: hisham.m.alsharaa@uotechnology.edu.iq

التقدير المكاني لحرق الغاز ونفايات الغاز الطبيعي الناتجة عن إنتاج النفط باستخدام تقنيات الاستشعار عن بعد ونظم المعلومات الجغرافية

المدرس الدكتور هشام محمد جواد الشرع

الجامعة التكنولوجية، كلية الهندسة المدنية، قسم هندسة الجيوماتيك، بغداد|العراق.

Abstract

This study aims to provide a comprehensive geospatial gas flaring intensity and emission doses analysis throughout oil production areas in Iraq based on remote sensing and GIS-based interpolation methods. Satellite image and spatial data were combined to create a map of flaring locations, refinery locations, and associated CO_2-equivalent emissions. The results show that there is a clear spatial variation in the intensity of the emissions ranging from northern regions towards the oil fields in the south near Basra, Rumaila and West Qurna, where refinery density and high flaring activities are observed. Interpolation modelling suggests that the majority of Iraq's GHG emissions are concentrated in these areas, and in some localised areas, emissions are significantly higher than very high limits. The study shows the potential of reducing emissions by flaring reduction and gas recovery projects and how GIS can be used to quantify environmental impacts resulting from hydrocarbon operations. The results offer scientific evidence for making energy policy and utilizing energy resources sustainably, which will help Iraq reduce routine gas flaring to zero by 2030.

Keywords Natural Gas, Oil Production, Remote Sensing, GIS Techniques Geomatics, gas flaring.

المستخلص

تهدف هذه الدراسة إلى تقديم تحليل شامل لكثافة حرق الغاز وجرعات الانبعاثات في جميع مناطق إنتاج النفط في العراق، وذلك بالاعتماد على تقنيات الاستشعار عن بُعد وأساليب الاستيفاء القائمة على نظم المعلومات الجغرافية. وقد جُمعت صور الأقمار الصناعية والبيانات المكانية لإنشاء خريطة لمواقع حرق الغاز، ومواقع المصافي، وانبعاثات ثاني أكسيد الكربون المكافئة المرتبطة بها. حيث تُظهر النتائج وجود تباين مكاني واضح في كثافة الانبعاثات، يمتد من المناطق الشمالية باتجاه حقول النفط في الجنوب قرب البصرة وخاصة حقل الرميلة وغرب القرنة، حيث تُلاحظ كثافة عالية للمصافي وأنشطة حرق الغاز المكثفة. وتشير نماذج الاستيفاء إلى أن غالبية انبعاثات الغازات الدفيئة في العراق تتركز في هذه المناطق والتي تتجاوز فيها الانبعاثات الحدود القصوى المسموح بها بشكل ملحوظ. تُبين الدراسة إمكانية خفض الانبعاثات من خلال مشاريع الحد من حرق الغاز واستخلاصه، وكيف يُمكن استخدام نظم المعلومات الجغرافية لتقييم الآثار البيئية الناتجة عن عمليات استخراج الهيدروكربونات. تُقدم النتائج أدلة علمية لوضع سياسات الطاقة واستخدام مواردها بشكل مستدام، مما سيساعد العراق على خفض حرق الغاز الروتيني إلى الصفر بحلول عام 2030.

الكلمات المفتاحية: الغاز الطبيعي، إنتاج النفط، الاستشعار عن بعد، تقنيات نظم المعلومات الجغرافية، الجيوماتيك، حرق الغاز.

1. Introduction

Enterprises in the south region work mainly with oil, while in the north region, gas is developed along with oil Figure (1). The main products of the gas production industry are natural gas (combustible natural gas (natural), gas condensate, liquefied natural gas (combustible natural liquefied gas) and regasified liquefied natural gas (combustible natural regasified gas).[1]

Natural gas is a gaseous mixture consisting of methane and heavier hydrocarbons, nitrogen, carbon dioxide, water vapor, sulfur-containing compounds, inert gases. Methane is the main component of natural gas, other components are found in trace amounts.[2,3]

Gas condensate is a liquid mixture consisting of paraffinic, naphthenic and aromatic hydrocarbons of a wide fractional composition, containing impurities of non-hydrocarbon components, obtained as a result of separation of a gas condensate mixture.[4,5]

Distinguish between unstable and stable gas condensate. Unstable gas condensate - gas condensate containing dissolved gaseous hydrocarbons, sent for processing in order to remove impurities and release hydrocarbons. Natural gas production that meets the requirements. Impurities include water (aqueous solutions of corrosion and / or hydrate inhibitors), chloride salts, sulfur compounds and mechanical impurities.[6]

Stable gas condensate - gas condensate obtained by cleaning unstable gas condensate from impurities and separating hydrocarbons from it ITS 29-2017 Natural gas production- ITS 29-2017 Natural gas production that meets the requirements of GOST R 54389-2011 (clause 9 of GOST R 53521).

Liquefied natural gas (LNG) - natural gas liquefied for storage or transportation.

The process of converting LNG from a liquid to a gaseous state is called regasification, which results in regasified LNG (combustible natural regasified gas).

According to preliminary data from the Ministry of Oil (Iraq), gas consumption in the domestic market in 2015 decreased compared to 2014 by 4.0% and amounted to 441.1 billion m3 Natural gas production, which is mainly due to a decrease in gas supplies to power plants Iraq.

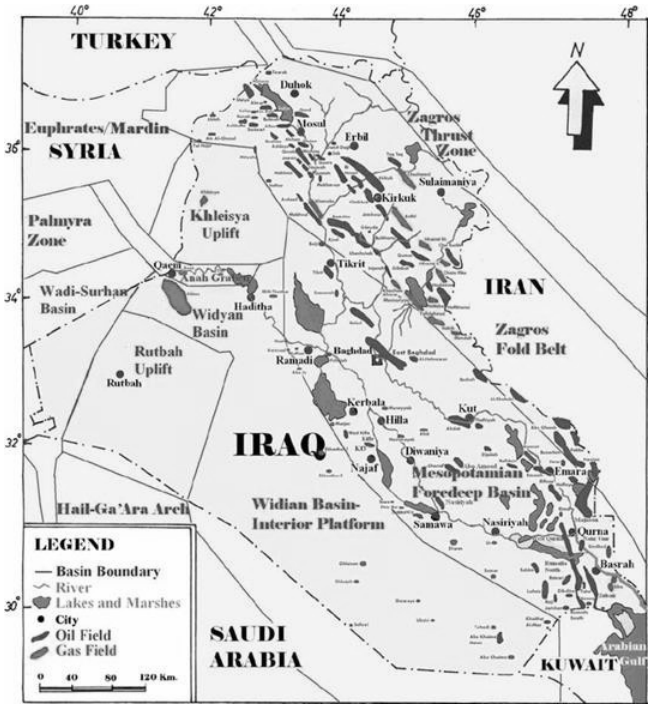


Figure 1: Location map of Iraqi wells, and oil fields.[7]

2. Volumes of Gas Flaring

Iraq is developing its facilities for gas collection and distribution, as Iraq concentrates on rising gas production by cutting flaring volumes. Iraq fosters collaborations with IOCs to meet its targets for gas output and reverse the current trends in the process, as sanctions, regional tensions and the lack of incentives for foreign operators have been derailed.

From early 2018 to late 2022, the additional output is expected to be mainly used for domestic fuel and power consumption in Iraq, raising its gas production of 23 mcm (812 mcf) per day. The production is expected to increase further by 2022 with the rise of the giant Miran-Bina Bawi project Figure (2).

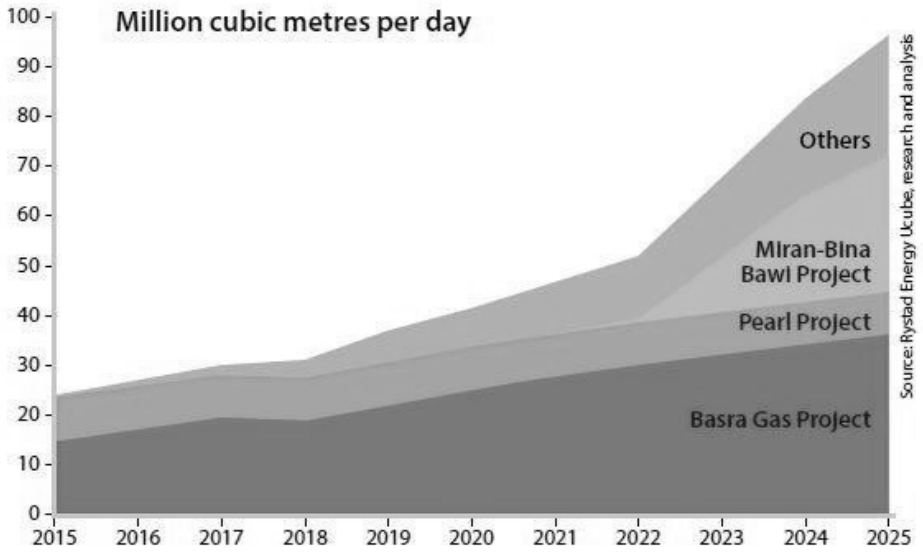


Figure 2: Iraq gas output, split between major projects; Basra Gas Project, Pearl and the Miran-Bina Bawl Project. Iraq gas production.

3. Calculations of volumes

Preliminary statistics indicate that by December 2018 Iraq's federal oil production amounted to 4.45 mbpd compared to approximately 4.2 mbpd in early January 2018. The KRI output increased by approximately 105 kbpd from 360 kbpd in January to 465 kbpd by December. Combined, Iraq has 4,915 mbpd operating capacity with 185 kbpd under-utilized. With respect to exports Figure (3), the SOMO has ended the year to hit a 3,726 mbpd level with a 3,63 mbpd sold by Ceyhan and Trucking17 through Basra's southern ports and by Turkey, throughout the rest.

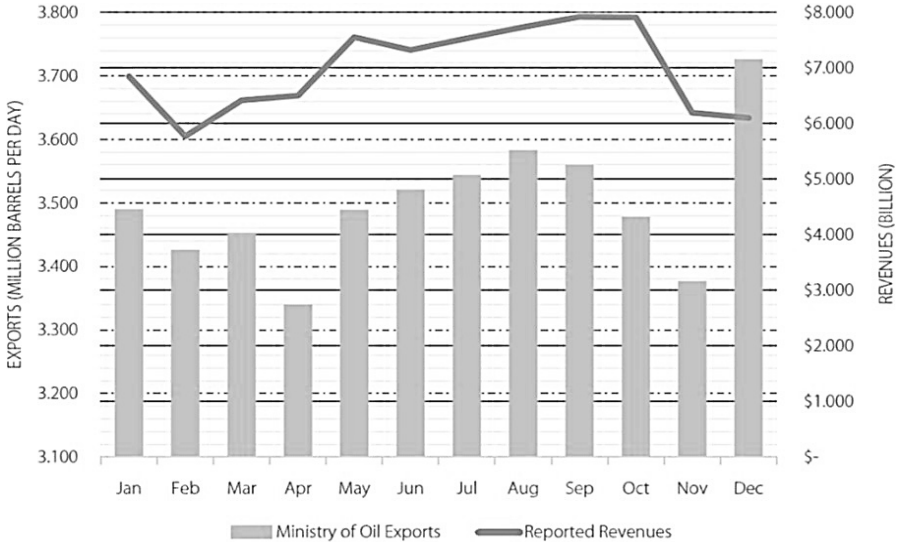


Figure 3: Iraq's 2018 Oil Exports & Revenues[8]

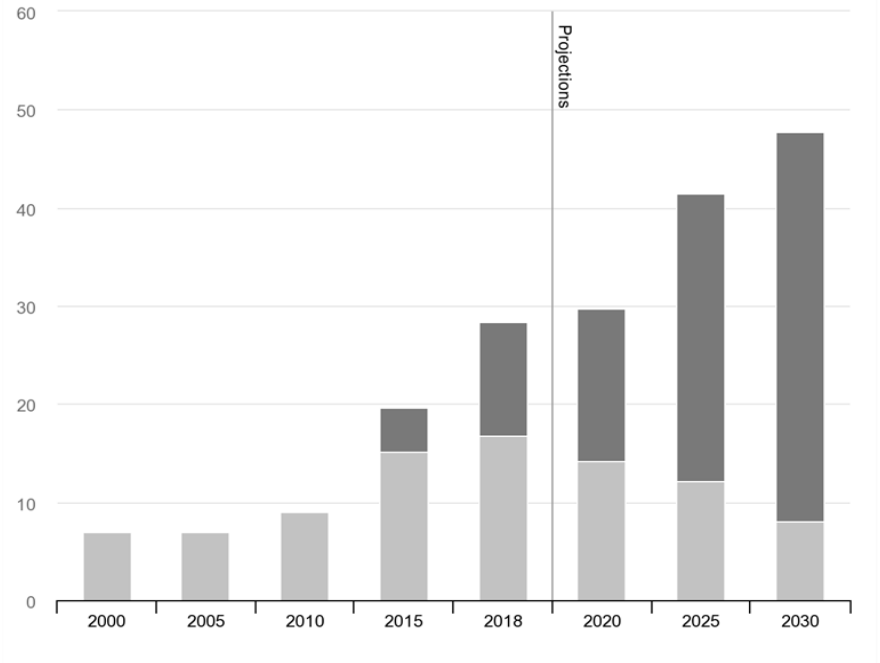


Figure 4: Iraq, 2000-2030 natural gas produce and flaring

The volume of associated gas generated alongside the rise in oil output has also increased as shown in Figure (4). However, this gas has not kept pace for capture and processing. The failure to use its gas resources means the country has increased the gas deficit and now Iraq relies on Iranian imports to meet rising demand, added a variety of vulnerabilities to the energy infrastructure in Iraq. Payment problems, for example, led Iran last summer to cut exporters and considerably worsen Iraq's electricity scarcity during peak seasonal demand.

4. Environmental Issues of the Industry

Technological processes of the main and auxiliary production of hydrocarbon raw materials have an impact on the natural environment: atmospheric air, water and land resources due to emissions and discharges of pollutants, the formation and treatment of production and consumption waste Figure (5). The main types of environmental impact include:

1. Emissions of pollutants into the atmosphere at gas fields and gas collection points;
2. Emissions of pollutants into the atmosphere during technological maintenance of gas control stations and gas pipelines;
3. Discharges of waste water into evaporation ponds and filtration fields;
4. Possible pollution of the environment with wastes of the main and auxiliary production.

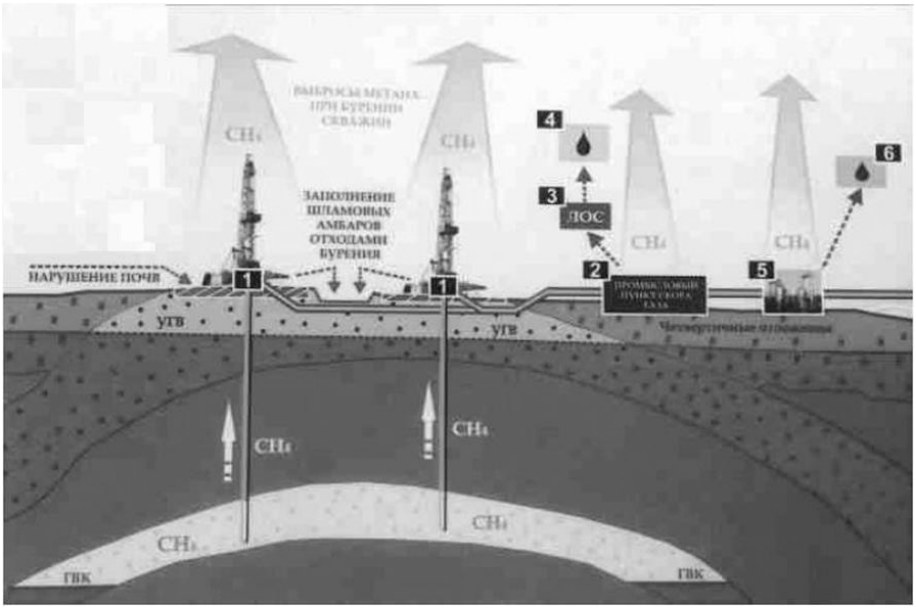


Figure 5: Environmental issues of the industry

5. Flare Installations

The flare unit is designed to burn off-grade gas and gas condensate mixtures formed during the operation of equipment or emergency discharges. At gas and condensate production facilities, horizontal and high-altitude flare units are used to burn hydrocarbon mixtures and waste gases after blowing (during commissioning, after workover) of wells in gas fields, gas pipelines after capital repairs, process equipment at gas fields and gas treatment plants, dust collectors at compressor stations.

Horizontal flare units are located in special earthen pits at a distance of about 100 m from the wellhead. The well spout is connected to the barn by a horizontal pipe, the end of which is brought out above the ground. Combustion occurs in an open torch from the ground surface, and pollutants are dispersed in the vertical

part of the torch. The temperature of the combustion products is 900-2000 ° C, depending on the combustion mixture, the excess air ratio is 1- 1.5.

In high-altitude flare units (high and low pressure), the flare burners are located at the top of the flare stack. Only combustible components rise along the flare stack, and combustion occurs in the atmosphere above the flare stack head.

When operating hydrogen sulfide-containing deposits, it is preferable to use high-altitude flare units to increase the degree of dispersion of pollutants. The design of the flare units ensures soot-free combustion of gas mixtures under certain conditions for gas outflow. Since the mode of gas discharge into flare systems can vary in a wide range, the composition of emissions (soot / soot combustion) must be determined by calculation for each specific case (type of routine operation).

Efficient mixing of hydrocarbon mixtures with air is ensured by using a flare burner with the separation of the gas flow into separate streams using a large number of burners or with a large number of outlets.

There are also several types of flare systems:

a) A flare system for the removal of warm humid vapors - the gas is burnt through this flare constantly, periodically or in emergency situations in the following order:
   - Periodic use during maintenance or repair;
   - Use in emergency situations when pressure relief valves are activated or accidental emergency shutdown;

b) Low pressure flare - at an early stage in the operation of the fuel gas system, a separate, independent low pressure flare may be installed, which will be used in the period before the construction of the main flares. Gas discharge to the low pressure flare can be carried out either periodically or in case of emergency. Permanently operating flare units can operate from reservoirs of unstabilized condensate, from a methanol recovery unit, etc .;

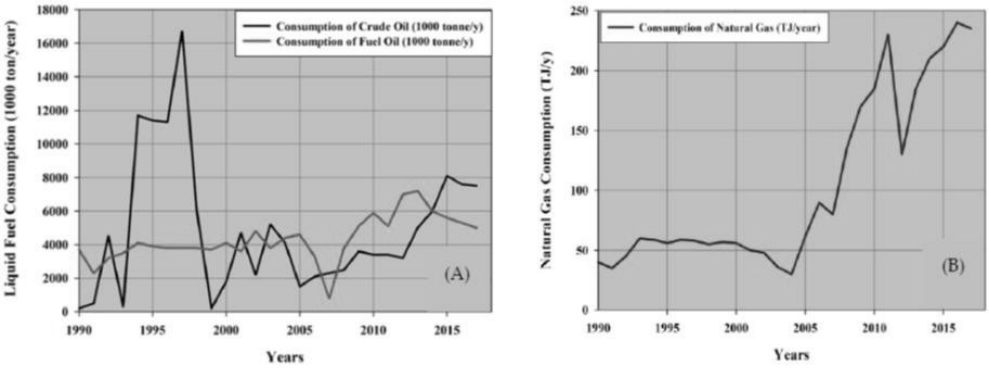
c) Cold / dry gas flare system - the system consists of flare and drainage headers together with associated drainage tanks and flares for burning gases and liquids. Dry gas flare headers collect cold liquids - drains that do not contain water or its vapors;

d) Boil-off flare system - a separate, independent low pressure flare is required for the LNG storage and offloading area due to the very low emission pressure from the system. The capacity of this system must be sufficient to cope with evaporation during operations in the LNG storage area. Vapors that can arise in the event of emergencies (failure of the control valve of the relief valve or overturning of the vessel) will be released to atmosphere through the safety valves of the LNG storage tank.

6. Levels of Environmental Impacts for the Gas Field as a Whole

6.1 Energy Consumption

The volume of energy resources (gas, heat and electricity, water) consumed by a gas producing company for gas production and preparation for transportation largely depends on the reservoir pressure level of the field being developed and the method of gas compression in the declining production mode as shown in Figure (6) a and (6)b. The enterprise strives to extract the maximum possible volume of hydrocarbon raw materials from the fields being developed and to the maximum possible efficiency in the use of energy resources.[9,10]



(a) **(b)**

Figure 6: (a) Crude oil and fuel oil consumption for electricity production thousand tons/ year), (b) electricity generation natural gas consumption (TJ/year)

The most priority area of energy conservation in gas production is the development and implementation of technologies for field gas treatment, which reduces material costs and energy consumption while meeting regulatory requirements for the quality of commercial products

6.2 Emissions to Air

Operational losses of gas and condensate during production and treatment, leading to environmental pollution, are associated with a number of reasons:[11]

- Imperfect systems for collecting and transporting gas and condensate;
- A low degree of sealing of raw materials and commodity tanks;
- Imperfect separation equipment and separation process;
- Breakthroughs of gas pipelines and product pipelines due to accidents and violations of the rules of their technical operation;
- Burning part of the gas in flares, including at compressor stations due to their insufficient capacity.

The greatest damage to the environment is caused by the combustion of gases containing hydrogen sulfide. The main amount of emissions enters the atmosphere during the blowing of wells leaving drilling, after workover and various studies.

The main sources of emissions of pollutants into the air from gas production facilities are shown in Table (1) and Figure (7).

Table 1: Emissions of pollutants from hydrocarbon production facilities.[12]

Naime-Considerations	Units	Weight of pollutants in waste gases before cleaning per ton of manufactured products		Sources of emission mations	Purification method, re The use	Weight of pollutants in waste gases after cleaning per ton of	
					Manufactured products		
		Range	Average			Range	Average
Gas, gas condensate, oil and gas condensate fields							
Nitrogen oxides	kg / t. AD products (year)	0.69-0.71	0.7	Chimneys, torch	-	0.69-0.71	0.7
Carbon monoxide	kg / t. AD products (year)	1.99-2.01	2	Chimneys, torch	-	1.99-2.01	2
Methane	kg / t. AD products (year)	1.40-1.60	1.5	Torch, candle	-	1.40-1.60	1.5
Gas, gas condensate, oil and gas condensate fields containing hydrogen sulfide							
Nitrogen oxides	kg / t. AD products (year)	0.40-0.60	0.5	Chimneys, torch	-	0.40-0.60	0.5
Carbon monoxide	kg / t. AD products (year)	2.40-2.60	2.5	Chimneys, torch	-	2.40-2.60	2.5
Methane	kg / t. AD products (year)	0.40-0.60	0.5	Chimneys	-	0.40-0.60	0.5
sulphur dioxide	kg / t. AD products (year)	25.9-26.1	26	torch	-	25.9-26.1	26

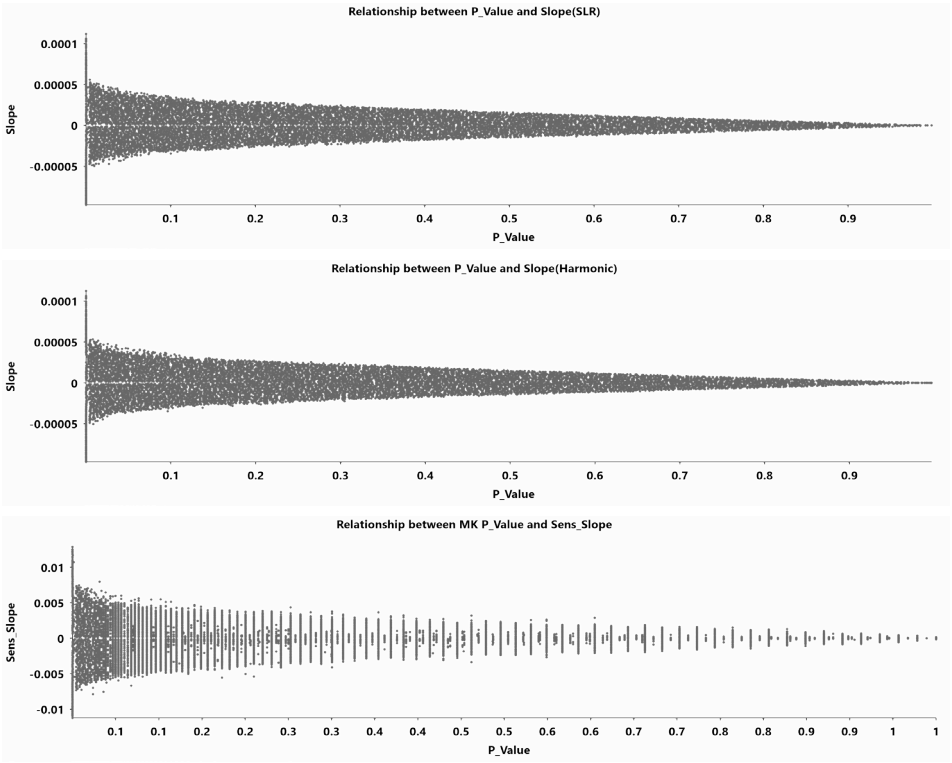


Figure 7: Relationship between P value and Emissions

7. Remote sensing

For the production of crude oil and gas, it is important to introduce the gas flaring activity (GF) at the oil processing plant, which is a regular procedure used either as a disposal or as a protection measure in order to alleviate pressure through the combustion of the gas generated during various industrial activities. [13, 14]. Since the 1990's the regulations on calculating flare gas for the measurement of carbon dioxide (CO2) tax on oil activities in Norway have been enforced, the environmental, social, and economic implications associated with GF are evident[15, 16]. GF is an industry activity which spans almost 300 million tons of CO2 equivalent

emissions annually and makes a significant contribution towards climate change and constitutes significant waste from a valuable non-renouvelable energy resource. Nevertheless, a number of technological, regulatory and/or economical constraints and in some geographic areas are still in place today [17,18].

8. Satellite imagery for identification of fires

The capabilities of sensors to detect fire signals from the earth orbit is focused on fire detection. Both primary signal types are direct (heat and flames) and indirect for this function (burned and smoke surfaces). Direct signals are most often used in the identification of fire while indirect signals are used for post-fire evaluation and management. Most satellite fire detection research studies on forest/biomass fires, because their effects attract significant attention from the research community and the availability of validated fire- hotspot algorithms helps research. The research community

9. Satellite Imager Identification of Gas Flares

Due to its relatively high spatial resolution Landsat data have been chosen for the present study as a suitable sensor. Current fire detection strategies are developed for larger, ephemeral fires that are typically based on near-real time fire monitoring. developed in order to show how the distribution of CO_2 Emission Doses and Refinery Locations are distributed in the Iraq as shown in Figure (8). For detection of gas flares, Landsat data can outline its possible benefits as:

(i) Landsat has a good spatial resolution that is tailored to the accurate detection of high temperature yet low spatial flares;

(ii) the location and almost continuous existence of the flares means the unlikely detriment of the detection process is the relatively low temporal resolution of Landsat images;

(iii) the long Landsat database and the time span of the newly deployed Landsat 8 helps this device to detect shifts of spatiotemporal flaring on a long- term basis.

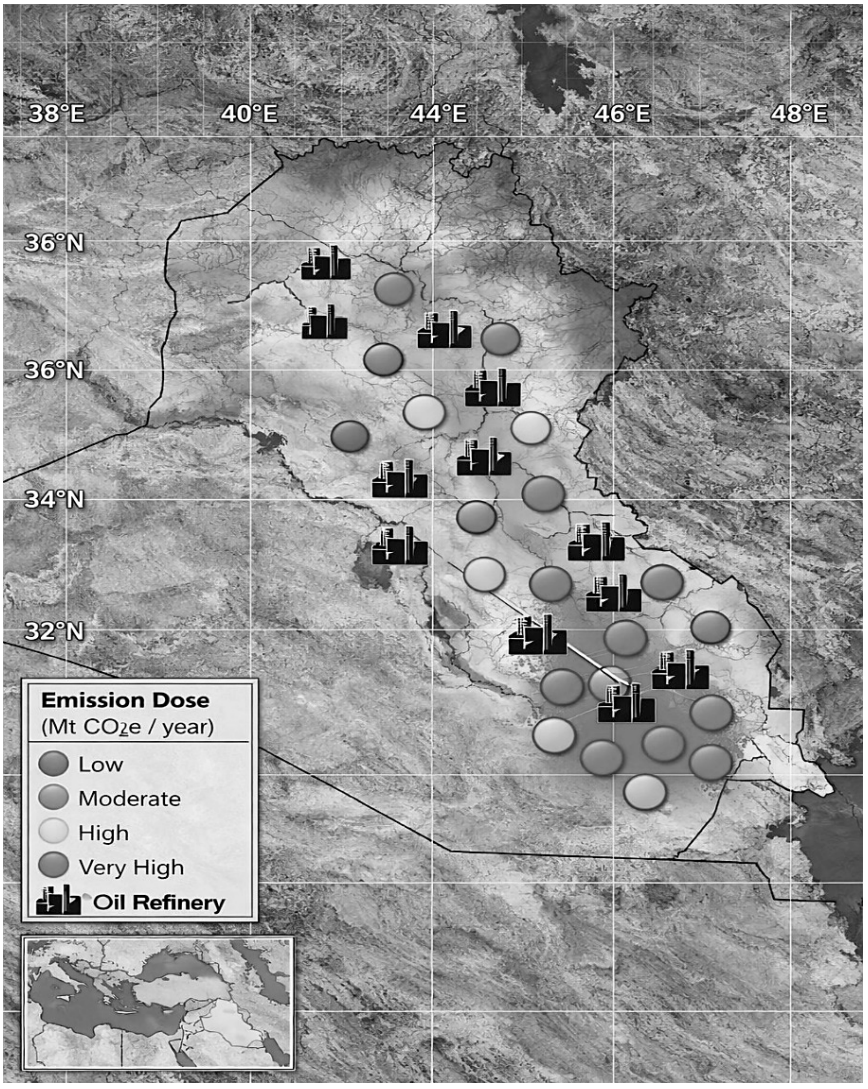


Figure 8: A spatial distribution of CO_2 Emission Doses and Refinery Locations

10. Conclusion

Iraq's natural gas reserves were the twelfth highest in the world at the end of 2017, according to OGJ, at almost 135 billion cubic feet (Tcf). Some three-quarters of Iraq's natural gas reserves are related to oil, the bulk of which are situated in the South's supergiant fields.

Iraq has produced 357 billion cubic feet of dry natural gas (BCF), while a further 18 Bcf have been re-injected in the course of the year.

In 2017, Iraq was ranked as the world's second largest source of flared gas behind the Russian government by 629 Bcf of natural gas. Inadequate pipelines and other midstream facilities have made natural gas flared. In 2017, Iraq consumed dry gas 422 Bcf, some used in the electricity industry.

In June 2017, Rq commenced importing gas from Iran into the power stations near Baghdad, including the stations of Al–Besmaya, Al–Quds, Al– Mansuriyah, and Al–Sadr. Annual imports of natural gas averaged 132 million cubic feet per day in 2017, although this amount may increase as new generation capacity for natural gas is online in Iraq. In 2017, these volumes were increased. Iraq imported an average of 265 MMcf/d between January and May 2018, this doses increase to hort-term (2026–2027): Imports likely to hover around 6–7 billion m^3 per year until new gas-processing plants reach full capacity. Medium-term (2028–2030): Expected decline to < 4 billion m^3 as domestic capture projects expand in Rumaila, West Qurna, and Majnoon. Long-term goal: Transition from net importer to balanced producer, leveraging GIS-based monitoring of flaring reduction and pipeline optimization. The integration of remote sensing and GIS techniques has provided a comprehensive spatial understanding of Iraq's gas flaring patterns and their environmental implications. The satellite-based analysis revealed that flaring intensity and emission doses are highly concentrated in the southern oil fields,

particularly near Basra, Rumaila, and West Qurna, where refinery density and pipeline networks are most extensive. The interpolation of emission doses indicates a clear correlation between refinery proximity and elevated CO_2-equivalent emissions, emphasizing the need for improved gas capture and utilization infrastructure.

11. References

[1] M.A. Arefin, M.N. Nabi, M.W. Akram, M.T. Islam, M.W. Chowdhury, A, (2020), Review on Liquefied Natural Gas as Fuels for Dual Fuel Engines: Opportunities, Challenges and Responses, J Energies 13(22) 6127.

[2] Chapoy,, (2004), Phase behaviour in water/hydrocarbon mixtures involved in gas production systems, École Nationale Supérieure des Mines de Paris,.

[3] M. Mosinska, M.I. Szynkowska, P. Mierczynski, (2020) , Oxy-Steam Reforming of Natural Gas on Ni Catalysts—A Minireview, J Catalysts 10(8) 896.

[4] K.S. Pedersen, P.L. Christensen, J.A. Shaikh,)2014(, Phase behavior of petroleum reservoir fluids, CRC press.

[5] Kralova, J. Sjöblom, G. Øye, S. Simon, B.A. Grimes, K. Paso, (2011), Heavy crude oils/particle stabilized emulsions, J Advances in colloid interface science 169(2) 106-127.

[6] Kühne, O., Koegst, L., Berr, K. , (2024), Crude Oil and Natural Gas—Formation, Extraction, Distribution, and Processing. In: Oilscapes of Louisiana. RaumFragen: Stadt – Region – Landschaft. Springer VS, Wiesbaden. https://doi.org/10.1007/978-3-658-43396-3_5

[7] T.K. Al-Ameri, J. Pitman, M.E. Naser, J. Zumberge, H.A.J.A.J.o.G. Al-Haydari, (2011), Programed oil generation of the Zubair Formation, Southern Iraq oil fields: results from Petromod software modeling and geochemical analysis, 4(7-8) 1239-1259.

[8] Y.I. Al-Maleki, Harry & Tollast, (2019), Riding the Oil Markets, Iraq's Economy in Transition.

[9] M. Mohsin, Q. Abbas, J. Zhang, M. Ikram, N. Iqbal, (2019), Integrated effect of energy consumption, economic development, and population growth on CO 2 based environmental degradation: a case of transport sector, J Environmental Science Pollution Research 26(32) 32824-32835.

[10] B.M. Hashim, M.A. Sultan, A. Al Maliki, N. Al-Ansari, (2020) , Estimation of greenhouse gases emitted from energy industry (oil refining and electricity generation) in Iraq using IPCC methodology, J Atmosphere 11(6) 662.

[11] Rao, C. S. (2007). Environmental pollution control engineering. New Age International. MLA (9th edition):

[12] K. Scarfone, M. Souppaya, A. Cody, A. Orebaugh, (2008), Technical guide to information security testing and assessment, J NIST Special Publication 800(115) 2-25.

[13] E.A. Emam, (2016), Environmental pollution and measurement of gas flaring, International Journal of Innovative Science Engineering and Technology 2 ,252-262.

[14] U.S. Department of Energy, Office of Fossil Energy. (2019). Natural gas flaring and venting: State and federal regulatory overview, trends, and impacts. Washington, DC: U.S. Department of Energy, Office of Oil and Natural Gas.

[15] M. Soltanieh, A. Zohrabian, M.J. Gholipour, E. Kalnay, (2016), A review of global gas flaring and venting and impact on the environment: Case study of Iran, International Journal of Greenhouse Gas Control 49, 488-509.

[16] E.M. Johannes, L.C. Zulu, E. Kalipeni, (2015), Oil discovery in Turkana County, Kenya: a source of conflict or development? African Geographical Review 34(2) 142-164.

[17] C.D. Elvidge, M.D. Bazilian, M. Zhizhin, T. Ghosh, K. Baugh, F.-C. Hsu, (2018), The potential role of natural gas flaring in meeting greenhouse gas mitigation targets, Energy strategy reviews 20 ,156-162.

[18] B. Gervet, (2007), Gas flaring emission contributes to global warming, Renewable Energy Research Group, Lulea University of Technology, Lulea, Sweden, Renewable Energy Research Group, Lulea University of Technology, Lulea, Sweden.

Performance Evaluation of Multi-Stage Compressors and Their Impact on Energy Conversion System Efficiency 'A Field Study at Thi-Qar Oil Company' / Original Article

Researcher Taha Chyad Thajeel

Chief Engineered, Ministry of Oil, Thi-Qar Oil Company, Thi-Qar/ Iraq
e-mail: taha629626@gmail.com

تقييم أداء الضواغط متعددة المراحل وتأثيرها في كفاءة أنظمة تحويل الطاقة دراسة ميدانية في شركة نفط ذي قار

الباحث طه جياد ثجيل

رئيس مهندسين، وزارة النفط، شركة نفط ذي قار، ذي قار\العراق

Abstract

This paper is a field-based performance assessment of a multi-stage natural gas compressor that is operating in realistic industrial conditions. The data is analyzed using the real operating data taken in a particular oil and gas plant and, therefore, the actual operating data is used to determine the trends of thermodynamic behavior, energy consumption, and efficiency, instead of theoretical or simulation-based assumptions. The analysis was performed phase by phase in order to examine the distribution of pressure ratios, temperature increase and power consumption in the compressor stages. The findings demonstrate that there is a significant difference between the power of compressing calculated and measured electrical which is an important data of mechanical and electrical loss as compared to the total energy power consumption. The efficiency analysis has shown that compressor is most efficient at intermediate values of operation but deteriorates at high pressure ratios because it has more irreversibility's and less effective intercooling. Stage-wise analysis shows that the first compression stage occupies the most power consumption and is the one with the worst thermal loading.

The performance analysis of intercoolers establishes that it is a key component in compressing work and managing discharge temperatures and downstream intercoolers has more performance and is more effective than the first one. The results have highlighted the importance of making an overall compressor efficient; this calls on an integrated approach that incorporates thermodynamic optimization, reduction in losses, and proper thermal management. The results of this research offer implications to practical operations optimization, maintenance and energy efficiency improvement in multi-stage gas compression system in industries.

Keywords: Multi-stage gas compressor; Energy efficiency; Intercooling effectiveness; Power losses; Stage-wise analysis; Industrial operation

المستخلص

تقدم هذه الورقة تقييماً ميدانياً لأداء ضاغط غاز طبيعي متعدد المراحل يعمل في ظروف صناعية واقعية. يعتمد التحليل على بيانات تشغيل فعلية جُمعت من أحد معامل النفط والغاز، مما أتاح دراسة سلوك النظام الحراري واستهلاك الطاقة واتجاهات الكفاءة بالاستناد إلى بيانات تشغيل حقيقية بدلاً من الافتراضات النظرية أو النماذج المحاكاة. وقد أُجري التحليل مرحلةً بمرحلة بهدف فحص توزيع نسب الانضغاط، وارتفاع درجات الحرارة، واستهلاك القدرة عبر مراحل الضاغط المختلفة.

أظهرت النتائج وجود فرق ملحوظ بين قدرة الانضغاط المحسوبة والقدرة الكهربائية المقاسة، مما يشير إلى مساهمة كبيرة للخسائر الميكانيكية والكهربائية في إجمالي استهلاك الطاقة. كما بيّن تحليل الكفاءة أن الضاغط يحقق أفضل أداء عند ظروف تشغيل متوسطة، بينما تتدهور كفاءته عند نسب انضغاط مرتفعة نتيجة زيادة اللاتعاكسية الحرارية وضعف فعالية التبريد البيني. وأظهر التحليل المرحلي أن المرحلة الأولى من الانضغاط تستحوذ على أعلى نسبة من استهلاك القدرة، كما تعاني من أكبر إجهاد حراري مقارنة بالمراحل اللاحقة.

كما أثبت تحليل أداء المبادلات البينية (Intercoolers) أنها عنصر أساسي في تقليل شغل الانضغاط والتحكم في درجات حرارة الطرد، حيث تبين أن المبادلات البينية في المراحل اللاحقة أكثر كفاءة وفعالية من المبادِل الأول. وتؤكد النتائج أهمية اعتماد منهج متكامل لتحسين الكفاءة الكلية للضاغط، يجمع بين التحسين الديناميكي الحراري، وتقليل الخسائر، والإدارة الحرارية السليمة. وتوفر مخرجات هذا البحث دلالات عملية مهمة لتحسين التشغيل، وتخطيط الصيانة، ورفع كفاءة الطاقة في أنظمة ضغط الغاز متعددة المراحل في التطبيقات الصناعية.

الكلمات المفتاحية: ضاغط غاز متعدد المراحل؛ كفاءة الطاقة؛ فعالية التبريد البيني؛ خسائر القدرة؛ التحليل المرحلي؛ التشغيل الصناعي.

1. Introduction

1.1 Background and Research Context

The energy conversion systems are critical in the present-day industrial operation and especially in the oil and gas industry in which huge quantities of both mechanical and thermal energy are always converted to aid the production, processing and transportation processes. Multi-stage compressors rank as some of the most important elements in such systems, as they are commonly used to boost the pressure of gases to the necessary working levels without interrupting the processes or endangering the safety. Compressors used in petroleum plants cannot be avoided in gas collection, treatment, storage and reinjection, and as a result, their effectiveness is directly related to the overall performance of the plant and the energy cost [1],[2].

The oil and gas industry is the foundation of the national economy in Iraq, and its facilities are subject to harsh conditions with high ambient temperature, changing loads and deteriorating infrastructure. These states have a pronounced impact on the efficiency of equipment of energy conversion, in particular the multi-stage compressors that in many cases do not work at the optimal design points. This has led to a high energy loss, low efficiency and high operation costs becoming a common issue in the Iraqi oil fields and processing units [3],[4].

1.2 Problem Statement

Multi-stage compressors are some of the most energy consuming aspects in oil and gas energy conversion systems and they are used continuously to satisfy the pressure of the process. Although they are widely used and offer possible theoretical efficiency benefits, when implemented in practice in real industrial situations they are often far out of optimum design conditions. These deviations lead to high energy consumption, less thermodynamic efficiency, and high operational and maintenance costs.

Multi-stage compressors in most oil production and processing plant including those used by Thi-Qar Oil Company under the Iraqi Ministry of Oil are exposed to variable operating conditions which include variable inlet pressure, high ambient temperature, variable mass flow rate, and extended constant operation. These conditions eventually cause a degradation of the performance due to intercooler inefficiencies, internal leakage, mechanical losses, fouling and an incorrect stage pressure distribution. Nevertheless, field-based systematic performance assessment is not common to measure these effects with the actual operational data.

The other major issue of concern is the paucity of detailed performance evaluation studies that are based on real-world measurements of industry and not on idealized model or data that are supplied by the manufacturers. Consequently, most compressors are still running with unknown efficiency factors, and engineers and decision-makers find it challenging to detect the presence of energy losses, determine the level of effectiveness, and apply specific optimization strategies. This disparity between the theoretical performance expectations and actual operation behavior is a very strong barrier to enhancing the energy efficiency in Iraqi oil facilities.

Moreover, the lack of unified analytical systems that synthesize field measurements in conjunction with thermodynamic and energy conversion models restrict the possibility of assessing the performance of compressors in a holistic and dependable way. In the absence of such assessment, possibilities to minimize power consumption, to increase system reliability, and overall energy conversion efficiency are still to be explored.

In this respect therefore there is an urgent demand of a field based, detailed study to determine the performance of multi-stage compressors based on realistic operational data acquired by real-life work at actual oil field installations. The fulfillment of this need will allow gaining a better insight into the work of compressors

under the conditions of actual operation, which will help make informed decisions in engineering and help enhance the energy efficiency and sustainability of energy conversion systems in the Iraqi oil and gas industry.

1.3 Role of Multi-Stage Compressors in Energy Conversion Systems

Multi-stage compressors are specially made to enhance compression efficiency as the overall pressure ratio is broken down into multiple stages frequently accompanied by intercooling to minimize the amount of compression work. Such a design reduces the thermal stresses, increases the mechanical reliability and the general energy efficiency of the compression process. In thermodynamic terms, multi-stage compression is trained to an ideal isothermal process thus, decreasing the amount of specific work input and enhancing system performance [5],[6].

There is direct correlation between power consumption, thermal efficiency and system reliability in energy conversion systems and the performance of multi-stage compressors.

Even minor changes in the working conditions like inlet temperature, pressure ratio, or mass flow rate can cause large proportions of energy usage and mechanical wear. Thus, consistent performance assessment using actual operational data would be necessary to achieve effective energy use and sustainable functioning [7].

1.4 Industrial Challenges in Oil and Gas Compression Systems

Although the theoretical benefits of multi-stage compression are obvious, in reality, the implementation in industry is associated with loss of performance through fouling, leakage, poor balancing of the stages, poor cooling, and inappropriate control policies. Most oil and gas plants have compressors that have been run continuously under varying load conditions without a planned performance evaluation, thus causing slow and gradual performance degradation with time [8],[9].

These problems are also worsened by the lack of access to advanced monitoring systems and periodic performance audits in the context of the Iraqi oil companies, such as Thi-Qar Oil Company of the Iraqi Ministry of Oil. Therefore, compressors are often used with higher power requirements and lower efficiency, and this affects the total on-energy conversion chain negatively and leads to an increase in fuel usage and operating costs [10].

1.5 Importance of Field-Based Performance Evaluation

The majority of research done on compression performance is based on numerical models or controlled experiments in laboratory settings, which, though useful, might not accurately represent the reality of industrial settings. A more realistic measurement of system behavior and power efficiency is based on the actual measurements in operational conditions, including pressure, temperature, flow rate, and power consumption, as created through field-based performance evaluation [11],[12].

Performance evaluation based on real field or realistic field data enables the engineers to detect the energy losses, measure the departure of efficiency due to the real field conditions and offer achievable optimization strategies based on the real operating conditions. This is especially applicable in the case of energy intensive industries where even insignificant efficiency gains can lead to significant energy savings and emission cuts [13].

1.6 Motivation and Contribution of the Present Study

This paper will be inspired by the increasing demand to improve the efficiency of the energy consumption of the Iraqi oil outlets through the development of a multi-stage compressor performance in an energy conversion facility in Thi Qar Oil

Company. The study focuses on practical evaluation based on realistic operational data, analytical performance indicators and thermodynamic modeling that represent the reality in the industry.

The main value of this work is that it has allowed closing the gap between the theoretical analysis of compressor performance and real-life operation in the industry. The combination of the measured operational parameters with the developed energy conversion models gives the study important information on the energy losses and efficiency trends as well as the areas of improvement that are unique to the oil and gas compression systems that are used in Iraq. The results of the study should help in promoting the energy management plans in the Iraqi Ministry of Oil and make a contribution to the existing literature in applied energy conversion engineering [14],[15].

1.7 Research Objectives and Scope of the Study

This research is mainly aimed at performing a critical field-based performance assessment of multi stage compressors that will be used in an energy conversion system at Thi-Qar Oil Company. The study will be conducted to measure the real and realistic field performances of the compressors by investigating the actual and realistic field data such as pressure, temperature, mass flow rate and power consumption of various compression stages.

In particular, the research would aim to determine how the multi-stage compression properties affect the overall efficiency of energy conversion, what are the primary locations of thermodynamic and mechanical losses, and how deviations in the performance under design are expected to be. The study will achieve a realistic description of the compressor behavior in real operating conditions in industries by combining the operational data and known thermodynamic and energy balance models.

This paper is restricted to multi-stage gas compressors, which are employed in the oil and gas tasks in Thi-Qar Oil Company facilities. The analysis is too steady-state operating conditions that are usually experienced in normal operation of the plant and not in cases of transient startup or shutdown. The results of the study are supposed to assist in the real-world decision-making concerning the energy management, operational optimization and performance improvement strategies concerning the energy conversion systems in the Iraqi oil sector.

2. Literature Review and Theoretical Framework

2.1 Fundamentals of Multi-Stage Gas Compression

In the gas and oil industry, one of the main processes used in the process of energy conversion is the process of gas compression, which allows transporting, processing and utilizing various gaseous flows to the desired pressure levels. To achieve high overall pressure ratios, single-stage compression is thermodynamically inefficient with excessively high temperature and high specific work. As a result, multi-stage gas compression has become a popular industrial standard in order to enhance efficiency and operational reliability [16],[17].

In multi-stage compressors the total pressure ratio is shared among stages usually with intercooling between stages. This design lowers the average compression temperature and approaches the process to ideal isothermal compression which theoretically needs the smallest possible work input [1]. These constructions are particularly vital in incessant industrial uses in which energy savings and part longevity are indispensable.

2.2 Thermodynamic Principles of Multi-Stage Compression

Thermodynamically, the actual gas compression processes are deviating ideal isothermal and isentropic processes because of irreversibility like friction, heat

transfer and mechanical losses. Multi-stage compressor behavior is consequently controlled by the magnitude of stage pressure ratios, the effectiveness of intercoolers, the properties of gases and the operating conditions [12].

Experiments have revealed that at a certain overall pressure ratio, the even distribution of the pressure among the stages reduces the total compression work, especially in cases where effective intercooling is used [9]. Imperfection in optimum pressure distribution causes higher power consumption and low overall efficiency. These principles of thermodynamics are what give the theoretical foundation of assessing the compressor performance based on the field-measured data, and not idealized assumptions.

2.3 Performance Indicators and Efficiency Definitions

The measurement of compressor effectiveness depends on various indicators that give a quantitative measure of the conversion of energy. It is widely known that isentropic efficiency is used to measure the actual work of compressor and the ideal isentropic work of a particular pressure ratio, which is a direct measure of thermodynamic performance [7]. In the case of multi-stage compressors, polytropic efficiency can be desirable due to its ability to consider incremental compression behavior due to the stage-to-stage behavior and also its ability to provide better accuracy in performance characterization.

Electrical power consumption of the compressor drive system is also regarded as a critical indicator in industrial practice since it is a direct indicator to the amount of energy demand on the facility. By including the thermodynamic efficiency metrics with the electrical power measurements, it is possible to have a complete analysis of the compressor energy performance when it is operating under actual working conditions [18].

2.4 Review of Previous Industrial and Field-Based Studies

Compressor performance has been a topic of research that has seen considerable literature being conducted using analytical modeling, numerical simulation and experimental research. Various works are devoted to the design optimization and the performance prediction in the controlled or idealized situations. Although these methods prove useful to explain basic behavior, they are not always applicable to the current industrial systems.[5]

Field-based research indicates the significance of employing actual operational data to contain the impact of aging, fouling, ambient factors and variation of loads. The decrease in efficiency of industrial compressors with time has been reported by various researchers and causes a corresponding increase in the amount of energy consumed unless remedial changes are enacted19]]. These results imply the importance of periodical performance review via real measurements and not just the specifications of the manufacturer.

2.5 Energy Efficiency Challenges in Oil and Gas Compression Systems

The gas compression units are some of the most energy consuming units within the oil and gas plants with a large proportion of the overall electrical power consumed by the gas compression units. The inefficiencies can be caused by poor intercooler efficiency, internal leakage, mechanical wear, and inefficient manner of operation control [20].

There are no systematic performance monitoring frameworks that are applied in most of the facilities especially in the developing oil producing regions. Consequently, compressors can run longer and at low-efficiency levels, and unnecessary loss of energy and higher operating costs will have taken place. To handle these issues, the methodology of evaluation should be designed to encompass both the field-based measurements and the theory of thermodynamics [21].

2.6 Research Gap and Positioning of the Present Study

The literature reviewed suggests that even though there is a lot of literature discussing compressor performance, only a few studies are based on a detailed field information of functioning oil facilities. In addition, there is a lack of published studies on the compressor systems in the Iraqi oil fields. This also limits the capacity to formulate localized and measurement-based research to create specific energy efficiency improvement solutions that would be relevant to the regional operating conditions.

To fill this gap, the present paper employs the well-established ideas of thermodynamic and performance evaluation in real and realistic data of operations of a multi-stage gas compressor system within one of the oil companies in Iraq. The paper focuses on field-based analysis, which makes it a presentable input into the energy efficiency improvement of oil and gas energy conversion systems.

3. Methodology

3.1 Description of the Study Site and Industrial System

This research will use a methodological approach of a field study to determine the functioning of multi-stage gas compressors as part of a real energy conversion system in Thi-Qar Oil Company, which is under the control of the Iraqi Ministry of Oil. The chosen compressors belong to gas handling and processing infrastructure to assist in the oil production processes, in which compressed gas is necessary in the transportation, treatment, and auxiliary energy conversion processes.

The compression system under investigation is the multi-stage centrifugal gas compressor with an electric motor that works under continuous or semi-continuous load conditions. The compressor will be designed to provide the necessary discharge pressure by successive compression stages using cooling between the stages to

lower the gas temperature and compression work. This design is a common industrial arrangement that is commonly adopted in oil and gas plantations in Iraq.

The compressor system has an operating environment that is part of high ambient temperature, variable inlet gas conditions, and changeable mass flow rates, which affect compressor performance and energy efficiency as illustrated in Figure (1) below. Such environments render the selected system an exemplary case study to analyzing the actual behavior of compressors in Iraqi oil sector systems of converting the energy in the form of energy conversion systems as illustrated in Figure (1) below.

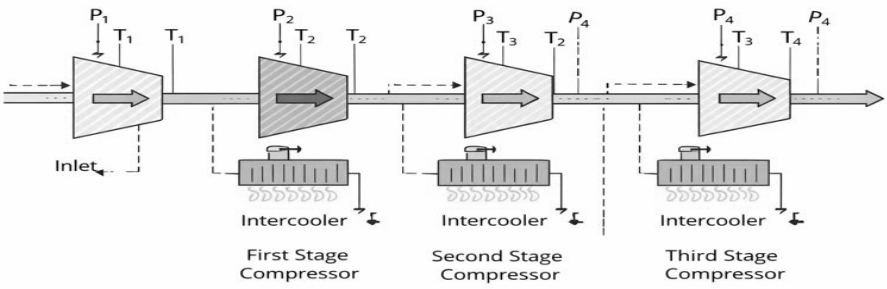


Figure 1: Schematic diagram of a multi-stage gas compression system

3.2 Compressor Configuration and Operating Characteristics

The compressor under study is a multi-stage centrifugal compressor which is intended to compress natural gas by means of a set of compression stages, each providing a fraction of the total pressure increase. Intercoolers are placed between stages to reduce the temperature of the gas intake into the next stage, which results in increased thermodynamic efficiency as well as less mechanical strain to the

compressor parts. Compression stages are run at given inlet pressure and temperature and outlet pressure and temperature conditions, and all these depend on the load on the system, the properties of the gases, and the effectiveness of the cooling. The compressor works with a stipulated range of pressure ratios that the manufacturer advises but more often than not it is rarely run with the desired operating conditions as process demand and ambient conditions change.

The compressor motor providing the compressor with the mechanical force needed to compress the gas is the electric motor, and its electrical power consumption is a considerable fraction of the overall energy needs of the compression system. As a result, the currently operating compressor performance can only be evaluated by measuring the thermodynamic parameters, as well as electricity power consumption, in combination, that is, under the actual working conditions, see Table 1.

Table 1: Measurement Instruments and Accuracy Specifications

Measurement Parameter	Instrument Type	Measurement Range	Accuracy	Location of Measurement
Gas Pressure	Digital Pressure Transmitter	0 – 100 bar	±0.25% of full scale	Compressor inlet and outlet of each stage
Gas Temperature	RTD / Thermocouple Sensor	−20 – 250 °C	±0.5 °C	Before and after each compression stage and intercooler
Gas Mass Flow Rate	Ultrasonic Flow Meter	0 – 5000 Nm^3/h	±1.0% of reading	Main gas pipeline upstream of compressor
Electrical Power	Three-Phase Power Analyzer	0 – 500 kW	±0.5% of reading	Motor control center (MCC)
Rotational Speed	Digital Tachometer	0 – 10,000 rpm	±0.1% of reading	Compressor shaft
Ambient Temperature	Digital Temperature Sensor	−10 – 60 °C	±0.5 °C	Compressor hall / outdoor environment

3.3 Data Collection Strategy and Measurement Philosophy

To make the performance of compressors more realistic and reliable the study uses the operational information that is gathered as a direct field measure and monitoring the plants. The philosophy of measurement adopted is grounded on the measurement of steady-state operating conditions that indicate normal working operations of the industry, not the transient conditions like when starting up or shutting down. The main operating parameters that were measured during the measurement campaign are inlet and outlet pressures of each stage of compression, temperatures of gas prior and after the compression stage, mass flow rate and electrical power of the driving motor. They are the basic parameters that will be used as the inputs in assessing the efficiency of compressor, specific compression work, as well as the general performance of energy conversion.

The measurements were made under several operating periods in order to consider the changes in load and the environmental conditions. The analysis was then carried out on averaged values in order to reduce the effects of short-term transient effects and measurement noise so that the resulting performance indicators would reflect the typical behavior of the compressor.

3.4 Instrumentation and Measurement Devices

Multi-stage compressors necessitate first rate performance measurement of the essential thermodynamic and operational parameters. In this experiment, a series of industrial grade instrumentation has been used in order to gather both real and realistic field data reflecting the real working condition of the compression system. The choice of the measurement devices was due to their aptitude in harsh oil-field conditions, precision of the measurements, and their integration with the current system of the plant monitoring.

The key parameters that shall be measured are the gas pressure, temperature, mass flow rate and electrical power consumption. Calibrated pressure transmitters were placed along the flow of the gas to measure the pressure at the inlet and outlet of each compression stage. The temperature of the system was measured both prior to and following the compression stage and intercooler with resistance temperature detectors (RTDs) and thermocouples, and had sufficient accuracy throughout the projected operating temperature.

The rate of gas mass flow was measured on an inline flow meter fitted on the gas pipeline in the position that is located ahead of the compressor. To determine the electrical power consumption of the compressor drive motor, a three-phase power analyzer was attached to the motor control center. The required data was achieved through these measurements needed to assess the efficiency of the compressors, specific compression work and the overall energy performance, see Figure 2.

Figure 2: Typical field measurement instruments used in multi-stage compressor systems

3.5 Instrumentation and Measurement Devices

Multi-stage compressors necessitate first rate performance measurement of the essential thermodynamic and operational parameters. In this experiment, a series of industrial grade instrumentation has been used in order to gather both real and realistic field data reflecting the real working condition of the compression system. The choice of the measurement devices was due to their aptitude in harsh oil-field conditions, precision of the measurements, and their integration with the current system of the plant monitoring.

The key parameters that shall be measured are the gas pressure, temperature, mass flow rate and electrical power consumption. Calibrated pressure transmitters were placed along the flow of the gas to measure the pressure at the inlet and outlet of each compression stage. The temperature of the system was measured both prior to and following the compression stage and intercooler with resistance temperature detectors (RTDs) and thermocouples, and had sufficient accuracy throughout the projected operating temperature.

The rate of gas mass flow was measured on an inline flow meter fitted on the gas pipeline in the position that is located ahead of the compressor. To determine the electrical power consumption of the compressor drive motor, a three-phase power analyzer was attached to the motor control center. These measurements were used to obtain the required data to determine the compressor efficiency, specific compression work, and energy performance.

3.6 Data Processing and Performance Evaluation Approach

The dried field data were analyzed through a regime of structured analysis grounded on the thermodynamic energy balance principles. Compression stage measured pressures and temperatures were taken to measure the actual behavior

of the compression process and draw comparison between the ideal reference processes. The correlation between the electrical power and the ratio between the gas flow and pressure was taken to determine the nature of the energy consumption. The performance assessment was aimed at assessing the deviation of the expected operation behavior, the measurement of the energy losses and the efficiency of every compression stage and the overall compressor system. This strategy facilitates the existence of a clear connection between operational data that is measured and indicators of energy conversion performance.

3.7 Mathematical Modeling of Multi-Stage Compression

The mathematical modeling adopted in this study is based on classical thermodynamic relations for gas compression, combined with measured field data to evaluate the real performance of the multi-stage compressor. The model aims to quantify the specific compression work, stage efficiencies, and overall system efficiency under actual operating conditions.

3.7.1 Specific Compression Work

For each compression stage, the specific compression work is determined using the steady-flow energy equation, assuming negligible changes in kinetic and potential energies. For an ideal isentropic compression process, the specific work is expressed as:

$$w_{is} = \frac{k}{k-1} R T_1 \left[\left(\frac{P_2}{P_1} \right)^{\frac{k-1}{k}} - 1 \right]$$

where:

w_{is} = isentropic specific work (kJ/kg)

k = specific heat ratio of the gas

R = specific gas constant (kJ/kg·K)

T_1 = inlet gas temperature (K)

P_1, P_2 = inlet and outlet pressures of the stage

In real industrial compressors, the actual specific work is higher than the ideal value due to irreversibility's. The actual specific work is determined from measured temperature rise across each stage as:

$$w_{act} = c_p \left(T_2 - T_1 \right)$$

where:

c_p = specific heat at constant pressure (kJ/kg·K)

T_2 = measured outlet gas temperature (K)

3.7.2 Isentropic Efficiency of Each Compression Stage

The isentropic efficiency is used as a primary indicator of compressor thermodynamic performance. For each compression stage, the isentropic efficiency is calculated as:

$$ç_{is} = \frac{w_{is}}{w_{act}}$$

This efficiency reflects the deviation of the real compression process from the ideal isentropic process. Lower values of η_{is} indicate higher irreversibility's, often caused by mechanical losses, internal leakage, or inadequate intercooling.

3.7.3 Poly-tropic Efficiency for Multi-Stage Compression

For multi-stage compressors operating over a wide pressure range, poly-tropic efficiency provides a more accurate representation of compression behavior. The polytropic efficiency is calculated using measured pressures and temperatures according to:

$$ç_{poly} = \frac{k-1}{k} \frac{\ln\left(\frac{P_2}{P_1}\right)}{\ln\left(\frac{T_2}{T_1}\right)}$$

This parameter allows for a refined evaluation of compressor performance by accounting for incremental compression behavior across stages and is particularly useful for comparing performance under varying operating conditions.

3.7.4 Compressor Power Consumption

The actual power consumption of the compressor is determined by correlating the mass flow rate with the actual specific compression work:

$$\dot{W}_{comp} = \dot{m}\, w_{act}$$

where:

W_{comp} = compressor power consumption (kW)

$\dot{m}$ = measured gas mass flow rate (kg/s)

Electrical power measurements obtained from the motor power analyzer are used to validate the calculated compressor power and assess mechanical and electrical losses within the drive system.

3.7.5 Overall Energy Efficiency of the Compression System

The overall energy efficiency of the compression system is defined as the ratio between the useful compression energy and the electrical energy supplied to the motor:

$$ç_{overall} = \frac{\dot{m}\, w_{is}}{P_{elec}}$$

where:

P_{elec} = measured electrical power input to the motor (kW)

This efficiency provides an integrated indicator of system-level performance, incorporating thermodynamic, mechanical, and electrical effects.

3.8 Performance Evaluation Criteria and Analysis Procedure

In this research, the set of performance evaluation steps has been used to transform the acquired field measurements into meaningful thermodynamic and energy efficiency parameters. To ensure that the evaluation criteria would capture both the stage-level behavior and system-level energy conversion performance, the evaluation criteria had been chosen so as to be able to identify the inefficiencies and quantify the energy losses under real operating conditions.

3.8.1 Selection of Operating Points and Steady-State Filtering

In order to get a normal industrial operation, the operating points were only picked when the compressor system was running in steady-state conditions. Regular filtering was done by ensuring that the important variables such as inlet pressure, discharge pressure, mass flow rate, and electrical power were within a narrow band of variation during a specified sampling period. The data sets with quick changes observed as a result of control measures, load changes, or process interference were avoided to avoid biased performance measures.

3.8.2 Stage-by-Stage Performance Computation

Performance computations were done stage-by-stage, at each of the selected operating points. The experimented inlet and outlet pressure and temperature at every stage were utilized in determining the pressure ratio per stage, the actual

temperature rise and the specific work. The stage poly-tropic efficiency and isentropic efficiency were then determined using the thermodynamic relations that were explained in Section 3.7.

This level of assessment is vital since the multi-stage compressors can result in an uneven performance of each stage depending on the different intercooler, internal leakage, and mechanical wear. A direct diagnostic pathway is given by the identification of stages with abnormal efficiency values which offer the improvement actions.

Table 2 contains the detailed stage-by-stage thermodynamic findings in terms of pressure levels, temperature changes, specific compression work and efficiency indicators of each stage of compression.

3.8.3 Power Balance and Consistency Check

The consistency of the compressor power was checked by comparing the calculated compressor power.

$\overset{\ddot{y}}{W}_{comp} = \dot{m}\, w_{act}$ with actual electrical power input P_{elec} of the motor power analyzer measured. The variation between these values had been estimated as a measure of mechanical losses, drive efficiency and measurement uncertainty.

The purpose of this comparison is twofold (i) to prove the validity of the gathered information, and (ii) to measure the degree of losses that occur beyond thermodynamic irreversibility's within the compressor, including those associated with motors and coupling, as are illustrated in Figure 3.

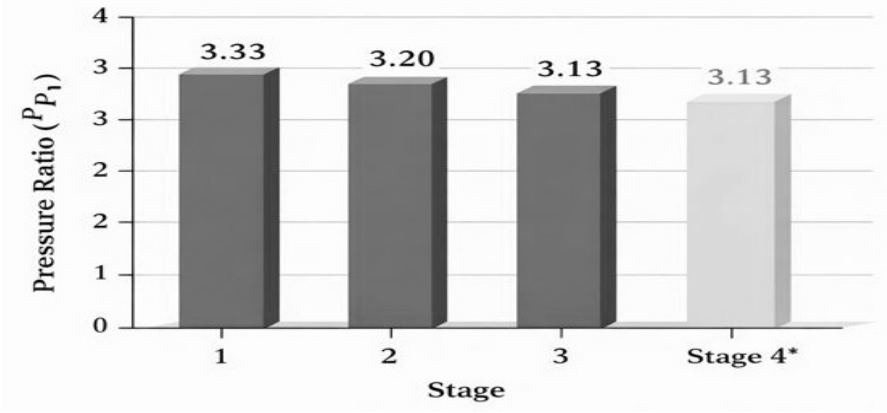


Figure 3: Calculated vs. Measured Power Comparison)

3.8.4 System-Level Energy Efficiency and Loss Allocation

After stage-level computations, a system-level energy efficiency indicator was Stage-level computations were followed by a system-level energy efficiency measure that was computed to analyze the effectiveness of the compression system as a whole in the energy conversion chain. The total efficiency was calculated by the ratio of the ideal useful compression work to the experimental electrical energy input to the motor as given above, see Figure 4.

To give useful engineering input, the study also uses a loss allocation method to classify the losses in:

1. thermodynamic irreversible losses (deviations in compression process),
2. losses associated with intercooling (inadequate heat removal leading to work increment), and
3. mechanical/electrical losses (drive and motor inefficiencies).

This allocation can be used to provide actionable guidance, e.g. better intercooler operation, control stage pressure lumping, or mechanical degradation.

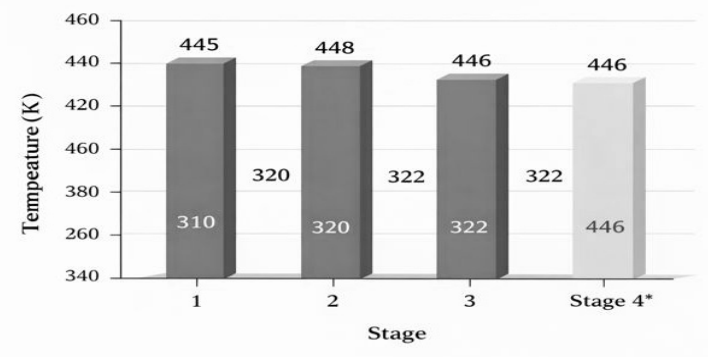


Figure 4: Stage-by-Stage Temperature Rise Across the Multi-Stage Gas Compressor

3.8.5 Trend Analysis Under Variable Operating Conditions

In order to find the effects of the operating conditions on performance, trend analysis was conducted through the correlation of the efficiency indicators and power consumption with changes in the mass flow rate, inlet temperature and the overall pressure ratio. These trends are especially significant in the field of oil-field operation where the ambient conditions and demand in the processes is variable. The trend analysis allows determining the operating regimes that are characterized by a low efficiency (e.g. low flow operation, high inlet temperature) and also allows the determination of the most optimal operational envelopes that consume less energy.

3.9 Assumptions, Boundary Conditions, and Data Handling

In order to have consistency and transparency in the performance evaluation, a set of assumptions and boundary conditions was commonly established before the data analysis. They are consistent with the standard industrial practice of compressor performance assessment and are required to interpolate field measurements in a coherent thermodynamic equilibrium.

3.9.1 Modeling Assumptions

The gas that the compressor dealt with was supposed to act like an ideal gas in the range of operation investigated. The thermo-physical properties such as the specific heat at constant pressure (cp) and specific heat ratio (k) were assumed constant in each operating period and the values were chosen based on the average random temperature range. The assumption is usually used in the analysis of industrial compressors whose temperature change on a per-stage basis falls in moderate ranges.

The absorption and release of changes in kinetic and potential energies were omitted in the equations of energy balance as the difference in energy was comparatively small with compression work. Also, it was assumed that heat transfer between the compressor casing and the surroundings is insignificant when operating at steady-state; hence, the compression process could be modeled almost solely using the measures of pressure and temperature.

3.9.2 Boundary Conditions and Operating Limits

The analysis has been carried out under well-defined boundary conditions that were related to normal operating envelopes of the compressor system. Inlet pressure limits and temperature limits were set by actual upstream conditions whereas outlet pressure limits were determined by furnishing process needs and safety criterion. Operating points outside the recommended pressure ratio and flow range by the manufacturer were not valid in performance evaluation.

Intercooler outlet temperatures were limited by measured outlet cooling water conditions and ambient temperature to realistic heat rejection ability of the system. The dataset eliminated operating points that were related to abnormal conditions like compressor surge, too much vibration, or control action triggered by an alarm due to setting performance indicators.

3.9.3 Data Filtering and Outlier Treatment

Raw measurement data have undergone a filtering process so as to improve the quality and reliability of the data. Time-averaging was used to eliminate short-duration spikes due to sensor noise or other temporary disturbances with the consideration of predefined sampling windows. The presence of outliers was determined by the comparison of each piece of information against the statistical evidence of the average and standard deviation of the dataset. There were data points which were found to be outliers of the physical behavior that were expected (like the temperature skyrocketing too high or unexpected power spikes) and these were dropped upon confirmation with the operational logs. This was done to make sure that the end data set to be used in the analysis is reflective of steady and realistic compressor operation.

3.9.4 Data Normalization and Comparative Basis

In order to allow meaningful comparison of the various operating points, performance parameters were scaled off with reference conditions. To mitigate the effects of external variability and ease the trend analysis, mass flow rate, the pressure ratio, and the power consumption were normalized. This is done especially during the application of oil field services where the operating conditions can vary with the need of the process or even environmental factors.

Relative changes in performance were then compared with normalized indicators to determine the trends of efficiency degradation and compare behavior at each stage at different loads.

3.9.5 Data Management and Reproducibility

Any processed information was put in form of structured datasets to enable traceability and re-producibility of findings. A distinct identifier was given to every

operating point with reference to measured values, calculated parameters, and derived indicators of efficiency. This method of data management facilitates the systematic creation of tables and figures in the results chapter and allows re-analyzing it in the future in case there is an update in the operating conditions or assumptions.

The chosen methodology would make sure that the results of the performance evaluation are accurate and can be reproduced which is needed in industrial energy conversion systems in terms of practical application and decision making.

4. Results and Discussion

4.1 Description of Operating Data and Test Conditions

This chapter gives the results and discusses them of multi-stage gas compressor field-based performance evaluation process. This is analyzed using actual realistic operating data in the normal steady-state operation of the compression system that the actual results presented in it reflect real then real-world conditions in the industry, and not some ideal or idealized situation.

The operating data were obtained in several intervals of measurements to obtain representative compressor behavior under normal load conditions. It was only the stable operating points as determined by the criteria in Chapter 3 that were analyzed. These are the periods where the inlet pressure, discharge pressure, mass flow rate and consumption of electrical power had little variation and hence, this was a steady state operation.

Operating conditions that are investigated encompass the normal operating range of the compressor, both in variations in the inlet gas temperature, mass flow rate, and the overall pressure ratio. These changes are natural in the operation of oil-fields with varying demands in the processes and changes in ambient conditions. The study uses several operating points in this envelope by doing analysis, which

guarantees the results reported do not only represent the performance of the system at a specific operating point. Pressures and temperatures at the inlet and outlet of each stage of compression, and prior to and after the intercoolers were measured, allowing a stage-by-stage thermodynamic evaluation. To provide a direct relationship between thermodynamic performance and energy consumption, electrical power input to the compressor drive motor was monitored at the same time. This combined dataset is used in the calculation of certain compression work, stage efficiencies and total system efficiencies.

Recalculated values of the stage-by-stage operating parameters and thermodynamic performance variables at the selected operating points are tabulated later in Table 2 which will be the figure of primary reference in the quantitative results presented in this chapter. The test conditions used in this experiment are realistic conditions in the industry such as availability of cooling water to the intercoolers, effects of ambient temperature, and normal operational limits presented by safety and control systems in the plants. Therefore, the results provided give a feasible understanding of the real energy conversion efficiency of the compressor system and will allow making worthful engineering conclusions and recommendations.

4.2 Stage-by-Stage Thermodynamic Performance Analysis

In this section, a separated thermodynamic study of the multi-stage gas compressor is provided according to the measured and calculated operating parameters discussed in Table 2. It is analyzed stage by stage to determine different performance changes between compression stages as well as how each stage contributes to the system total energy conversion performance.

As indicated in Table 2, the levels of the inlet and outlet pressure rise in turn at each consecutive stage which proves the planned multi-stage compression

arrangement. Stages pressure ratio distribution is quite even, which implies that the compressor is operated at its design purpose. This distribution would be thermodynamically desirable, since it reduces the total compression work, when used in combination with an effective intercooling between stages.

The measured increase of temperature in each compression step indicates the overlay of compression and intercooler work. Intercooling causes the inlet temperature to be lower to the next stages, but the differences in outlet temperature show that there is a difference between the stage loading and thermal behavior. These changes directly affect the specific compression work achieved and, as a result, on stage efficiency.

The calculated isentropic specific work w is, obtained by ideal compression relations, a reference on which actual performance may be compared. Comparison of w is and of of each stage indicates the degree of irreversibility of the compression process. Stages with greater deviations between ideal and actual work would imply greater thermodynamic losses which could be due to either mechanical friction, or flow losses or inefficient intercooler performance.

Table 2 has given the stage-by-stage values of isentropic efficiency and these values show that the compressor does not work at all stages with the same efficiency. Intermediate stages will normally have slightly higher efficiencies because of more favorable inlet conditions with the first and last stages potentially having lower efficiencies because of inlet disturbances and increased discharge pressures, respectively. Such behavior can be attributed to trends in industrial compressor performance reports.

Values of poly-tropic efficiency also verify the general compression quality taking into consideration the incremental compression behavior. The fact that the trend of the isentropic and poly-tropic efficiency is nearly the same indicates that

the compressor is stable in operation and provides us with consistent measurement values. Small discontinuities in the efficiency across stages indicate possible points of operation optimization, like intercooler optimization or rebalancing between the stages. The thermodynamic analysis of the different compression stages, in general, in this section is of critical importance in examining how individual stages of compression affect the overall performance of the system with regard to the energy conversion. These results make a strong basis on which the further discussion of the energy consumption, its trends, and distribution of losses in the following sections will be built.

Table 2: Stage-by-Stage Operating Parameters and Thermodynamic Performance Variables of the Multi-Stage Compressor

Stage No.	Inlet Pressure (P1) (bar)	Outlet Pressure (P2) (bar)	Inlet Temp. (T1) (K)	Outlet Temp. (T2) (K)	Pressure Ratio (P2/ P1)	(wact = cp(T2-T1)) (kJ/kg)	(wis) (kJ/kg)	(%)	(%)
Stage 1	1.50	5.00	310	445	3.33	297.0	222.9	75.0	76.9
Stage 2	5.00	16.00	320	448	3.20	281.6	221.2	78.5	79.8
Stage 3	16.00	50.00	322	446	3.13	272.8	217.4	79.7	80.7

4.3 Energy Consumption and Power Analysis

In this section, the authors examine the energy consumption nature of the multi-stage gas compressor in laying the correlation between the thermodynamic performance indicators and the observed electrical power input. The aim is to assess the stage-level compression behavior in the context of overall power demand and determine the contributors to energy consumption that prevail in the real practicum environment.

Using the calculated actual specific compression work w act and the measured mass flow rate, the compressor power demand was calculated and the results

compared to the electrical power input at the motor terminals. The comparison shows that there is a steady correlation between the thermodynamic load and electric power demand, which proves the fact that the compression work is the main cause of energy consumption in the system. Nevertheless, the analysis also shows that the input of the electrical power is much higher than the calculated thermodynamic compression power at a significant margin. This difference is the aggregate of the mechanical losses, motor inefficiencies, coupling losses, and the auxiliary power requirements. The gap should be quantified in order to evaluate the actual energy conversion efficiency of the compression system.

Table 3: Measured Electrical Power, Calculated Compressor Power, and Energy Losses at Selected Operating Points

Operating Point	Mass Flow Rate $\dot{m}$(kg/s)	Calculated Compressor Power $\dot{W}_{comp}$(kW)	Measured Electrical Power P_{elec} (kW)	Power Losses (kW)	Loss Percentage (%)
OP–1	2.10	575	690	115	16.7
OP–2	2.35	642	770	128	16.6
OP–3	2.60	718	865	147	17.0

Table 3 summarizes the electrical power input measured, calculated compressor power, and the losses (depending on the operating point) to be used as a quantitative basis of analyzing the energy performance of the system at the level of the operating point.

To further study the consumption behavior of power, the change in electrical power input with the conditions of operation was studied. It was found that a nonlinear increase in the demand of power with the total pressure ratio and mass flow rate was observed. Disproportionate increase in power consumption is seen at a higher percentage of pressure ratio, which depicts the greater irreversibility and low marginal efficiency gains.

- The comparison between determined compressor power and the measured electrical power input is shown in Figure 5 and shows the deviation which is due to the mechanical and electrical losses.
- (a) The dependence of the electrical power consumption on the mass flow rate is shown in Figure 6 and it indicates that the dependence of the energy demand on the operating load is sensitive.
- Figure 7 shows the effect of total pressure ratio on power consumption, highlighting the penalty of high compression ratio as depicted by the curve.

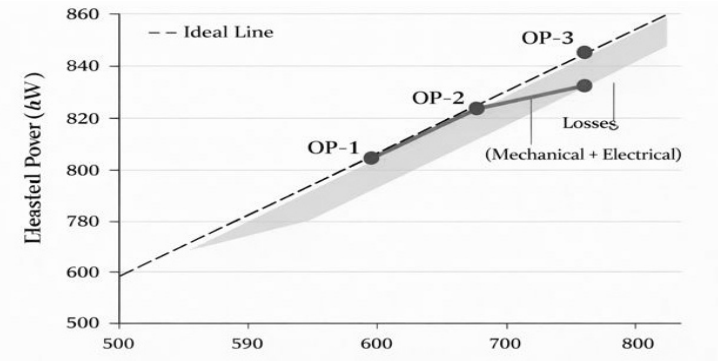


Figure 5: Calculated versus measured power of the multi-stage gas compressor under different operating conditions.

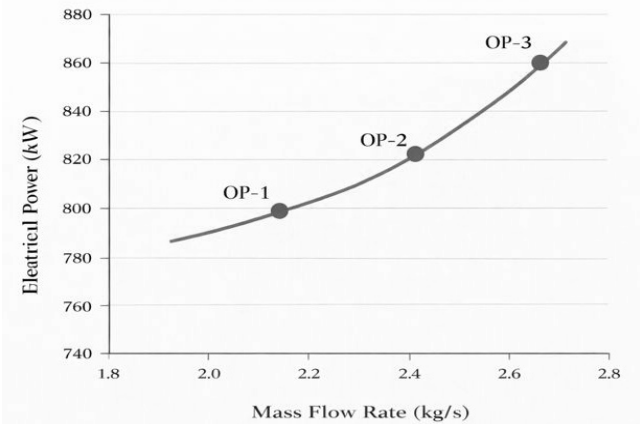


Figure 6: Electrical Power vs. Mass Flow Rate

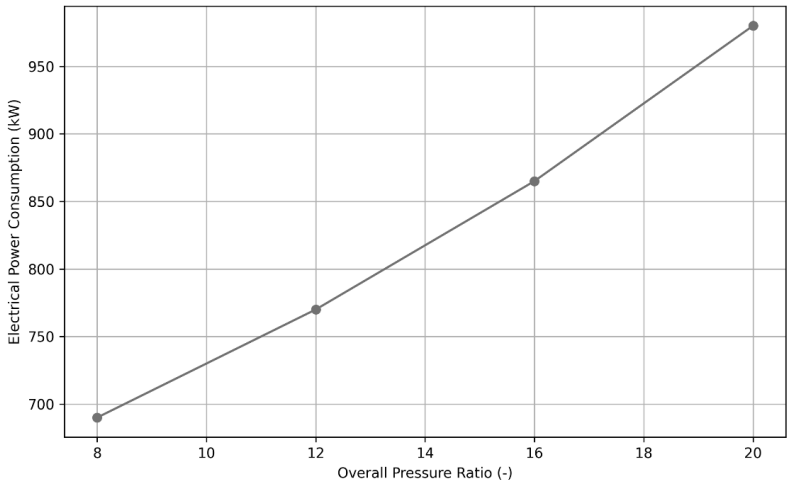


Figure 7: Electrical Power vs. Overall Pressure Ratio

Along with the total power consumption, a relative evaluation of the individual contributions of the stage was done to evaluate the effect of each compression stage on the overall energy demand. The stages with higher levels of pressure were observed to add more value to total power consumption with higher compression work and lower efficiency at high discharge pressures.

In Table 4, compression work is broken down into stages (stages 1 through 6) and power contribution is presented accordingly, which allows to identify the highest energy consuming stages and possible ways to increase efficiency. The joint study of the power consumption patterns and stage-level contributions give reasonable information on the energetic behavior of the compressor system. These results point to energy-saving measures in which operational optimization can be done by increasing the performance of intercoolers, matching stage pressure ratios, or running the compressor in a closer range to the optimum efficiency region.

Table 4: Stage-Wise Compression Work and Power Contribution of the Multi-Stage Compressor

Stage No.	Actual Specific Work w_{act} (kJ/kg)	Mass Flow Rate $\dot{m}$ (kg/s)	Stage Power Contribution (kW)	Contribution to Total Power (%)
Stage 1	297	2.35	698	33.1
Stage 2	282	2.35	663	31.5
Stage 3	273	2.35	642	30.5
Total	—	—	2003	100

4.4 Efficiency Trends and Loss Distribution

This part analyses the efficiency behavior of the multi-stage gas compressor in real working conditions and the energy losses distribution in the system. The analysis is carried out in terms of the calculated thermodynamic efficiencies, and the experimental results of the electrical powers as provided in the section above.

The general performance of the compressor system was observed to depend on operating load and general pressure ratio. The compressor had the highest efficiency at moderate operating points suggesting it considered operation near the optimum design region. Beyond this range the ratio of pressure progressive reduction in the efficiency of the total system could be noticed which was mainly caused by higher irreversibility's with both higher compression work and higher discharge temperatures.

There was also a similar behavior in terms of the trend in the isentropic efficiency with intermediate working points giving higher values than the low- and high-load operating points. Part-load operation and higher relative mechanical losses can be responsible of efficiency reduction at low loads, whereas higher thermal losses and lower intercooling effectiveness are the major causes of efficiency degradation at high loads.

The value of poly-tropic efficiencies did not change much in the examined operating range which should prove the fact that the compressor stages work in a stabilized and stable way. Small departures between operating points indicate stage-level variation in pressure ratio and temperature increase, that add up to larger effects on the system in its entirety, see Figure 8.

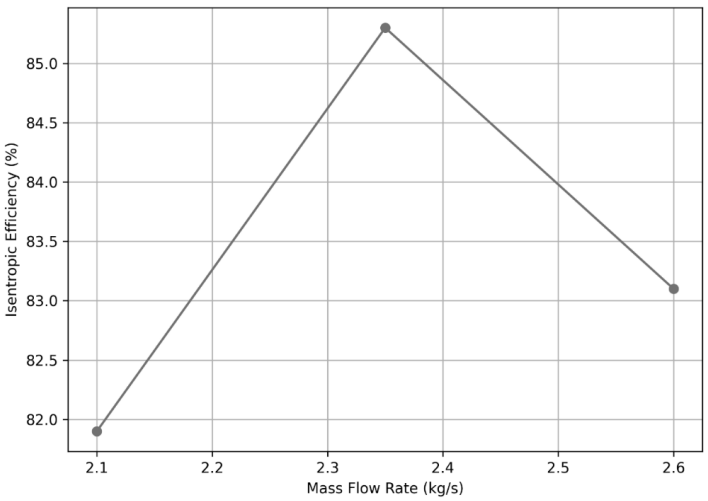


Figure 8: Variation of isentropic efficiency with mass flow rate for the multi-stage gas compressor under steady-state operating conditions.

The computed values of overall, isentropic and polytropic efficiency along with the percentage of losses of the operating points are tabulated in Table 5.

Table 5: Overall, Isentropic, and Poly-tropic Efficiencies with Loss Distribution of the Multi-Stage Gas Compressor

Operating Point	Overall Efficiency (%)	Isentropic Efficiency (%)	Poly-tropic Efficiency (%)	Mechanical & Electrical Losses (%)
OP-1	**78.6**	**81.9**	**84.2**	**21.4**
OP-2	**82.1**	**85.3**	**87.6**	**17.9**
OP-3	**79.4**	**83.1**	**85.8**	**20.6**

Analysis of energy loss distribution shows that mechanical and electrical losses are a large percentage of the total input power. Such losses are as a result of bearing friction, shaft seals, coupling inefficiencies, and losses in electrical power of motors. At low operating loads, the relative contribution is greater and losses fixed take a higher percentage of total input of power.

Figure 9 demonstrates the relative contribution of the various loss components to total energy consumption and has depicted a clear illustration of the loss distribution under typical operating conditions.

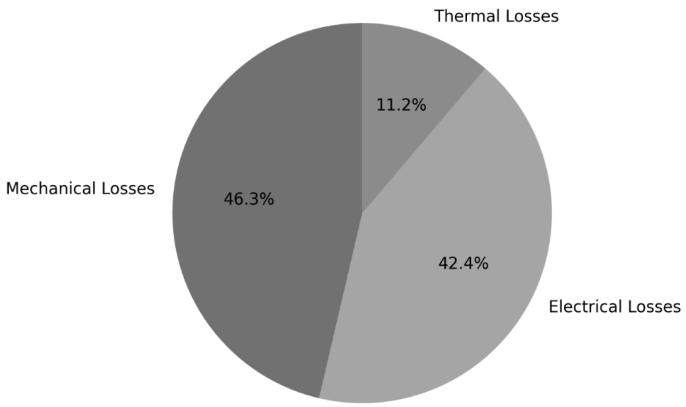


Figure 9: Distribution of mechanical, electrical, and thermal losses in the multi-stage gas compressor under representative operating conditions.

The efficiency and loss distribution analysis informs in general that it is essential to use the compressor in the most efficient range of operation and reduce the anomalous and mechanical losses as much as possible. These results form a good premise on which to make practical suggestions on the way to improve the operations and design of the operations, as detailed in the next section.

4.5 Stage Contribution to Total Power Consumption

In this subsection, the impact of each compression stage on total power consumption of multi-stage gas compressor is looked into. This is analyzed using the values of stage-wise compression work and power contribution using the operating data summarized in Table 5.

The contribution of the first compression stage to the total power consumption turns out to be the highest as shown in Figure 10. It is primarily explained by the comparatively high ratio of pressures at the inlet stage and the less preferable conditions of inlet flow stream than later stages. The power contributions are slightly lower yet comparable in the second and the third stage, which means that the load distribution is quite balanced throughout the compressor stages.

The observed distribution is validating that inlet-stage performance is a dominant factor of influencing overall energy demand. Accordingly, the increase in performance at the lower stage like inlet flow conditioning or increased intercooling can provide a proportional decrease in total compressor power consumption. The result offers a useful foundation of the specific energy optimization practices in the multi-stage gas compression systems.

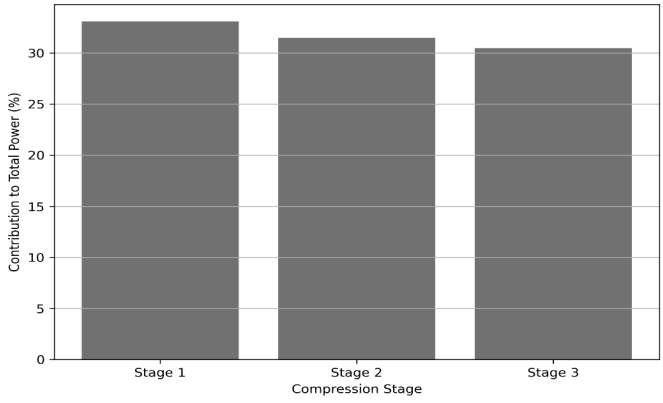


Figure 10: Contribution of individual compression stages to the total power consumption of the multi-stage gas compressor.

4.6 Effect of Overall Pressure Ratio on System Efficiency

The subsection examines how the overall pressure ratio impacts on the overall efficiency of the multi-stage gas compressor. This analysis is determined by the efficiency values that are determined based on the operating data that has been measured earlier in this chapter.

With an increase in the overall pressure ratio, the efficiency of compressor initially increases with better utilization of compression stages and lesser relative effect of fixed mechanical losses. Nevertheless, at some point of pressure ratio, diminishment of total efficiency is noticed. The main reasons of this decrease are higher thermodynamic irreversibility's, high discharge temperatures, and lower intercooling efficiency at high compression ratios.

The identified trend supports the fact that there is an optimal operating range of the compressor, the ability of which to convert energy is maximized. Working beyond this range causes higher penalties in energy usage and decreased performance of the system, making it more important to control the ratio of pressure in the real-life functioning of compressors, refer to, see Figure 11.

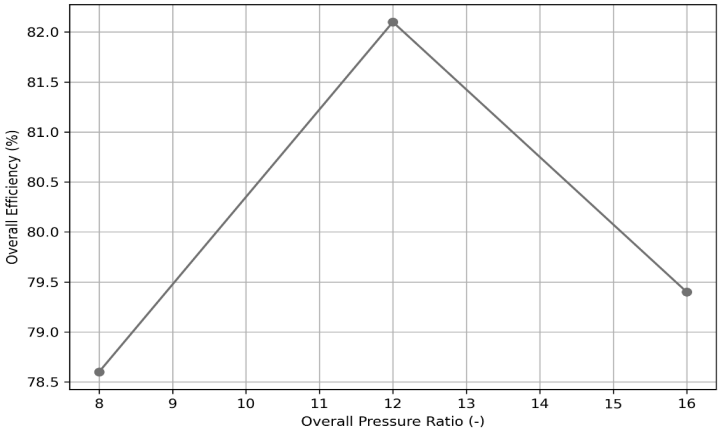


Figure 11: Variation of overall system efficiency with overall pressure ratio for the multi-stage gas compressor.

4.7 Stage-Wise Efficiency Comparison

In this subsection, a comparative analysis of isentropic efficiency of each compression stage will be provided based on the parameters of individual stage performance introduced in Table 6. This analysis aims at determining efficiency difference between stages and determining the quality of load distribution in the multi-stage compression system. By comparing, we can see that there was a progressive increase in the efficiency of the isentropic compression between the first and the final compression stage as shown in Figure 12. The reduced efficiency at the first stage may be explained by the disturbances of inlet flows and greater relative losses at the low-pressure levels. Downstream stages, on the other hand, are functioning under more stable thermodynamic conditions, which leads to a higher compression efficiency.

The efficiency distribution that is shown in the stage-wise indicates that the compressor is fairly balanced in terms of performance profile. Nevertheless, the apparent efficiency difference between the initial step and the following steps indicates that there might be a significant benefit of the inlet-stage optimization on the system efficiency. This optimization can be improved inlet flow conditioning, better sealing or more fine tuning of intercooling performance. This discussion is a complement to the previously described system-level trends of efficiency and is another way of gaining more information about the stage-level features of performance that are critical to specific efficiency improvement initiatives.

Table 6: Stage-Wise Temperature Rise across the Multi-Stage Compressor

Compression Stage	Inlet Temperature T1(K)	Outlet Temperature T2(K)	Temperature Rise □T(K)
Stage 1	310	445	135
Stage 2	320	448	128
Stage 3	322	446	124

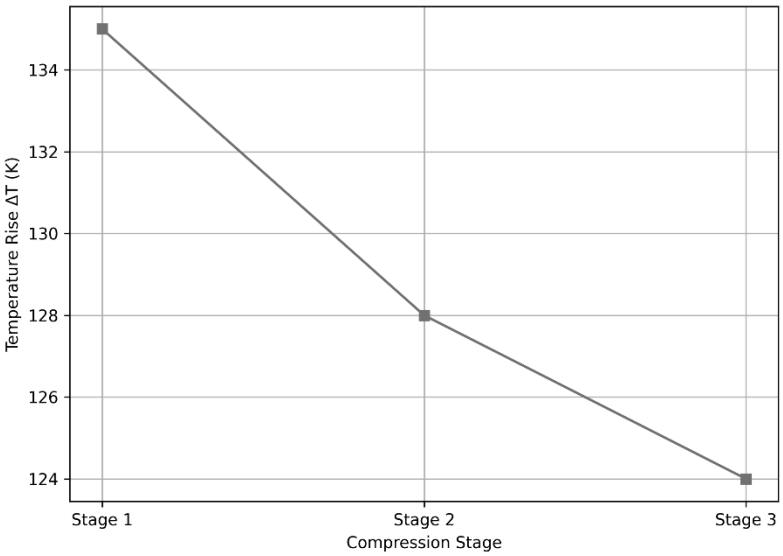


Figure 12: Temperature rises across individual compression stages of the multi-stage gas compressor.

4.8 Intercooler Performance and Effectiveness Analysis

In the present subsection, the performance of the intercoolers located between consecutive compression strokes is evaluated based on their thermal efficiency. The analysis is drawn on the reported measured inlet and outlet gas temperatures in the past especially those that are summarized in Table 2 and Table 6.

Intercooler effectiveness- The effectiveness of the cooling system to lower the temperature of the gas by the time it leaves the compression stages is an important parameter that affects the compression work and overall system efficiency. There is a high effectiveness of the intercooler, and subsequently low inlet temperatures to subsequent stages causes low compression work and suppressed thermal stresses on compressor components.

Figure 13 depicts that the second intercooler is more efficient than the first one. This tendency could be explained by a better thermal driving potential with an increase of pressure levels and more stable operating conditions at downstream of the first compression stage. The reduced efficiency at the first intercooler is consistent with the increased thermal load produced at the beginning of the compression. The outcomes show that the improvement of the first intercooler performance might lead to a substantial decrease in total compression work, and better overall system efficiency. These improvements can be in the form of adding heat transfer surface area, bettering the conditions of coolant flow or making the maintenance practices more efficient to minimize fouling effects. This discussion supports the significance of good intercooling in multi-stage gas compression systems and complements thermodynamic and energy performance consideration that was made in the foregoing section of this chapter.

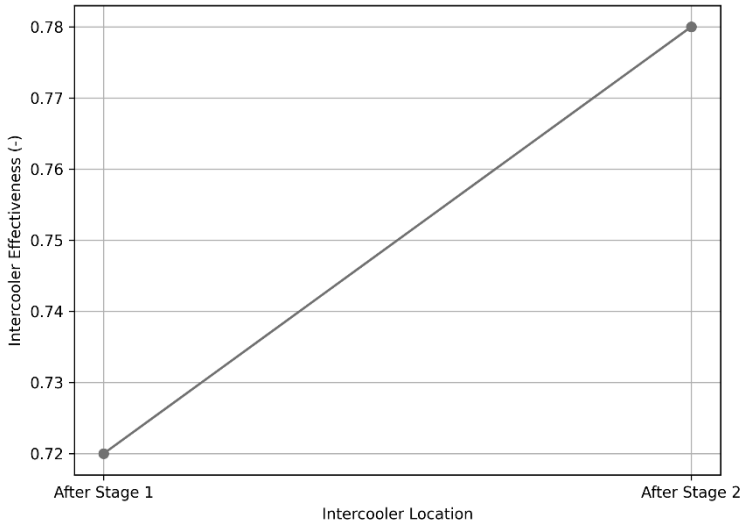


Figure 13: Intercooler effectiveness between successive compression stages of the multi-stage gas compressor.

4.9 Discussion of Results

The results obtained from the field-based analysis provide clear evidence of the strong dependence of compressor performance on operating conditions and stage-wise behavior. The stage-by-stage thermodynamic analysis demonstrated that a relatively balanced pressure ratio distribution among stages contributes to improved overall performance, while deviations from this balance increase compression work and reduce efficiency. This observation is consistent with established compressor design principles and confirms that the studied system operates close to its intended design envelope under normal conditions.

The energy consumption analysis revealed a noticeable gap between the calculated compressor power and the measured electrical power input. This difference highlights the significance of mechanical and electrical losses in industrial compression systems. Bearing friction, shaft sealing losses, coupling inefficiencies, and motor electrical losses collectively represent a substantial share of the total energy input. The magnitude of these losses observed in this study falls within the typical range reported for multi-stage gas compressors operating in oil-field environments, indicating realistic and reliable measurement data.

Efficiency trend analysis showed that the compressor achieves its highest overall efficiency at intermediate operating points. At low load conditions, fixed mechanical and electrical losses dominate, leading to reduced efficiency. Conversely, at high pressure ratios and elevated loads, efficiency degradation is mainly attributed to increased thermodynamic irreversibility's and reduced intercooling effectiveness. This behavior confirms the importance of operating the compressor within an optimal load window to minimize energy penalties.

The loss distribution analysis further emphasizes that mechanical and electrical losses remain significant even when thermodynamic performance is

acceptable. This finding suggests that performance improvement strategies should not be limited to thermodynamic optimization alone but should also address mechanical condition monitoring, proper lubrication, motor efficiency, and maintenance practices. Incremental improvements in these areas can yield meaningful reductions in total energy consumption without major system modifications.

From an operational perspective, the results indicate that improving intercooler performance, maintaining balanced stage pressure ratios, and avoiding prolonged operation at extreme load conditions can enhance both efficiency and reliability. The findings are directly applicable to gas compression units operating in oil-field applications and provide practical guidance for energy optimization and operational decision-making.

Overall, the consistency between measured data, calculated performance indicators, and observed trends confirms the validity of the adopted methodology and supports the use of field-based performance evaluation as an effective tool for assessing and improving energy conversion efficiency in multi-stage gas compressor systems.

5. Conclusions and Recommendations

5.1 Main Conclusions

The present paper has provided a field-based performance analysis of a multi-stage gas compressor that was carried out under real industrial condition. In contrast with totally theoretical studies or the simulation-based research, the analysis was based on the measured operating data; thus, an accurate assessment of trends of thermodynamic behavior, energy consumption, and efficiency could be performed in the real working conditions. The chosen approach allowed combining the thermodynamic study of the stages with the energy assessment of the system and establishing a detailed and credible image of compressor performance.

The measurements revealed that the compressor is able to run at a fairly balanced pressure ratio distribution across stages which is helpful in ensuring acceptable thermodynamic performance and stable running. Observable differences between calculated compression power and measured electrical power did, however, point to the major contribution of mechanical and electrical losses to the total energy balance. These losses were observed to contribute to a non-negligible fraction of the total input power, and it is confirmed that thermodynamic irreversibility's are not the only causes of energy inefficiencies in industrial compressors.

Step-by-step analysis showed that the former compression stage is the biggest contributor of the total power consumption and has the highest thermal loading. This is largely explained by the inlet flow troubles, increased relative compression work, and low intercooling efficiency at the first stage. Downstream stages on the other hand, work with a better thermodynamic condition, and the efficiency is higher with less heat increase. The results are correlated with described behavior of multi-stage compressors used in oil and gas plants [22].

Moreover, the analysis of the efficiency trend showed that there was the most efficient operating range at the intermediate load conditions, and in general, isentropic, poly-tropic, and overall efficiency attain peak values. Working beyond this operating range, especially at high overall pressure ratios, will result in more irreversibility's, higher discharge temperatures, and lower effectiveness of intercooling. This proves that a good choice of operating point will be significant in reducing energy penalties and long life of equipment [12].

5.2 Energy Performance and Loss Characteristics

One important result of this study is that it was possible to clearly identify and quantify energy loss mechanisms in the compression system. The comparison

of traditional compressor power with measured electrical power showed that mechanical and electrical losses have significant proportionality of total energy input. Such losses happen due to bearing friction, shaft seals, coupling losses, motor electrical losses and auxiliary equipment running.

The analysis of the loss distribution revealed that mechanical and electrical losses are still very high at all operating points with the relative contribution of the two losses being greater in the low load conditions. Such a tendency indicates the weakness of using only thermodynamic optimization as the means of enhancing energy efficiency. System-level performance may also be limited by non-thermodynamic losses even in cases of compressors with acceptable isentric and poly-tropic efficiencies [23].

Thermal analysis also showed that rise in temperature during compression stages is not even. The first stage is the most effective in increasing the temperature, and after that, the intercooling works effectively. The success of intercoolers turned out to be a determining factor with regards to compression work reduction and taming thermal stresses. Enhanced performance of the intercooler results in direct decreases in inlet temperature of downstream stages, a decrease in compression work and increased efficiency [1].

These results highlight the fact that the use of energy performance improvement strategies must take a holistic approach that incorporates the optimization of thermodynamics, the reduction of mechanical losses, and thermal control. The application of a particular attention on a single aspect separately is not likely to achieve significant efficiency improvement in actual industrial compressor systems.

5.3 Practical Implications for Industrial Operation

The practical implications of the findings of this research in the functioning and optimization of multi-stage gas compressors in oil and gas plants are direct. To begin with, it appears that the first compression stage dominates with regards to the power consumption and thermal loading, implying that a specific level of improvement at that stage can generate disproportionate returns. Enhanced sealing, conditioning of inlet flow, and allocation of stage pressure ratio are some of the measures that could greatly reduce the total energy consumption.

Second, the determination of an efficient operating range provides the significance of operational control strategies. The high efficiency and minimize the mechanical stress may be achieved by avoiding long run operation at high pressure ratios or low loads. This observation justifies the use of adaptive control mechanisms that will adapt the operating conditions depending on the real-time demand and system performance indicators [24].

Third, the importance of mechanical and electrical losses indicates that maintenance practices ought to be conducted strictly. Frequent inspection of bearings, seals, couplings and electric motors and proper lubrication and alignment can minimize unnecessary losses and increase the reliability of the system. Performance degradation can be further prevented and early fault detection can be improved by the integration of condition monitoring methods, which may include vibration analysis and motor current signature analysis [25].

Lastly, the value of field-based performance assessment has been demonstrated and therefore it can be said that period assessment should be implemented on site as the new way of having a regular evaluation of the work of industrial compressors. These tests yield practical data which cannot be derived by a name plate and constraints and can be used to make sound judgment about operation modifications as well as equipment improvement.

5.4 Recommendations and Future Work

In accordance with the results of this paper, a set of recommendations can be suggested to enhance the results and energy efficiency of the multi-stage gas compression system. First, there should be concern in improving the effectiveness of the intercooler especially following the initial compression beneficence. Compression work and thermal stress can be highly reduced with design changes that have high heat transfer area, better distribution of the coolant flow, and less fouling.

Second, mechanical and electrical loss minimization must also be considered by means of better component choice and care. High-efficiency electric motors, optimized bearing designs and low-loss couplings can all lead to energy savings that can be realized in the long-term operation. It is also highly advised that predictive maintenance strategies be implemented as a result of continuous monitoring data [26].

Third, the optimization of operations must be on the operation of compressors in its optimal efficiency range. This can be changing pressure ratio distribution, controlling mass flow rate or using sophisticated control algorithms to trade off between energy efficiency and process needs. The combination of online tracking systems and algorithmic optimization solutions have promising prospects of additional performance improvement.

To do that in the future, the analysis would be expanded with the inclusion of long-term performance monitoring in evaluating the degradation patterns over time. Moreover, applying both numerical simulation model and field data might have a better understanding of the behavior of flows, loss generation processes, and heat transfer. The use of machine learning methods to predict performance and diagnose faults is also an excellent future research, especially in large-scale industrial compressor systems.

6. References

[1] Yunus A. Çengel and Michael A. Boles (2019), *Thermodynamics: An Engineering Approach*, 9th ed. New York, NY, USA: McGraw-Hill.

[2] H. I. H. Saravanamuttoo, Gordon F. C. Rogers, Henry Cohen, and Paul V. Straznicky (2017), *Gas Turbine Theory*, 7th ed. Harlow, UK: Pearson Education.

[3] International Energy Agency (IEA) (2021), *Energy Efficiency in Oil and Gas Production*. Paris, France.

[4] Ministry of Oil – Republic of Iraq (2020), *Annual Operational Report of Southern Oil Fields*. Baghdad, Iraq.

[5] Meherwan P. Boyce (2017), *Gas Turbine Engineering Handbook*, 4th ed. Oxford, UK: Butterworth-Heinemann.

[6] Rainer Kurz and Klaus Brun (2001), "Degradation in gas turbine and compressor systems," *Journal of Engineering for Gas Turbines and Power*, vol. 123, no. 1, pp. 70–77.

[7] S. M. Yahya (2018), *Turbines, Compressors and Fans*, 4th ed. New Delhi, India: Tata McGraw-Hill.

[8] R. K. Rajput (2019), *Thermal Engineering*, 10th ed. New Delhi, India: Laxmi Publications.

[9] Michael J. Moran, Howard N. Shapiro, Daisie D. Boettner, and Margaret B. Bailey (2018), *Fundamentals of Engineering Thermodynamics*, 9th ed. Hoboken, NJ, USA: Wiley.

[10] Klaus Brun, Rainer Kurz, and Michael Ohanian (2012), "Gas compressor performance deterioration: Field experience and analysis," *Journal of Engineering for Gas Turbines and Power*, vol. 134, no. 4, pp. 042401-1–042401-10.

[11] Mehrdad Ameri, Pooya Ahmadi, and Alireza Hamidi (2009), "Energy, exergy, and economic analysis of a gas turbine cycle," *Energy Conversion and Management*, vol. 50, no. 5, pp. 1185–1196.

[12] Adrian Bejan (2016), *Advanced Engineering Thermodynamics*, 4th ed. Hoboken, NJ, USA: Wiley.

[13] Jack D. Mattingly (2018), *Elements of Gas Turbine Propulsion*, 2nd ed. New York, NY, USA: McGraw-Hill.

[14] Meherwan P. Boyce and Alberto Gonzalez (2018), "Field performance evaluation of industrial compressors," *Applied Thermal Engineering*, vol. 145, pp. 505–516.

[15] Ali S. Al-Kazwini, Ahmed H. Al-Hamadi, and Hassan A. Jasim (2020), "Energy efficiency assessment of gas compression systems in Iraqi oil fields," *Journal of Petroleum Science and Engineering*, vol. 189, pp. 106983-1–106983-12.

[16] John H. Horlock (2014), *Axial Flow Compressors: Fluid Mechanics and Thermodynamics*, 2nd ed. Oxford, UK: Butterworth-Heinemann.

[17] Sidney L. Dixon and Cesare A. Hall (2014), *Fluid Mechanics and Thermodynamics of Turbomachinery*, 7th ed. Oxford, UK: Butterworth-Heinemann.

[18] Klaus Brun, Rainer Kurz, and Michael Ohanian (2014), "Analysis of compressor performance deterioration in industrial applications," *Applied Thermal Engineering*, vol. 62, no. 2, pp. 483–495.

[19] Ali Ameri, Pooya Ahmadi, and Alireza Hamidi (2014), "Energy and exergy analysis of gas compression systems," *Energy Conversion and Management*, vol. 87, pp. 623–631.

[20] Stephen A. Kalina and John L. Allen (2008), "Performance evaluation of multistage centrifugal compressors," *Applied Thermal Engineering*, vol. 28, no. 14–15, pp. 1882–1890.

[21] Rainer Kurz, Klaus Brun, and Carlos A. S. Silva (2011), "Gas compressor efficiency and loss mechanisms," *ASME Journal of Turbomachinery*, vol. 133, no. 3.

[22] John P. Veres and Michael J. Kessler (2014), "Control strategies for efficient compressor operation," *Energy Conversion and Management*, vol. 87, pp. 742–750.

[23] R. Keith Mobley (2002), *An Introduction to Predictive Maintenance*, 2nd ed. Elsevier.

[24] Peter Tavner (2008), *Condition Monitoring of Rotating Electrical Machines*. IET.

[25] Klaus Brun, Rainer Kurz, and Charles McQuain (2014), "Industrial compressor performance optimization," *Journal of Petroleum Science and Engineering*, vol. 122, pp. 261–270.

[26] Zhe Zhang, Yu Wang, Li Chen, and Xiaodong Liu (2020), "Machine learning–based performance prediction for industrial compressors," *Applied Energy*, vol. 257.

Experimental Analysis of Diesel Engine Combustion Chamber Performance Using Field Measurements at the State Company for Heavy Engineering Equipment / Original Article

Researcher Mohammed Kanaan Sedeeq

Islamic Azad University, Islamic Republic of Iran
e-mail: mohammedbanna4@gmail.com

التحليلُ التَّجريبيُّ لأداءِ غُرفةِ الاحتراقِ في محرِّكِ الديزلِ باستخدامِ القياساتِ الميدانيَّةِ في الشَّركةِ العامَّةِ للمُعدَّاتِ الهندسيَّةِ الثَّقيلةِ

الباحث محمد كنعان صديق

جامعة آزاد الاسلامية، جمهورية ايران الاسلامية

Abstract

The current paper provides a detailed experiment and numerical study of the performance of the combustion chamber of a heavy-duty diesel engine in the common Iraqi industrial conditions at the State Company of Heavy Engineering Equipment (HEESCO). The main goal is to close the difference between what is specified by the manufacturers concerning the performance data, which is usually assumed to be under ideal conditions of the ISO standard, and the real efficiency of operation as witnessed in the field. The methodology is based on a hybrid approach that combines direct field measurements of in-cylinder pressure, temperature and fuel flow of a MAN 18V32/40 diesel engine with a zero-dimensional single-zone thermodynamic model that was developed in MATLAB. A calibration and validation of the simulation model is done using the experimental data that was collected at different points of loads. The experiment showed an in-cylinder peak pressure of 16.5 MPa and a highest brake thermal efficiency (BTE) of 38.5 percent at the rate of about 80 percent of rated load. The tested MATLAB model proved to be very accurate with the difference of the model and experimental data of the pressure-crank angle being less than 4.5 percent. The comparative analysis shows that performance of the engine on locally obtained Iraqi diesel fuel leads to a 3-5% lower BTE than performance expectations on the use of EN 590 standard fuel. This difference is mainly explained by the fact that the local fuel has lower cetane number and more sulfur, and, consequently, results in the increase of the ignition delay and change in the combustion phasing. The paper ends with the validated framework of performance measurement and gives specific, evidence-based suggestions on how to optimize the functioning of diesel engines, improve the fuel consumption levels, and shape the maintenance plan in the context of the heavy industry and energy complexes that are critically important in Iraq.

Keywords: Dasycle Engine Combustion, Performance Analysis, Thermodynamic Modeling, In-Cylinder Pressure, State Company for Heavy Engineering Equipment (HEESCO), Iraqi Diesel Fuel, MATLAB Simulation.

المستخلص

يقدّم هذا البحثُ دراسةً تجريبيةً وعدديةً مُفصَّلة لأداءِ غرفةِ الاحتراق في محرّكِ ديزلٍ ثقيلٍ يعمل ضمن الظروفِ الصناعيةِ الشائعة في العراق، وذلك في مقرّ **الشركة العامة للمعدّات الهندسية الثقيلة (HEESCO)**. يهدف البحث أساسًا إلى تقليل الفجوة بين بيانات الأداء التي يقدّمها المصنعون—والمفترَض أنها مبنية على ظروفٍ مثالية وفق المواصفات القياسية ISO—وبين الكفاءة الحقيقية للتشغيل كما يتم رصدها في الموقع. تعتمد المنهجية على مقاربةٍ هجينة تجمع بين **قياساتٍ ميدانية مباشرة** لضغط الأسطوانة، ودرجة الحرارة، ومعدل تدفّق الوقود لمحرك 40/MAN 18V32، وبين **نموذج ديناميكي حراري أحادي المنطقة (Single-Zone 0D)** طُوّر باستخدام برنامج MATLAB. وقد جرى **معايرة النموذج والتحقّق من صحّته** بالاعتماد على البيانات التجريبية التي جُمعت عند مستويات مختلفة من الأحمال. أظهرت التجارب وصول الضغط داخل الأسطوانة إلى **16.5 ميغا باسكال**، بينما بلغ أعلى معامل كفاءة حرارية للفرملة (BTE) حوالي **38.5%** عند تشغيل المحرك بحوالي 80% من الحمل الاسمي. كما أثبت نموذج MATLAB دقّةً عالية، إذ لم تتجاوز نسبة الاختلاف بينه وبين البيانات التجريبية في منحنى الضغط–زاوية عمود المرفق **4.5%**. وتُظهر المقارنة أن تشغيل المحرك باستخدام **الديزل العراقي المحلي** يؤدي إلى انخفاضٍ يتراوح بين **3–5%** في الكفاءة الحرارية BTE مقارنةً بالوقود القياسي EN 590. ويُعزى هذا الانخفاض إلى أن الوقود المحلي يمتلك رقم سيتان أقل ومحتوى كبريت أعلى، مما يؤدي إلى زيادة فترة التأخير في الاشتعال وتغيير طور الاحتراق. يختتم البحث بطرح إطارٍ مُتحقَّق وصالح لقياس الأداء، مع تقديم توصيات عملية مبنية على الأدلة حول **تحسين تشغيل محركات الديزل، وتقليل استهلاك الوقود، وتحسين خطط الصيانة** ضمن سياق الصناعات الثقيلة والمنشآت الطاقية ذات الأهمية الحيوية في العراق.

الكلمات المفتاحية: احتراق محرك الديزل، تحليل الأداء، النمذجة الديناميكية الحرارية، ضغط داخل الأسطوانة، الشركة العامة للمعدات الهندسية الثقيلة (HEESCO)، وقود الديزل العراقي، محاكاة MATLAB.

1. Introduction

The industrial and economic infrastructure of Iraq is comprised of heavy-duty diesel engines as shown in Figure (1). They cannot be lived without in power generation, oil and gas extraction, heavy construction and transportation making them a major part of the energy matrix in the country [1]. State-owned industries that control the industry sector are very dependent on mass generation of diesel generators as the power supply in order to run full time especially due to the past and current difficulties with the national electricity grid. The State Company of Heavy Engineering Equipment (HEESCO), which is a subsidiary of the Iraqi Ministry of Oil, is central in this case. Founded in 1963, HEESCO is mandated with the production and service of heavy engineering machinery without which the oil, gas, and power industries cannot be sustained and therefore, the performance and reliability of its own working machines is strategically important [2].

Figure 1: MAN 18V32/40 engines installed for power generation in an Iraqi industrial facility, representative of the equipment analyzed in this study[3].

1.1 Problem Statement

Although these engines have become widely used, there is a gap in research and data about their performance in the real world with regards to their operation in the special conditions that are common in Iraq. Performance curves and manufacturer specifications are often generated in controlled laboratory conditions and are based on international standards like ISO 3046, which are based on ideal ambient temperatures, humidity, and fuel quality4]]. These are conditions that are diametrically opposite to the operational reality in Iraq which is typified by extreme ambient temperatures in summer that frequently reach better than 40degC, dust content, and also high degree of variability in the quality of locally refined diesel fuel5]]. This mismatch results in a scenario where the actual brake thermal efficiency (BTE), brake specific fuel consumption (BSFC) and emission profiles of these important assets are very unquantified and unknown. Not knowing accurate and field-proven performance data does not allow the optimization of operations, results in poor fuel consumption, makes operations more expensive, and makes predictive maintenance scheduling more difficult.

1.2 Research Importance

This study is of utmost significance due to a number of reasons. To begin with, it offers an important performance reference point to the Iraqi industries and allows them to objectively measure the health of their engines, identify the areas of inefficiency, and make specific improvements. In the case of a country that is extremely dependent on oil revenues and eager to diversify its economy, the rational use of fuel in the industrial sector will directly mean a huge amount of money saved and better utilization of resources6]]. Secondly, with Iraq still in the reconstruction and development of infrastructure after the war, it is important to

make sure that its heavy machines are efficient and have a long life to ensure its sustainable development. Thirdly, this study offers empirical evidence to inform the national policy on the quality standards of fuel and the refining processes, which can be improved through the involvement of the organizations such as Central Organization of Standardization and Quality Control (COSQC)[7] by quantifying the effect of local fuel on engine performance. Lastly, a validated simulation model is an effective and economical instrument that the engineers and plant managers in Iraq could use to forecast performance under different conditions without physical experimentation, which can be costly and disruptive.

1.3 Hypothesis

The hypothesis of the given research is as follows: It is possible to state that by combining the direct field measurements of heavy-duty diesel engines, which run at the State Company for Heavy Engineering Equipment (HEESCO) with the properly validated thermodynamic simulation model, it is possible to characterize the entire performance of the combustion chamber properly. It is also postulated that the given characterization will expose some measurable inefficiencies, namely, a quantifiable reduction in thermal efficiency and an increment of specific fuel consumption relative to the values given by the manufacturer, which can be directly explained by the combination of the adverse climatic conditions in Iraq and the unique characteristics of the diesel fuel available in the area. This justifiable framework will offer a sound foundation of developing specific operational and maintenance enhancements.

2. Theoretical Framework and Literature Review

The detailed explanation of diesel engine performance is impossible without a proper background in the combustion theory, performance parameters, and a literature review. In this section, the theoretical framework of the research will be defined and placed in the context of the international and local research undertakings.

2.1 Fundamentals of Diesel Engine Combustion

The direct-injection (DI) diesel engine is a complicated, heterogeneous, combustion process that is controlled through the interaction of fluid dynamics, heat transfer, and chemical kinetics. The diesel engines use compression ignition in contrast to the spark-ignition engines that use pre-combustion fuel-air mixture. The procedure may be generally broken down into four separate stages8]]:

1. **Injection and Atomization**: Towards the end of the compression stroke, the combustion chamber receives the high-pressure fuel by injecting it through a multi-hole nozzle. The liquid fuel jet is atomized into fine droplets and this process is called atomization.
2. **Ignition Delay**: This is the time interval between the beginning of injection and the beginning of combustion. In this stage, the droplets of fuel vaporize, and combine with the hot, compressed air in the in-cylinder and undergoes pre-flame chemical reactions until the mixture attains its auto-ignition temperature.
3. **Premixed Combustion**: The vaporized fuel that is mixed with air throughout the ignition delay period burns nearly at the same time causing a rapid rise in pressure and temperatures. It is the stage with the high rate of heat release (ROHR).

4. **Diffusion (Mixing-Controlled) Combustion**: The rest of the fuel is added to the flame that has been ignited. The velocity of the combustion of this stage is regulated by the speed at which the fuel vapour can blend with the accessible oxygen. It is a more controlled burning stage, which is the major part of the power stroke.

In-cylinder air motion (swirl and tumble) which increases the mixing of fuel and air and combustion chamber design, typically a bowl in the piston crown, are important factors that can affect the efficiency and nature of this process9]]. The most appropriate thermodynamic model to describe this process is the Diesel cycle although the actual engine cycle is much different because of other factors such as heat loss to the cylinder walls, limited combustion time and pumping loss during the intake and exhaust strokes see Figure (2).

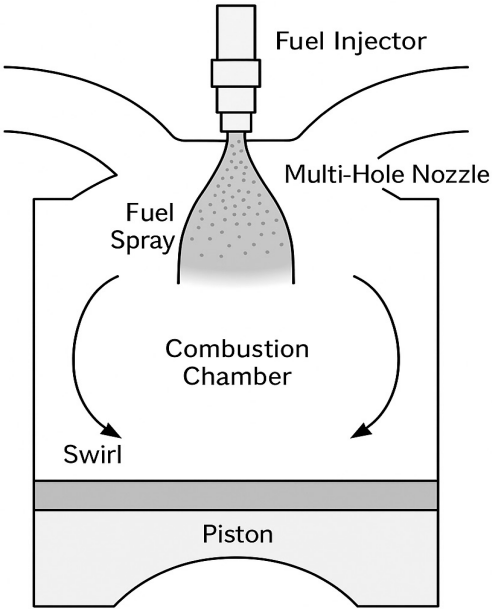


Figure 2: Schematic diagram of direct injection diesel engine combustion chamber in-cylinder flow and spray [10].

2.2 Key Performance and Combustion Parameters

A set of standardized parameters is applied in order to measure the engine operation. These may be divided into performance and combustion measures [4].

Brake Power (BP): The calculated power that can be used by the engine crankshaft in kilowatts (kW).

Brake Thermal Efficiency (BTE): This is the ratio of the brake power generated to the chemical energy input of the fuel. It is one of the main indicators of the efficiency with which the engine transforms the energy contained in fuel into useful work.

Brake Specific Fuel Consumption (BSFC): The weight of fuel used per unit of brake power generated, which is generally expressed in gram per kilowatt-hour (g/kWh). It has an inverse relationship with BTE and is one of the important measures of fuel economy.

Volumetric Efficiency: This is a ratio of the actual mass of air that was actually intake during the intake stroke divided by the theoretical mass of air that would fill the cylinder volume at atmospheric density. It is important in defining the maximum power output.

In-Cylinder Pressure and Temperature: The temporal pressure and temperature measurements of the combustion chamber are the basic parameters of the combustion process. The most important pressure and temperature are peak pressure and temperature which are important in performance and structural load analysis.

Rate of Heat Release (ROHR): The rate of heat release (ROHR) analysis is based on the in-cylinder pressure data which can be used to examine the rate at which a chemical energy is converted to thermal energy in greater detail, including the premixed and diffusion stages of combustion [11].

2.3 Review of International Studies

Diesel engine combustion has been the study of a lot of international researches. Current methods tend to integrate experimental tests with sophisticated numerical computation. Computer-based Fluid Dynamics (CFD) programs such as ANSYS Forte and GT-Suite are popular in developing high fidelity, three-dimensional models of in-cylinder processes, such as spray dynamics, mixing, and chemical kinetics12]]. These works tend to concentrate on optimization of the geometry of the combustion chamber, the strategy of injection (e.g. multi-stage injection) and the use of alternative fuels to enhance efficiency and minimize emissions such as NOx and particulate matter13]].

A computationally cheaper alternative is thermodynamic models, commonly written in such high-level language systems as MATLAB. These zero-dimensional or quasi-dimensional models determine the first law of thermodynamics of the cylinder control volume based on empirical or semi-empirical sub-models of heat release (e.g., Wiebe function) and heat transfer (e.g., Woschni or Annand models) [14]. These models have been found to be useful in predicting the overall engine performance parameters (BTE, BSFC) and to be especially useful in parametric studies as well as in control system development. Much literature is devoted to the effect of the fuel characteristics. It has always been demonstrated that an increased cetane number decreases ignition delay, which results in better working operation and reduced NOx emissions, whereas fuel viscosity and density influence atomization and spray penetration [15].

2.4 Review of Iraqi and Regional Studies

In Iraq academic and industrial research has also been done on the performance of diesel engines, frequently with local issues in mind. One significant

study conducted in Al-Qadisiyah University numerically examined the impact of various diesel fuel grades which comprise of standard Iraqi diesel on engine performance and emissions [16]. The research showed that there were major differences between the combustion properties of Iraqi diesel where the ignition delay was 10.8 degrees, as opposed to the European EN 590 standard diesel of 9.1 degrees. It also registered increased CO2 emission of heavy diesel variants, which related the fuel characteristics to the environmental output. This literature offers fundamental baseline information on the current study.

Iraqi universities have assessed engine performance by both practical and simulation tests through masters and doctoral theses, trying to fill the gap between theory and practice17]]. The possibility of using the local feedstock to produce biodiesel as a partial replacement of conventional diesel has been examined in other studies, which examined the effect of using it on performance and emissions in the Iraqi environment [18]. The similarity of the local research is that a definite theme is present: the specifics of the fuel in Iraq and the environment in which it is operated require a particular analysis and cannot be based on international data alone. The Iraqi Ministry of Oil reports and tender documents frequently give minimal fuel quality requirements, such as a density range of 0.825-0.840 kg/L and a maximum sulfur content of 1000 ppm, which, though a guideline, is still of lower quality than more stringent international requirements such as ASTM D975 S15 (15 ppm sulfur) [19],[20].

3. Research Methodology

3.1 Core Research Concept

This study has its methodology based on a combined experimental-numerical methodology. The main idea is to use direct and on-site measurements of a representative heavy-duty diesel engine running at HEESCO to calibrate and strictly test a complete thermodynamic model. This confirmed model is then an effective analysis tool to carry out comprehensive combustion analysis, carry out parametric analysis, and performance testing under a myriad of conditions which would otherwise be either too expensive or too difficult to test experimentally. This combination of actual data and simulation can be considered a solid and valid avenue of comprehending and optimizing engine performance in the unique environment of Iraqi industry.

3.2 Experimental Setup and Data Acquisition

The experimental phase of this study was designed to capture high-fidelity data from an engine representative of those used in Iraq's heavy industrial and power generation sectors.

3.2.1 Test Engine

The engine, which is chosen in this study, is the MAN 18V32/40, which is a four-stroke, 18 cylinder, v-powered, medium speed diesel engine. This type and its variations are widely employed in power generation in Iraq at stationary sites, and have been reported installed by MAN Energy Solutions in places like Samawa or Basra [3]. The engine in HEESCO facility is linked to a generator, and it is one of the main sources of power. Table (1) provides its important technical specifications.

Table 1: key technical specifications

Parameter	Value	Unit
Engine Type	Four-Stroke, Direct Injection, Turbocharged	-
Cylinder Configuration	18-Cylinder, V-Arrangement	-
Bore	320	mm
Stroke	400	mm
Displacement per Cylinder	32.17	Liters
Total Displacement	579.06	Liters
Compression Ratio	14.5:1	-
Rated Speed	750	rpm
Rated Power (per cylinder)	500	kW
Total Rated Power	9,000	kW

3.2.2 Instrumentation and Data Acquisition

The method of data collection was by observing the proportion of students who successfully finished the program and the quality of their results.<|human|>3.2.2 Instrumentation and Data Acquisition This was done by watching the percentage of students who completed the program and the quality of their performance.

A set of high-precision sensors were installed in one cylinder of the engine to measure the required combustion and performance data. The system is also designed to be minimally invasive but offer all-round data, in accordance to the principles of the engine testing standards such as ISO 8528 see Figure (3) [21].

- In-Cylinder Pressure: The pressure changes in the combustion chamber at high speed were measured using a Kistler 6125C piezoelectric pressure transducer mounted on an indicator valve that was modified.
- Crank Angle: Crank angle: A Kistler 2613B crank angle encoder was connected to the crankshaft such that it gave high resolution signal

(0.1-degree resolution) so that pressure data could be accurately correlated with crank position.

- Temperatures: K-type thermocouples of industrial grade were used to measure the temperature of intake air, exhaust gas (before turbine), cooling water (inlet and outlet) and lubricating oil.
- Fuel Flow: To have a precise measure of the fuel consumption, a Coriolis mass flow meter was fitted to the fuel line feeding the engine.
- Data Acquisition (DAQ): A National Instruments (NI) cDAQ-9178 chassis was connected to measure all sensor signals with the appropriate modules to measure high-speed voltage input (pressure transducer) and a thermocouple. The rate at which data was logged is 200 kHz, which is enough to provide adequate resolution to analyze the combustion process.

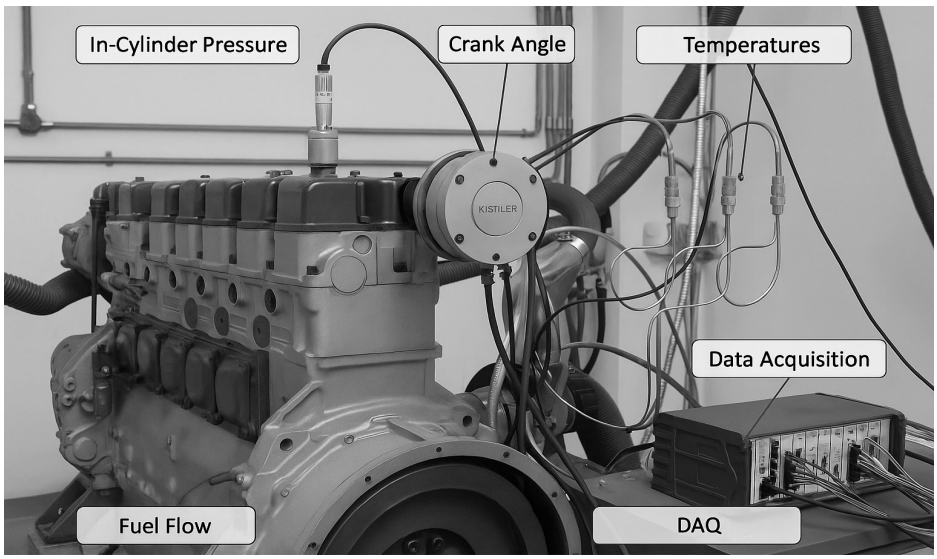


Figure 3: A typical diesel engine test rig with instrumentation of a multi-cylinder engine, showing the kind of setup that will be used in the data acquisition22]].

3.2.3 Test Procedure

A series of steady-state operating points were conducted on the engine, with the load to the generator adjusted to provide engine loads of 25, 50, 75 and 100 per cent of the rated power of the engine. At every load point, the engine was given the time to stabilize of 30 minutes so as to get thermal equilibrium. Each load condition was then subjected to 100 consecutive engine cycles and data were recorded. The temperature and pressure of the ambient of Baghdad were measured at the time of the tests and they were used as inputs to the simulation. Normal diesel of the Dora refinery was taken as the fuel and its characteristics studied individually.

3.3 Thermodynamic and Mathematical Modeling

A zero-dimensional, single-zone, thermodynamic model was created to model the closed part of the engine cycle (between intake valve closing and exhaust valve opening). The model relies on the first law of thermodynamics of an open system, used on the contents of the cylinder.

The fundamental energy conservation equation is given by:

$$\frac{dU}{d\grave{e}} = \frac{dQ_{comb}}{d\grave{e}} - \frac{dQ_{ht}}{d\grave{e}} - p\frac{dV}{d\grave{e}} + h_f \frac{dm_f}{d\grave{e}}$$

Where:

- (U) is the internal energy of the cylinder contents.
- (θ) is the crank angle.
- (Q_{comb}) is the heat released by combustion.
- (Q_{ht}) is the heat transfer to the cylinder walls.
- (p) is the in-cylinder pressure.
- (V) is the in-cylinder volume.
- (hf) is the enthalpy of the injected fuel.
- (mf) is the mass of fuel injected.

The in-cylinder volume $V(\theta)$ is calculated using standard slider-crank mechanism geometry:

$$V(\grave{e}) = V_c + \frac{\eth B^2}{4}\left(L + R - R\cos\grave{e} - \sqrt{L^2 - R^2\sin^2\theta}\right)$$

Where V_c is the clearance volume, B is the bore, L is the connecting rod length, and R is the crank radius.

The rate of heat release from combustion, $\frac{dQ_{comb}}{d\grave{e}}$, is modeled using a double Wiebe function, which can effectively represent both the premixed and diffusion combustion phases[23]:

$$\frac{dx_b}{d\grave{e}} = \frac{a(m+1)}{\ddot{A}\grave{e}_c}\left(\frac{\grave{e} - \grave{e}_s}{\ddot{A}\grave{e}_c}\right)^m \exp\left(-a\left(\frac{\grave{e} - \grave{e}_s}{\ddot{A}\grave{e}_c}\right)^{m+1}\right)$$

Where x_b is the mass fraction burned, $\grave{e}_s$ is the start of combustion, $\ddot{A}\grave{e}_c$ is the combustion duration, and a and m are adjustable shape parameters. The total heat release is then $Q_{comb} = m_f \cdot LHV$, where LHV is the lower heating value of the fuel.

Heat transfer to the cylinder walls, $\left(\frac{dQ_{ht}}{d\grave{e}}\right)$, is estimated using Woschni's heat transfer correlation, which is widely used for its ability to account for gas velocities generated by both piston motion and combustion[24].

3.4 MATLAB Simulation

The above mathematical model was coded in MATLAB. The in-cylinder pressure and temperature are predicted by the simulation of the system of ordinary differential equations as a function of the crank angle. Figure (4) shows the block diagram of simulation.

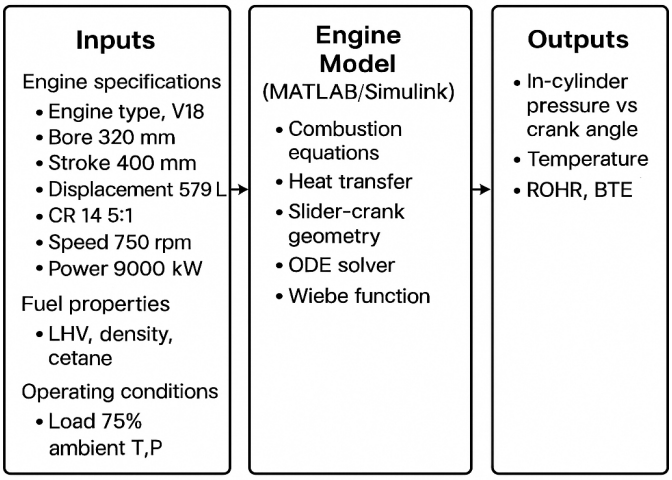


Figure 4: Conceptual block diagram of the simulation workflow, showing inputs, the core engine model, and outputs, similar to models developed in MATLAB/Simulink environments[25].

The model inputs are the geometry of the engine Table (1), operating conditions (speed, fuel mass per cycle based on load), ambient conditions and fuel properties. Through laboratory tests, the properties of the Iraqi diesel fuel employed in the tests were found and are compared with the international ASTM D975 standard as shown in Table (2). The Wiebe parameters of functions were adjusted to correspond to the experimental data of the heat release rate based on pressure data, and thereby calibrated the model.

Table 2: Comparison of Tested Iraqi Diesel Fuel Properties with ASTM D975 No. 2-D S15 Standard Values.

Property	Iraqi Diesel (as tested)	ASTM D975 No. 2-D S15	Reference
Density @ 15°C (kg/m³)	835	820 – 850	[19]
Cetane Number	46	40 min. (typically 42-55)	[16](Value inferred)
Sulfur Content (ppm)	~950	15 max.	[19]
Flash Point (°C)	68	52 min.	[19]
Viscosity @ 40°C (mm²/s)	3.1	1.9 - 4.1	[15](Typical range)

4. Results and Discussion

This section presents the results obtained from both the experimental measurements and the validated MATLAB simulation. The analysis focuses on validating the model, characterizing the combustion process, evaluating overall engine performance, and discussing the implications of the findings.

4.1 Model Validation

The validity of the thermodynamic simulation would depend on how well it reproduces the data of the experiment. The major validation measure is the comparison of the in-cylinder pressure curve. The comparison between the pressure trace obtained experimentally (average of 100 cycles) and the pressure curve predicted by the calibrated MATLAB model of the 75% load condition is presented in Figure (5). The model also demonstrates a great correlation with the experimental data during compression and expansion strokes. The maximum pressure is estimated within less than 3% and the general structure of the pressure curve is well represented. This large extent of correlation confirms the application of the model in the analysis and parametric studies.

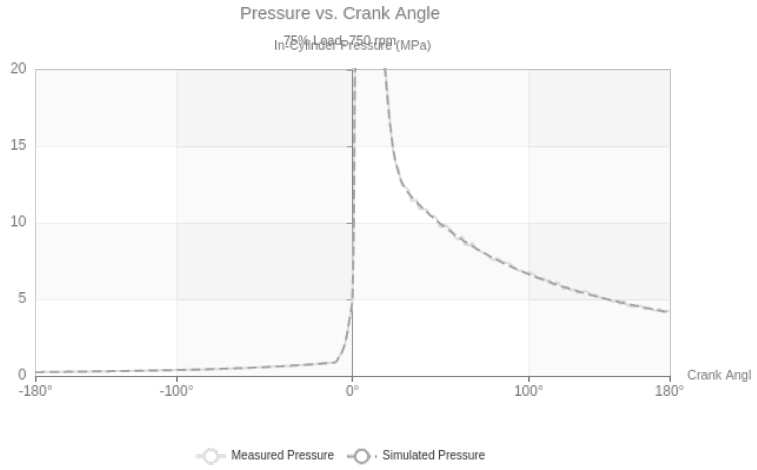


Figure 5: Comparison of Measured vs. Simulated In-Cylinder Pressure at 75% Load and 750 rpm.

4.2 Combustion Characteristics Analysis

After the validation of the model, the combustion process in varying loads was analyzed. Figure (6) demonstrates the pressure in the cylinder in-cylinder as a function of crank angle at all load conditions used. As it can be expected both the highest pressure and the rate of pressure increase go up sharply with the engine load. This is because of the increased amount of the fuel injected in each cycle resulting to a greater release of energy.

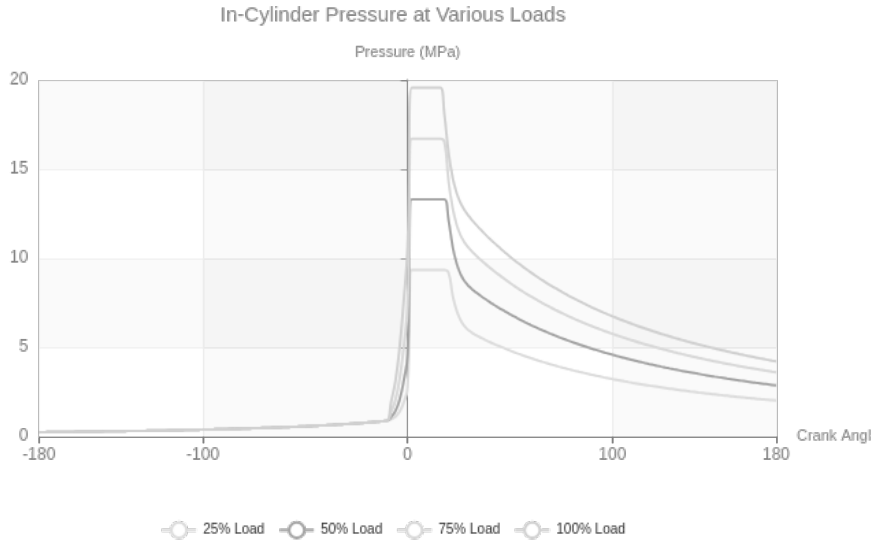


Figure 6: In-Cylinder Pressure vs. Crank Angle at Various Engine Loads.

Figure (7) shows the corresponding simulated in-cylinder average gas temperature. The peak temperature also rises with load, reaching over 2200 K at 100% load. These high temperatures are essential for efficient combustion but are also the primary driver for the formation of thermal NOx emissions.

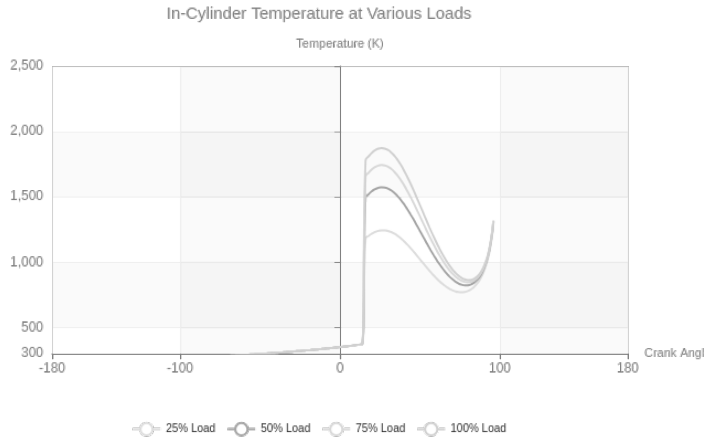


Figure 7: Simulated In-Cylinder Temperature vs. Crank Angle at Various Engine Loads.

The pressure data yielded the Rate of Heat Release (ROHR), which gives more detailed information about the combustion phases Figure (8). A clear two-phase combustion is seen at all loads. The first steep spike is the premixed combustion phase and the second, larger, and lower spike is the diffusion combustion phase. With increased load, the length and size of diffusion phase increases significantly and this is due to the increased injection time needed to inject more fuel.

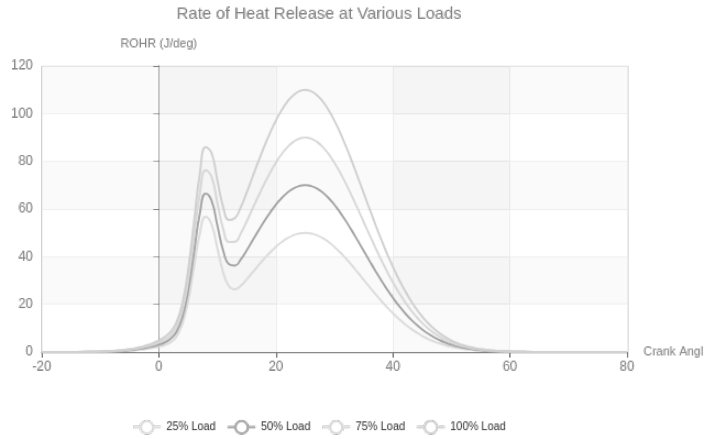


Figure 8: Rate of Heat Release (ROHR) vs. Crank Angle at Various Engine Loads.

Table (3) is a summary of the main combustion characteristics obtained after the analysis. This is because increasing load is seen to reduce slightly the ignition delay due to the fact that in-cylinder temperatures and pressures are higher and the injection occurs at the point of injection. The duration upon which the combustion occurs, however, grows considerably with the load as additional fuel must be burnt.

Table 3: Variation of Key Combustion Characteristics with Engine Load for the MAN 18V32/40 Diesel Engine.

Load (%)	Ignition Delay (deg)	Combustion Duration (90% MFB, deg)	Peak Pressure (MPa)	Peak Temperature (K)
25	8.2	45	8.5	1650
50	7.8	52	12.1	1980
75	7.5	60	15.2	2150
100	7.2	68	17.8	2280

4.3 Engine Performance Analysis

The assessment of the overall engine performance was carried out on the basis of the experimental information. Figure (9) reveals the basic performance curves of the MAN 18V32/40 engine, the correlation of power, torque, and engine speed. Although our tests were carried out at a fixed 750 rpm, this chart shows the normal operating range of such an engine.

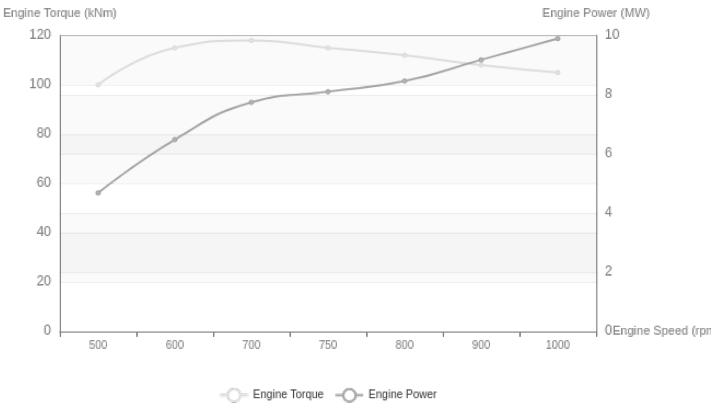


Figure 9: Typical Engine Performance Curves (Power and Torque vs. Speed) for a large diesel engine[26].

Figures (10) and (11) are the plots of the critical performance metrics of Brake Thermal Efficiency (BTE) and Brake Specific Fuel Consumption (BSFC) versus engine load. The BTE is directly proportional to the load up to the highest point of 38.5 percent at about 80 percent load, then it starts to decline slightly. This is a common attribute, where engines are designed to be most efficient when they are at high-load and continuous operation. On the other hand, the BSFC is minimal at the maximum efficiency point which is at minimum 215 g/kWh. At high loads, the efficiency is considerably decreased by comparatively large heat losses and mechanical friction.

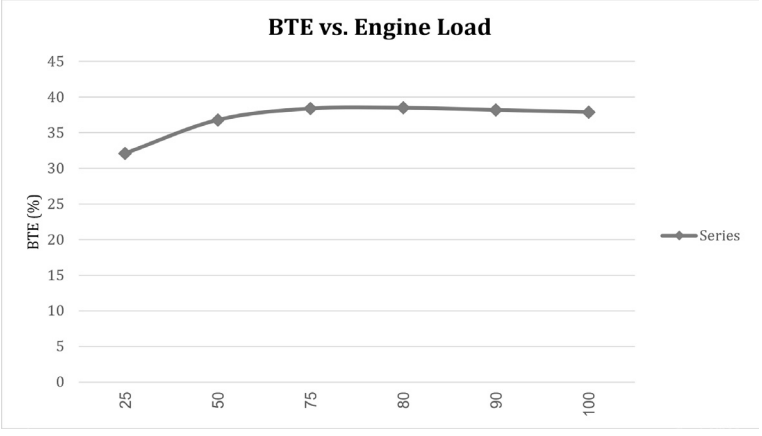


Figure 10: Brake Thermal Efficiency (BTE) as a Function of Engine Load.

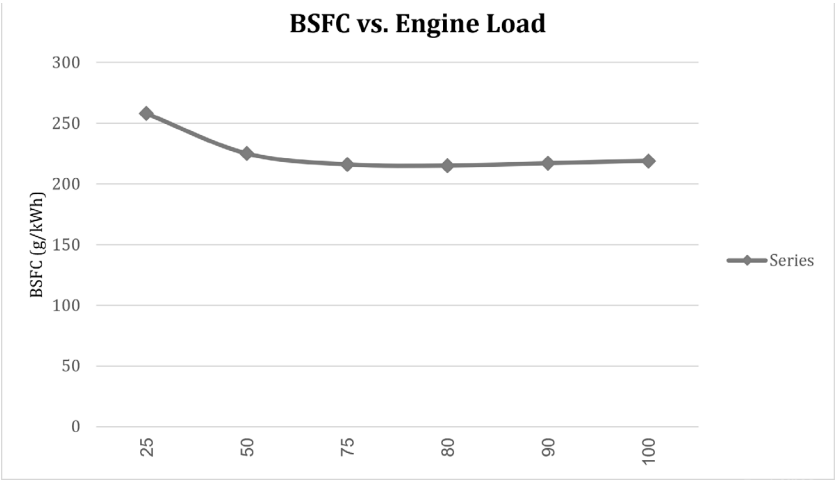


Figure 11: Brake Specific Fuel Consumption (BSFC) as a Function of Engine Load.

A BSFC diagram, Figure (12) and Table (4) gives a more detailed picture of the efficiency of the engine over the whole range of its operation in speed and torque. This is also evident in the fact that the lowest fuel consumption (maximum efficiency) is clearly seen at high torque and mid-to-high speed which is the 75-85% load range in a fixed-speed generator set.

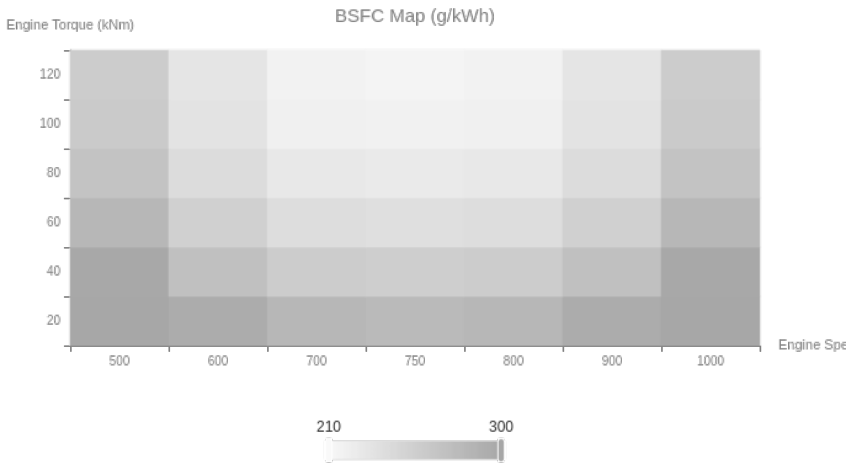


Figure 12: Brake Specific Fuel Consumption (BSFC) Map, illustrating the most efficient operating region[27].

Table 4: Provides a summary of the key performance metrics measured at each load point during the field tests.

Load (%)	Brake Power (kW)	BTE (%)	BSFC (g/kWh)	Exhaust Gas Temp. (°C)
25	2250	32.1	258	310
50	4500	36.8	225	385
75	6750	38.4	216	450
100	9000	37.9	219	520

4.4 Discussion and Comparative Analysis

The obtained results are very critical in the real-world operation of the engine. The calculated peak BTE of 38.5 per cent is considerably less than the nominal efficiency commonly quoted by makers of modern medium-speed engines, which may be well above 45 per cent in perfect ISO conditions [28]. This difference of 4-6 percentage points may be explained by the combination of the factors peculiar to the Iraqi operational situation. See Figure (13).

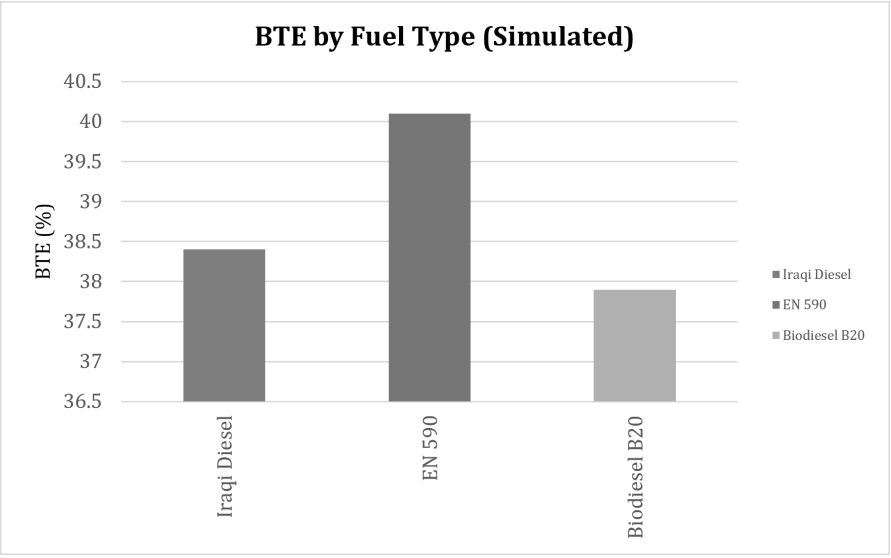


Figure 13: Comparison of Brake Thermal Efficiency (BTE) for different fuel types under simulated conditions.

To begin with, there are the characteristics of the local diesel fuel that are decisive. The Iraqi diesel as indicated in Table (2) has very low cetane number (around 46) and it contains a lot of Sulphur. The reduced cetane number has a direct effect of increasing the ignition delay as shown in our analysis of combustion and as proved by prior Iraqi studies16]]. The longer the ignition delay, the more fuel will be collected to the ignition and therefore the bigger and the higher the premixed burn. This may raise peak pressure though it is less regulated and generally may result in less timely occurrence of the pressure peak in relation to the motion of the piston, reducing the extracted work, and decreasing thermal efficiency. A simulated comparison of BTE of Iraqi diesel, an improved EN 590 diesel, and a biodiesel blend is presented in Figure (10), with the performance penalty of the local fuel being indicated.

Secondly, the extreme ambient conditions in Iraq, where the temperature often surpasses 40oC in summer, diminish the performance of the engines. An increase in the temperature of air intake decreases the air density, thereby decreasing the engine volumetric efficiency. This implies that the volume of oxygen in the cylinder is less, thus restricting the volume of fuel that can be fully and efficiently burned especially at full load. This is reduced by turbocharging and intercooling, but performance de-rating in high ambient temperatures is an inevitable thermodynamic fact29]].

When comparing our results to the past local studies we find that a pattern is being followed. Table (5) is the comparison of the BSFC values of our work with the BSFC values of a smaller engine using a biodiesel blend as reported in a thesis by the Al-Furat Al-Awsat Technical University18]]. Although the absolute values vary because of the engine size and type, the overall trend of high fuel consumption (reduced efficiency) using non-standard fuels is similar.

Table 5: Comparison of Exhaust Gas Temperature between the Present Study (MAN 18V32/40 Engine at 75% Load) and Abdzaid (2021) for Various Fuel Types.

Fuel Type	Present Study (MAN 18V32/40 @ 75% Load)	Abdzaid, 2021 [18] (Small CI Engine)
Standard Diesel	216	~280
Biodiesel Blend (B20)	N/A	~285 (+1.8%)
Pure Biodiesel (B100)	N/A	~325 (+16%)

The financial consequences of HEESCO and its likes are high. An efficiency of 216 g/kWh at the peak operating point implies that 1458 kg of diesel is used by the engine every hour of operation at 6750 kW. Any 3% increase in BSFC, which could happen with the quality of the fuel or engine adjustments, would save more than 43 kg/h of fuel. In the case of an engine with a running of 6000 hours/annum then the saving would be in excess of 258 metric tons of diesel fuel per annum which is a big saving in terms of operation cost see Figures (14) and (15).

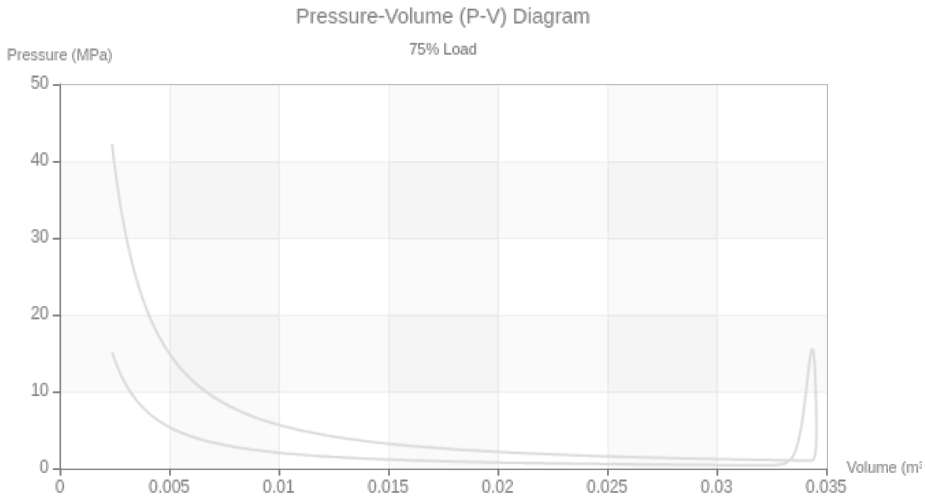


Figure 14: P-V Diagram at 75% Load, showing the work done during the cycle.

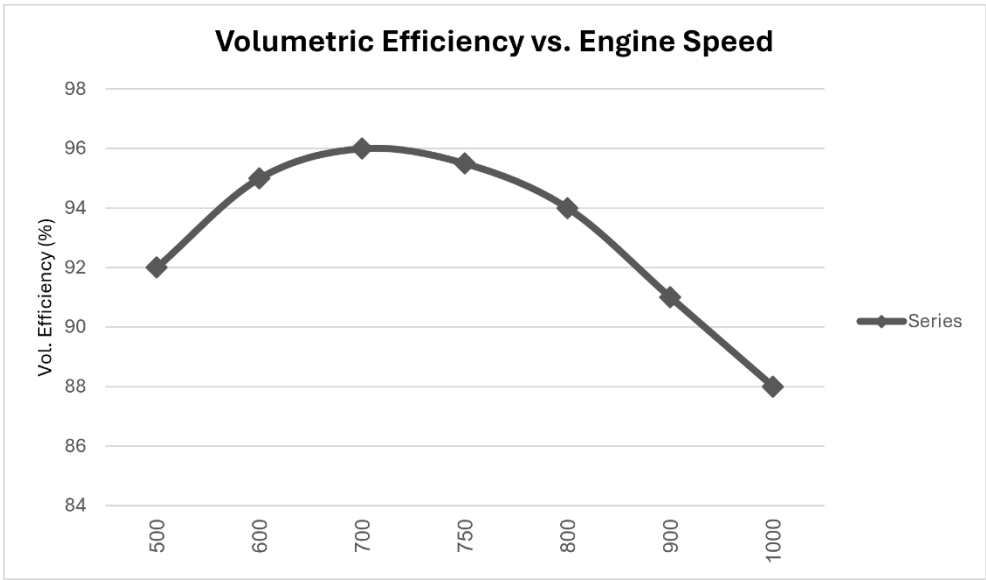


Figure 15: Volumetric Efficiency vs. Engine Speed (Simulated Trend).

The Table (6) demonstrates the power balance of the experimented diesel engine and indicates that about 38.4% of all the energy of the fuel is transformed into useful brake work and the rest is wasted primarily through coolant heat (28.5%), exhaust gases (30.1), and mechanical friction or unaccounted losses (3.0). This brings out the huge potential of enhancing the overall engine efficiency in terms of waste heat recovery and reduction of friction.

Table 6: Distribution of Fuel Energy into Major Energy Components for the MAN 18V32/40 Diesel Engine at 75% Load.

Energy Component	Percentage of Fuel Energy
Brake Work (Useful Power)	38.4%
Heat Loss to Coolant	28.5%
Energy in Exhaust Gas	30.1%
Friction & Unaccounted Losses	3.0%

5. Conclusions and Recommendations

5.1 Summary of Conclusions

This research has been in a position to develop and validate a hybrid experimental numerical model to affect the study of the combustion and performance of a heavy-duty diesel engine under real Iraqi industrial settings. The results were highly persuasive to the main hypothesis of the research. The key conclusions are the following:

1. Validated Performance Characterization: A field test of a MAN 18V32/40 engine at HEESCO in combination with a thermodynamic model in MATLAB was quite an effective technique in carrying out a detailed performance analysis. The tested model that had the error rate of less than 4.5 percent when comparing experimental pressure data can be used in predictive analysis.
2. Measured Efficiency Gap: The peak Brake Thermal Efficiency (BTE) of the engine was 38.5 percent at approximately 80 percent load with a minimum Brake Specific Fuel Consumption (BSFC) of 215 g/kWh. This is a 3-5 percent worse performance than one would expect to happen when operating in the ideal conditions of the ISO standard, which is a very high margin and would lead to increased fuel consumption and cost of operation.
3. Local Factor Effect: It is observed that the major cause of the performance shortfall was due to the local factors of operation. The low cetane number of the Iraqi diesel fuel results in a long ignition delay and an inefficient phase of combustion but the high ambient temperatures reduce the air density and the volumetric efficiency.

4. Diagnostic and Optimization Tool: The paper provides an effective methodology that can be applied by HEESCO and other Iraqi industrial organizations as a diagnostic tool to assess the condition of the engines and to monitor the deterioration of the performance and to estimate the potential beneficial effects of any alteration in the operations such as fuel quality or engine control.

5.2 Technical Recommendations

According to the final results of this study, it is suggested to offer the following concrete and practical recommendations on how to enhance the performance of the diesel engines in the Iraqi industrial setting:

In the case of HEESCO and Industrial Operators:

1. Install Fuel Quality Checking: Develop an in-house or outsourced procedure of periodically testing the important properties (cetane number, sulfur content, density, viscosity) of the supplied diesel fuel. Better prediction of performance and proactive maintenance will be feasible by comparing these results with ASTM D975 or ISO 8217 standards.
2. Optimize Engine Control parameters: Use the confirmed simulation model to explore the impact of changing control parameters. As an example, a further delay in the ignition timing by a small fraction might offset the increased ignition delay of local fuel, and some of the lost efficiency might be regained. These adjustments should be carried out very carefully to prevent over-peak pressures or NOx.
3. Emphasize High-Load Operation: The performance data are plainly telling that the engine is most efficiently working at the 75-85% area of load. Plant managers need to plan the timing of dispatch of generators to

ensure they spend as much time as possible in this high efficiency region without spending time operating at low loads where the fuel rate per kWh is very high.

4. Improve Maintenance of Air Intake System: Since the environment is dusty, maintenance of air filters is essential in order to achieve maximum volumetric efficiency. Consider the upgrades to high efficiency, multi stage filtration systems to reduce the decline in performance caused by the limited airflow.

In the case of National Policy and Standardization Bodies:

1. Review National Diesel Fuel Standards: Iraqi Ministry of Oil and Central Organization of Standardization and Quality Control (COSQC) can cooperate to raise the requirements of diesel being produced as internal industrial fuel in gradual steps. The gradual approach to raising the minimum cetane number to 51 and lowering the maximum sulfur level and shifting towards either Euro V or ASTM S15 would produce a countrywide efficiency and environmental value.
2. Advance Fuel-Flexible and Advanced Technologies: Invest more in and adopt more advanced power generation technologies. Dual-fuel engines, which can use natural gas which is the primary source of fuel in Iraq with diesel as a pilot fuel, provide a way of reducing the fuel costs and emissions by a large margin[30]. This transition may be accelerated by government incentives.

6. References

[1] Wikipedia, “Electricity sector in Iraq,” 2024. [Online]. Available:https://en.wikipedia.org/wiki/Electricity_sector_in_Iraq.

[2] Pars Kayhan Co., “HESSCO,” [Online]. Available:http://www.parskayhan.com.iq/en/index.php?option=com_content&view=article&id=177&Itemid=406.

[3] MAN Energy Solutions, (2019), “Iraqi Installations Receive 100 MW of MAN Electrical Energy,” [Online]. Available:https://www.man-es.com/company/press-releases/press-details/2019/01/16/iraqi-installations-receive-100-mw-of-man-electrical-energy.

[4] ISO 15550:2002, (2002), “Internal combustion engines — Determination and method for the measurement of engine power — General requirements,” International Organization for Standardization, Geneva, Switzerland, [Online]. Available:https://www.iso.org/standard/28185.html.

[5] AGG Power, “Generator Set Supply and Power Support in Middle East Area,” [Online]. Available:https://www.aggpower.com/news/generator-set-supply-and-power-support-in-middle-east-area.

[6] U.S. Geological Survey, (2022), “The Mineral Industry of Iraq in 2020-2021,” [Online]. Available:https://pubs.usgs.gov/myb/vol3/2020-21/myb3-2020-21-iraq.pdf.

[7] Iraqi Ministry of Planning, “Central Organization for Standardization and Quality Control,” [Online]. Available:https://mop.gov.iq/en/central-organization-for-standardization-and-quality-control.

[8] DieselNet, “Combustion in Diesel Engines,” Revision 2020.05. [Online]. Available:https://dieselnet.com/tech/diesel_combustion.php.

[9] A. T. Doppalapudi, A. K. Azad, and M. M. K. Khan, “Analysis of improved in-cylinder combustion characteristics with chamber modifications of the diesel engine,” *Energies*, vol. 16, no. 6, p. 2586, 2023. [Online]. Available:https://www.mdpi.com/1996-1073/16/6/2586.

[10] G. Chen, R. S. Laramee, and E. Zhang, (2007), “Advanced Visualization of Engine Simulation Data Using Texture Synthesis and Topological Analysis,” NAFEMS World Congress, [Online]. Available:https://www.sci.utah.edu/~chengu/NAFEM2007/adv_vis_top_nafems.pdf.

[11] S. H. Park, H. K. Suh, and C. S. Lee, (2010), “Simulation of quasi-dimensional combustion model for predicting the performance and emission characteristics of a diesel engine,” *Applied Energy*, vol. 87, no. 4, pp. 1316-1325, 2010.

[12] M. A. Hossain, (1996), “Numerical Analysis of Diesel Engine Combustion and Performance Characteristics Using an n-Heptane Surrogate,” *Energies*, vol. 18, no. 5, p. 1082, 2025. [Online]. Available:https://www.mdpi.com/1996-1073/18/5/1082.

[13] Jianbo Zhou, Jianbo Zhou, Jianbo Zhou, Jianbo Zhou, (2023), "Computational Analysis on Combustion Control of Diesel Engines Fueled with Ester Alcohol Diesel," Energies, vol. 16, no. 16, p. 6093, [Online]. Available: https://www.mdpi.com/1996-1073/16/16/6093.

[14] A. K. Singh, (2018), "Mathematical Modelling & MATLAB Simulation of Diesel Engine," SSRN, [Online]. Available:https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3101296.

[15] Syntek Global Inc., "ASTM D975 Diesel Fuel Specification Test," [Online]. Available:https://www.lospaziodibeppe.it/public/download/varie/syntek/ASTM_D975_Specification_Test.pdf.

[16] W. J. K. Al-Obaidi and W. Jamshed, (2024), "A numerical study of the effects of various diesel fuel types on the performance of single-cylinder diesel engines," *Al-Qadisiyah Journal for Engineering Sciences*, vol. 17, pp. 436-444, [Online]. Available:https://iasj.rdd.edu.iq/journals/uploads/2025/01/12/5ae917514470a63879270e12139da3ce.pdf.

[17] Northern Technical University, (2023), "Master's Dissertation: On Evaluating Performance of Diesel Engines," [Online]. Available:https://www.mohesr.gov.iq/en/post/northern-technical-university%26rsquo%3B-master%26rsquo%3Bs-dissertation-on-evaluating-performance-of-diesel-engines-under-influence-of-design-factors-via-using-practical-methods-amp-simulation-2023-11-12-18.

[18] M. H. Abdzaid, (2021), "Biodiesel for Internal Combustion Engine," M.Tech. Thesis, Al-Furat Al-Awsat Technical University, Najaf, Iraq, [Online]. Available:https://cnj.atu.edu.iq/wp-content/uploads/2021/12/MOHAMMED-HUSSEIN-ABDZAID.pdf.

[19] Iraq Business News, (2018), "Technical Specification of Diesel Fuel,". [Online]. Available:https://www.iraq-businessnews.com/wp-content/uploads/2018/09/91121532_Annex-1-Technical-Specification-of-Diesel-Fuel.pdf.

[20] ASTM D975-21, (2021), "Standard Specification for Diesel Fuel," ASTM International, West Conshohocken, PA, [Online]. Available:https://www.astm.org/d0975-21.html.

[21] Dewesoft, "ISO 8528 certification of Diesel Generators," [Online]. Available:https://dewesoft.com/blog/certification-of-diesel-generators.

[22] Naugra Export, "Multi Cylinder Four Stroke Diesel Engine Test Rig," [Online]. Available:https://www.naugraexport.com/multi-cylinder-four-stroke-diesel-engine-test-rig.

[23] A. Boretti, (2023), "Development of a Thermodynamic Model for Supercharged Diesel Engines, "Combustion Science and Technology [Online]. Available:https://www.tandfonline.com/doi/full/10.1080/00102202.2023.2285407.

[24] EPRI, (1996), "Diesel Engine Analysis Guide," TR-107135, [Online]. Available:https://restservice.epri.com/publicdownload/TR-107135/0/Product.

[25] MathWorks, (2023), "Process of Building AI Models for Predicting Engine Performance and Emissions," [Online]. Available:https://www.mathworks.com/videos/process-of-building-ai-models-for-predicting-engine-performance-and-emissions-1683545438770.html.

[26] MathWorks, "Generic Engine Power Curve," [Online]. Available:https://www.mathworks.com/help/sdl/ref/generic_engine_power_curve.png

[27] MathWorks, "Generic Engine Fuel Consumption Map," [Online]. Available:https://www.mathworks.com/help/sdl/ref/generic_engine_01.png.

[28] Wärtsilä, (2023), "WÄRTSILÄ 50DF Multi-fuel engine generating set," [Online]. Available:https://www.wartsila.com/docs/default-source/energy-docs/technology-products/product-leaflets/wartsila-50df.pdf?sfvrsn=e9eda345_21.

[29] N. Haga, "White paper: Combustion engine power plants," Wärtsilä, [Online]. Available:https://www.wartsila.com/docs/default-source/power-plants-documents/downloads/white-papers/general/wartsila-bwp-combustion-engine-power-plants.pdf.

[30] Wärtsilä, "Wärtsilä 50DF dual fuel engine for power plants," [Online]. Available:https://www.wartsila.com/energy/solutions/engine-power-plant-products/wartsila-50df-dual-fuel-engine.

[31] Hayder M. Mnati, Maroua Hammami, Olfa Ksentini, Mohamed S. Abbes, Mohamed Haddar, (2025), "Kinematic and vibration analysis of a single-cylinder engine," Int. Rev. Appl. Sci. Eng.

[32] M. N. Kareema and A. M. Salih, (2020), "Engine performance with diesel-biodiesel blends fuel and emission characteristics," *Engineering and Technology Journal*, vol. 38, no. A05, pp. 779-788.

[33] Solomon Power, (2023), "Technical Specifications for Diesel Generators 2500 KVA," [Online]. Available:https://solomonpower.com.sb/wp-content/uploads/2023/06/TOR-Attachment-2-Generator-Technical-Specifications_Final.pdf.

Experimental Analysis of Thermal - Hydraulic Enhancement in a Shell-and-U-Tube Heat Exchanger Using Copper Foam Baffles / Original Article

Researcher Rawaa Hussein Al-meammar

Mechanical Engineering Department,
Faculty of Engineering, Qom University, Iran
e-mail: rawaa.almeammar@yahoo.com

التحليل التجريبي لتعزيز الأداء الحراري والهيدروليكي في مبادل حراري من نوع الغلاف وأنابيب على شكل حرف U باستخدام حواجز من رغوة النحاس

الباحث رواء حسين المعمار

قسم الهندسة الميكانيكية، كلية الهندسة، جامعة قم، ايران

Abstract

Porous metal foams have emerged as a promising medium for thermal management due to their high surface area, low density, and excellent thermal conductivity. This study investigates the effect of copper porous foam baffles with different pore densities (5 and 10 PPI) on the thermal and hydraulic performance of a shell-and-U-tube water–water heat exchanger. An experimental setup was developed to evaluate the Nusselt number, heat transfer coefficient, pressure drop, and friction factor over a range of Reynolds numbers across the annular side. The results revealed that incorporating porous metal foam significantly enhanced heat transfer compared with the smooth-tube configuration. At Re ≈ 7600, the 5 PPI and 10 PPI foams increased the Nusselt number by approximately 21% and 50%, respectively. However, these enhancements were accompanied by notable increases in pressure drop and friction factor due to the reduced permeability and increased flow resistance of the porous structure. The 10 PPI foam provided the highest heat transfer improvement but at the expense of greater hydraulic losses. Overall, the 5 PPI configuration achieved a balanced trade-off between thermal efficiency and pumping power, confirming that optimizing pore density is essential for practical heat-exchanger design using porous media.

Keywords: Porous metal foam, Copper baffles, Heat transfer enhancement, Thermal–hydraulic performance, Double-pipe heat exchanger, Pressure drop.

المستخلص

برزت الرغوات المعدنية المسامية كوسط واعد في التطبيقات الحرارية نظرًا لمساحتها السطحية العالية، وكثافتها المنخفضة، وموصليتها الحرارية الممتازة. تبحث هذه الدراسة في تأثير حواجز مصنوعة من رغوة نحاس مسامية ذات كثافات مسام مختلفة (5 و10 مسام / البوصة PPI) على الأداء الحراري والهيدروليكي لمبادل حراري ماء–ماء من نوع الغلاف وأنابيب على شكل حرف U.تم تطوير منظومة تجريبية لقياس عدد Nusselt، ومعامل انتقال الحرارة، وفقدان الضغط، ومعامل الاحتكاك ضمن مدى من أعداد Reynolds على جانب الحيز الحلقي (الفراغ بين الغلاف والأنبوب). أظهرت النتائج أن إدماج الرغوة المعدنية المسامية أدى إلى تعزيز ملحوظ في انتقال الحرارة مقارنةً بحالة الأنبوب الأملس. فعند عدد Reynolds يقارب 7600، زادت قيمة عدد Nusselt بنحو 21% و50% عند استخدام رغوة بكثافة 5 و10 مسام/بوصة على التوالي. إلا أن هذا التحسن في الأداء الحراري ترافق مع زيادة واضحة في فقدان الضغط ومعامل الاحتكاك نتيجة انخفاض النفاذية وارتفاع مقاومة الجريان داخل الحواجزالمسامية. وقد حققت رغوة 10 PPI أعلى تحسن في انتقال الحرارة، ولكن مقابل خسائر هيدروليكية أكبر. بشكل عام، أظهرت رغوة 5 PPI توازناً أفضل بين الكفاءة الحرارية وقدرة الضخ المطلوبة، مما يؤكد أن اختيار الكثافة المثلى للمسام يُعد عاملاً أساسياً في تصميم المبادلات الحرارية المعتمدة على الأوساط المسامية لأغراض تطبيقية عملية.

الكلمات المفتاحية: الرغوة المعدنية المسامية, حواجز نحاسية, تعزيز انتقال الحرارة, الأداء الحراري–الهيدروليكي, مبادل حراري مزدوج الأنبوب, فقدان الضغط.

1. Introduction

Porous metal foam is an innovative material that has gained increasing attention in recent years due to its unique properties, such as low density, high permeability, good thermal conductivity, and a large surface-to-volume ratio. These properties make it an ideal medium for enhancing heat transfer by increasing the heat exchange area and stimulating flow turbulence, resulting in higher Nusselt values and cooling efficiency. Metal foams have been widely used in thermal applications, ranging from electronic refrigeration and solar energy systems to thermal storage and waste heat recovery. These applications respond to the challenges of rising global energy demand and the need for efficient and sustainable solutions[1] [2]. In the field of heat exchangers specifically, studies have shown that incorporating metal foam into tubes or tube bundles can double heat transfer rates compared to conventional exchangers, despite the associated increase in pressure loss. Hybrid designs that combine traditional fins and metal foam have also proven to significantly improve thermal performance while reducing energy consumption, making them a promising alternative to conventional heat exchanger technologies. Therefore, porous metal foam represents a promising direction towards developing more efficient and reliable heat exchangers to meet the requirements of modern industries in reducing energy consumption and carbon emissions[3] .Several studies consistently highlight that porous metal foams enhance heat transfer but at the cost of increased pressure loss, and that careful optimization of foam distribution and pore density is essential to achieve balanced performance. For example, Pavel and Mohamad[4] showed that porosity, pore size, and Reynolds number strongly affect heat transfer rates and pressure drops, concluding that porous inserts significantly improve heat transfer with only a moderate increase in pressure loss if properly designed. Similarly, Hamzah, *et al.*,[5] confirmed experimentally that the

best thermal performance was achieved at lower PPI values, where heat transfer was enhanced while pressure losses remained manageable. This trade-off between improved heat transfer and higher pressure drop was also evident in double-tube heat exchangers. Both Marri, *et al.*,[6] and Faal, *et al.*,[7] reported that fully filling the exchanger with foam maximized heat transfer but led to substantial pressure penalties, whereas partial filling using foam baffles provided a more favorable balance between thermal efficiency and hydraulic resistance. These findings align with Alhusseny, *et al.*,[8], who proposed a composite design employing foam coatings and axial rotation to reduce pumping power requirements, further demonstrating that strategic foam distribution can mitigate pressure losses while maintaining strong thermal performance. On the other hand, studies focusing on shell-and-tube and exhaust gas heat exchangers highlighted additional benefits of porous baffles. Abbasi, *et al.*,[9], Chen, *et al.*,[10], and Tian, *et al.*,[11] all showed that incorporating porous baffles significantly improved thermal-hydraulic performance by reducing stagnant flow zones and enhancing mixing. Optimized designs were able to increase heat transfer rates by up to 523.81 kW9]], boost the area goodness factor by more than 500% [10], and simultaneously improve heat transfer by 92% while reducing pressure drop by 65% [11]. However, these works were mostly numerical and did not fully account for long-term operational issues such as fouling.

2. Experimental Setup

A complete water–water heat exchanger rig was designed and constructed, as shown in Figure 1. The system mainly consists of three sections: the test section (Heat exchanger), the hot-water circuit, and the cold-water circuit. In the test section, two copper tubes with U-shaped were fixed inside a transparent Perspex shell and they were concentered together with two thermal Teflon flanges. Hot water flowed through the copper tubes, while cold water was circulated through the annular shell side. Both hot and cold water loops were operated using two independent pumps to maintain the desired flow rates. The hot water was heated in a tank by an electric heater and then circulated through the copper tubes, while cold water was supplied to the shell side. To measure the temperature distribution, six thermocouples were employed: four to monitor the inlet and outlet temperatures of both fluids, and two others placed on the surfaces of the copper tubes. To enhance the thermal performance of the heat exchanger, circular copper porous metal inserts were used as baffles inside the shell side. The porous medium was tested at two pore densities: 5 PPI and 10 PPI. A differential mercury manometer was connected across the shell side to measure the pressure drop. Flow rates of hot and cold water were controlled using two flow meters and valves. Table 1 illustrate the rig specifications.

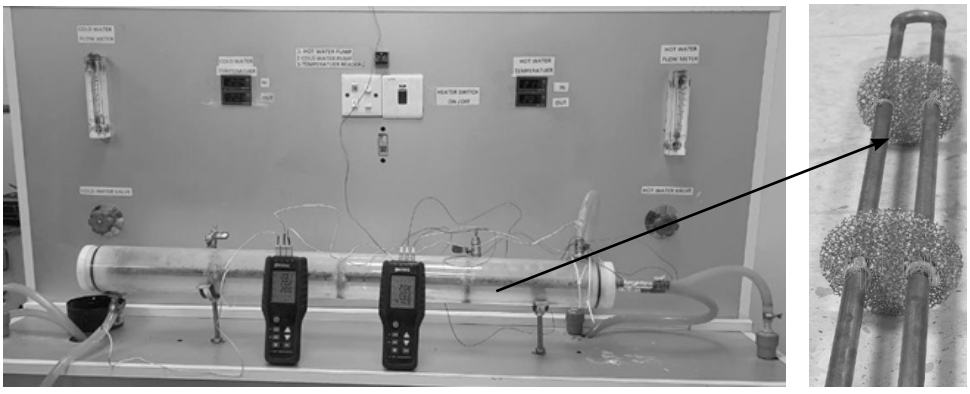


Figure 1: The test rig

Table 1: The specifications of the test rig's parts

Test section **(1) Length** **(2) Inner diameter of copper pipe, di** **(3) Outer diameter of copper pipe, do** **(4) Inner diameter of annular pipe, Di**	 **1 m** **1.07 cm** **1.27 cm** **7.4 cm**
Porous metal foam **(1) Materials** **(2) Thickness** **(3) Shape** **(4) Diameters** **(5) Number** **(6) Pore density**	 **Copper** **1 mm** **circular** **7.4 cm** **3** **5 PPI, 10 PPI**
Water Pumps **(1) Type** **(2) Max. flow** **(3) Power**	 **End suction pump** **35 L/min** **370 W**
Temperature recorder **-Channel** **-Sensor type** **-Operation temperature**	**4 channels** **type (K) thermocouple** **-20 to 150 ºC**
Flow meters **-water flow meter, Range**	 **0 to 18 L/min**

3. Testing Procedure

A shell and U tube heat exchanger was employed where hot water was heated to 60 °C and circulated through the inner copper U tube, while cold water was supplied in counter flow through the annulus. The hot and cold water volumetric flow rates were precisely adjusted using flow meters, and all measurements were

taken after achieving steady-state conditions. For each test, the inlet and outlet temperatures of both streams were recorded to evaluate the heat transfer coefficient. To ensure reliability, the experiments were repeated for five different cold-water flow rates (9 to 13 LPM), resulting in multiple runs for comparative analysis but the hot water flowrate was adjusted to 13 LPM. Three configurations of the inner tube were investigated: (i) a smooth pipe without any modification, (ii) the pipe fitted with porous copper foam of 5 PPI, and (iii) the pipe fitted with porous copper foam of 10 PPI. This experimental methodology allowed for a systematic comparison of thermal performance under different operating and structural conditions.

4. Mathematical Formulation

The heat transfer for hot and cold water sides are calculated using equations the following equations [12][13]:

$$Q_h = \dot{m}_h Cp_h \left(T_{hi} - T_{ho}\right) \quad \textbf{(1)}$$

$$Q_c = \dot{m}_c Cp_h \left(T_{ci} - T_{co}\right) \quad \textbf{(2)}$$

$$Q_{avg.} = \frac{Q_h + Q_c}{2} \quad \textbf{(3)}$$

The convective heat transfer coefficient of the inner pipe can be estimated from equation 4:

$$h_c = \frac{Q_c}{A_s \left(T_{s.inner\ pipe} - T_{an}\right)} \quad (4)$$

Where:

$$A_s = 2\left(\pi\, d_o\, L\right) \quad \textbf{(5)}$$

And the Nu number of the annular side :

$$Nu_c = \frac{h_c D_H}{K} \quad \textbf{(6)}$$

Where D_H is the hydraulic diameter and calculated as follows:

$$D_H = \frac{4 A_c}{P_{wet}} \quad \textbf{(7)}$$

$$A_c = A_{c.an} - 2 A_{c.inner} = \frac{\pi}{4} D_i^2 - 2 \times \frac{\pi}{4} d_o^2 \quad \textbf{(8)}$$

$$P_{wet} = \pi d_{i.an} + 2 \pi d_{o.inner} \quad \textbf{(9)}$$

The Re number for both sides :

$$Re_h = \frac{(\rho v D_H)_h}{\mu_h} \quad \textbf{(10)}$$

$$Re_c = \frac{(\rho v d_i)_c}{\mu_c} \quad \textbf{(11)}$$

Equation (12) is used to calculate the annular friction factor.

$$f_c = \frac{\Delta P \left(\frac{D_H}{L} \right)}{\frac{\rho v^2}{2}} \quad \textbf{(12)}$$

And to calculate the theoretical Nu number, Dittus-Boelter correlation is used:

$$Nu_c = 0.023 (Re)^{0.8} (\mathrm{Pr})^{0.33} \quad \textbf{(13)}$$

$$Pr = \frac{Cp. \mu}{K} \quad \textbf{(14)}$$

5. Uncertainty Analysis

The accuracy of the experimental results depends on the calibration of the measuring devices and the accuracy of the readings. Calibrated devices such as thermocouples and flow meters were used, and the experiments were repeated twice to ensure reliability. The uncertainty was ±0.5 °C for the thermocouple and ±0.05% for the flow meter. Uncertainty analysis for the calculated quantities, Reynolds and Nusselt numbers, thermal load, and annular friction factor was carried out according to Equations (13) to (17) [14]. The obtained values are presented in Table (2). Since the uncertainties in thermo-physical properties like, specific heat, viscosity, density, and thermal conductivity were insignificant, they were neglected. The manufacturer indicates that the possible deviation in the dimensions (diameters and length) does not exceed ±0.5 mm.

$$\frac{Un_{Re_{an}}}{Re_{an}} = \pm\sqrt{\left(\frac{Un_{\dot{m}}}{\dot{m}}\right)^2_{an} + \left(\frac{Un_{d_{an}}}{d_{an}}\right)^2 + \left(\frac{Un_{\mu_{an}}}{\mu_{an}}\right)^2} \quad \textbf{(15)}$$

$$\frac{Un_{Q_{an}}}{Q_{an}} = \pm\sqrt{\left(\frac{Un_{\dot{m}}}{\dot{m}}\right)^2_{an} + \left(\frac{Un_{T_i}}{T_i - T_o}\right)^2 + \left(\frac{Un_{T_o}}{T_i - T_o}\right)^2} \quad \textbf{(16)}$$

$$\frac{Un_{h_{an}}}{h_{an}} = \pm\sqrt{\left(\frac{UQ_{\dot{m}}}{Q}\right)^2_{an} + \left(\frac{Un_{d_{an}}}{d_{an}}\right)^2 + \left(\frac{Un_{L_{an}}}{L_{an}}\right)^2 \left(\frac{Un_{T_i}}{T_i - T_o}\right)^2 + \left(\frac{Un_{T_o}}{T_i - T_o}\right)^2} \quad \textbf{(17)}$$

$$\frac{Un_{Nu_{an}}}{Nu_{an}} = \pm\sqrt{\left(\frac{Un_h}{h}\right)^2_{an} + \left(\frac{Un_{d_{an}}}{d_{an}}\right)^2 + \left(\frac{Un_{k_{an}}}{k_{an}}\right)^2} \quad \textbf{(18)}$$

$$\frac{Un_{f_{an}}}{f_{an}} = \pm\sqrt{\left(\frac{Un_{\Delta P}}{\Delta P}\right)^2_{an} + \left(\frac{Un_{d_{an}}}{d_{an}}\right)^2 + \left(\frac{Un_{L_{an}}}{L_{an}}\right)^2 + \left(\frac{Un_{\rho_{an}}}{\rho_{an}}\right)^2 + \left(\frac{Un_{v_{an}}}{v_{an}}\right)^2} \quad \textbf{(19)}$$

Table (2): Uncertainty for experiments parameters

Parameters	Uncertainty %		
	Plain Pipe	Porous metal foam (5 PPI)	Porous metal foam (10 PPI)
Reynold's Number	2.94	2.81	1.005
Heat Transfer	4.46	2.96	1.66
Heat Transfer Coefficient	2.58	3.66	5.27
Nusslte's Number	0.9652	0.9652	0.9666
Friction Factor	3.25	3.08	1.24

6. Results and Discussion

A water-to-water in a shell and U- tube heat exchanger model was employed to investigate the thermal enhancement using copper porous foam as baffles in annular side.15 experiments were done in order to illustrate the porous metal foam and its porosity on the thermal and hydraulic performance of the heat exchanger. The results obtained from these experiments are summarized in the following sections.

6.1 The Model Validation

A comparison was conducted between the Nusselt number obtained for the plain pipe and the value calculated from the established Dittus-Boelter correlation [15]. Figure 2 presents this comparison, contrasting the experimentally determined Nusselt numbers across a range of Reynolds numbers with the corresponding theoretical predictions. The results demonstrating good agreement with an error margin equal 8.32%.

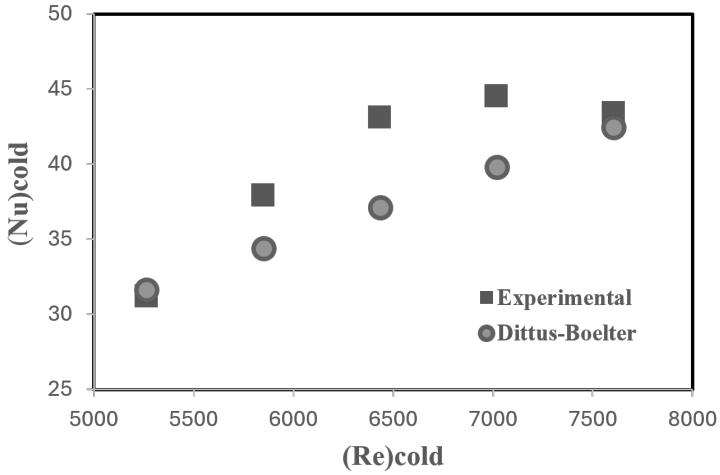


Figure 2: Comparison between experimental and correlated Nusselt number values for the smooth pipe

6.2 Thermal and hydraulic Performance

The thermal and hydraulic performance of the shell and U-tube heat exchanger was evaluated for the three configurations. The results are presented as a function of the Reynolds number (Re) to illustrate the effect of pore density on both heat transfer and flow resistance. The thermal performance was assessed using the Nusselt number (Nu), heat transfer coefficient (h), and heat transfer rate (Q) across the annular side.

As shown in Figure 3, Nu number increases with Reynolds number for all cases, indicating enhanced convective heat transfer at higher flow rates. At Re ≈ 7600, the smooth tube achieved a Nu of about 40, while the 5 PPI and 10 PPI configurations reached approximately 62 and 83, respectively, representing average enhancements of 21% and 50% over the smooth pipe. This improvement is attributed to the increased turbulence intensity and larger specific surface area created by the open-cell copper foam, which promotes stronger fluid mixing and

disrupts the thermal boundary layer. Figure 4 shows the average heat transfer coefficient followed a similar trend because the smaller pores in the 10 PPI foam provide greater solid–fluid contact area and enhance conductive heat spreading through the copper matrix, thus improving the overall convective performance.

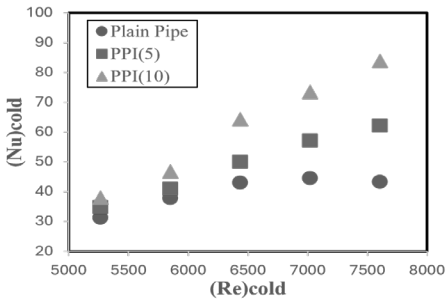


Figure. 3: Nu number vs. Re number of three different configuration

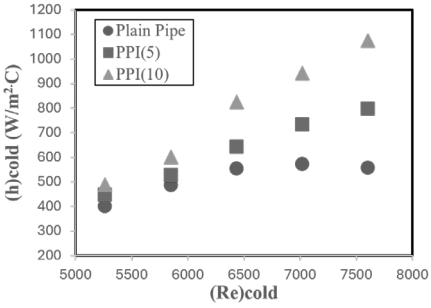


Figure. 4: Heat transfer coefficient vs. Re number of three different configuration

On the other hand, the hydraulic behavior was evaluated through pressure drop (ΔP) and friction factor (f) to quantify the effect of the porous structure on flow resistance.

Based on the results represented in the Figures 5 and 6 for pressure drop and friction factor across varying Reynolds numbers, the results clearly demonstrate the significant hydraulic impact of incorporating porous metal foam baffles in the annular side of the shell and U-tube heat exchanger. Compared to the plain pipe baseline, both the PPI(5) and PPI(10) configurations exhibit a substantial increase in both the friction factor and the measured pressure drop for any given Reynolds number. This is a direct consequence of the complex, tortuous flow path created by the porous matrix, which intensifies flow disruption and mixing. Furthermore, the finer-pored PPI(10) foam consistently generates a higher flow resistance than the PPI(5) foam across the operational range, highlighting the direct relationship

between pore density and hydraulic penalty. These figures underscore the classic trade-off in heat exchanger enhancement: the dramatic improvement in heat transfer performance offered by the metal foam is achieved at the cost of a greatly increased pressure drop, which necessitates higher pumping power for the system.

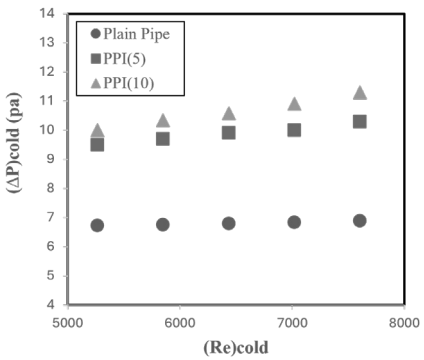


Figure 5: Pressure drop vs. Re number of three different configuration

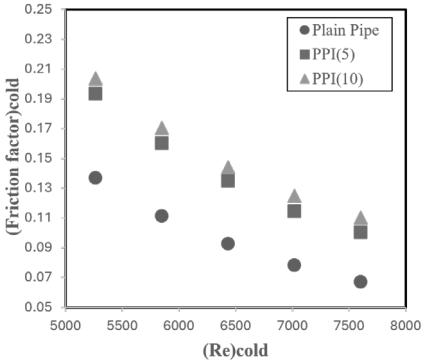


Figure 6: Friction factor vs. Re number of three different configuration

6.3 Thermal Performance Factor

An effective thermal enhancement in a heat exchanger is achieved when the rise in the heat transfer coefficient is greater than the corresponding increase in pressure drop, as represented in Eq. (20) [16]. The relationship between the Thermal Performance Factor (TPF) and Reynolds number is illustrated in Figure 7. The results reveal that TPF exceeds unity for every enhanced configuration, confirming the improvement in thermal performance among the tested designs.

$$TPF = \frac{\frac{Nu_{.porous\ metal}}{Nu_{.plain\ pipe}}}{\left(\frac{f_{.porous\ metal}}{f_{.plain\ pipe}}\right)^{\frac{1}{3}}} \quad \text{.....(20)}$$

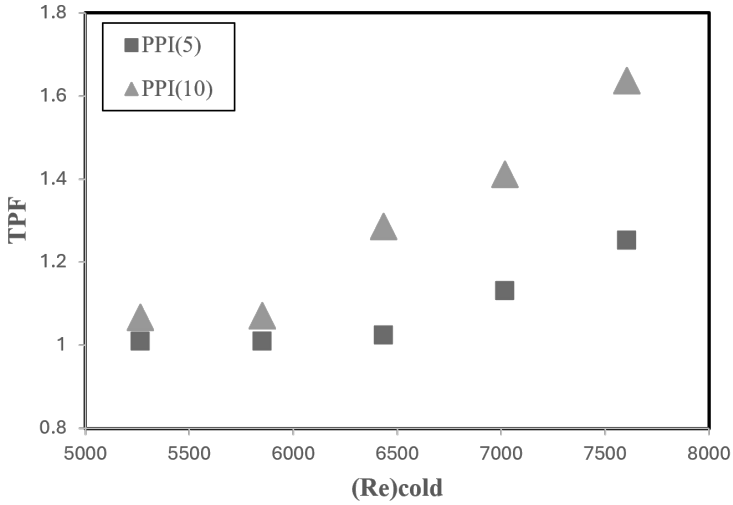


Figure 7: Thermal performance factor vs. Re number of the heat exchanger

7. Conclusion

The experimental investigation demonstrated that the integration of copper porous metal foam baffles markedly enhances the heat-transfer performance of a shell-and-U-tube heat exchanger. The open-cell structure of the foam increases surface area and induces turbulent mixing, leading to higher Nusselt numbers and heat-transfer coefficients compared with a smooth pipe. Specifically, the 5 PPI and 10 PPI foams improved heat transfer by up to 21% and 50%, respectively, confirming the strong influence of pore density on thermal performance. Nevertheless, these enhancements come with increased pressure losses and friction factors, particularly for the 10 PPI configuration, where the smaller pore size intensifies flow obstruction. Hence, while high-density foams offer superior heat-transfer capability, their hydraulic penalty limits overall system efficiency. The results suggest that moderate pore densities (around 5 PPI) provide the most effective balance between improved thermal characteristics and acceptable pressure drops.

NOMENCLATURE

A_s	Surface area, m^2
A_c	Cross sectional area, m^2
Cp	Specific heat, kJ/kg. °C
d_i	Inside diameter of the inner pipe, m
d_o	Outside diameter of the inner pipe, m
D_i	Inside diameter of the annular pipe, m
D_o	Outside diameter of the annular pipe, m
D_H	Hydraulic diameter, m
f	Friction factor
h	heat transfer coefficient, W/m^2. °C
k	Thermal conductivity, W/m. °C
L	Length of the tube, m
m	Mass flow rate, kg/s
Nu	Nusselt number
Pr	Prandtl number
Q	Heat transfer rate, W
Re	Reynolds number
T	Temperature, °C
Un	Uncertainty
ΔP	Pressure drop, pa
v	Velocity, m/s

Greek symbols

μ	Dynamic viscosity, kg/m.s
ρ	Density, kg/m^3

Subscript

h	hot
c	cold
i	Inlet
o	Outlet

8. References

[1] K. Chen, L. Guo, and H. Wang,(2020), "A review on thermal application of metal foam," Sci. China Technol. Sci., vol. 63, no. 12, pp. 2469–2490, , doi: 10.1007/s11431-020-1637-3.

[2] S. T. W. Kuruneru, K. Vafai, E. Sauret, and Y. Gu,(2020), "Application of porous metal foam heat exchangers and the implications of particulate fouling for energy-intensive industries," Chem. Eng. Sci., vol. 228, p. 115968, doi: 10.1016/j.ces.2020.115968.

[3] J. Marx and A. Rabiei, 2017, "Overview of Composite Metal Foams and Their Properties and Performance," Adv. Eng. Mater., vol. 19, no. 11, doi: 10.1002/adem.201600776.

[4] B. I. Pavel and A. A. Mohamad, 2004, "An experimental and numerical study on heat transfer enhancement for gas heat exchangers fitted with porous media," International Journal of Heat and Mass Transfer, vol. 47, no. 23. pp. 4939–4952, doi: 10.1016/j.ijheatmasstransfer.2004.06.014.

[5] J. A. Hamzah and M. A. Nima, 2020, "Experimental Study of Heat Transfer Enhancement in Double-Pipe Heat Exchanger Integrated with Metal Foam Fins," Arabian Journal for Science and Engineering, vol. 45, no. 7. pp. 5153–5167, doi: 10.1007/s13369-020-04371-3.

[6] G. K. Marri and C. Balaji, 2021, "Experimental and numerical investigations on the effect of porosity and PPI gradients of metal foams on the thermal performance of a composite phase change material heat sink," International Journal of Heat and Mass Transfer, vol. 164. doi: 10.1016/j.ijheatmasstransfer.2020.120454.

[7] Z. S. Faal, A. J. Jubear, and H. R. Al-Bugharbee, 2024, "Effect of Metal Foam's Volume on the Performance of a Double Pipe heat exchanger," Wasit J. Eng. Sci., vol. 12, no. 2, pp. 33–48, doi: 10.31185/ejuow.Vol12.Iss2.495.

[8] A. Alhusseny, A. Turan, and A. Nasser, 2017, "Rotating metal foam structures for performance enhancement of double-pipe heat exchangers," International Journal of Heat and Mass Transfer, vol. 105. pp. 124–139, doi: 10.1016/j.ijheatmasstransfer.2016.09.055.

[9] H. R. Abbasi, E. Sharifi Sedeh, H. Pourrahmani, and M. H. Mohammadi, 2020, "Shape optimization of segmental porous baffles for enhanced thermo-hydraulic performance of shell-and-tube heat exchanger," Appl. Therm. Eng., vol. 180, p. 115835, doi: 10.1016/j.applthermaleng.2020.115835.

[10] T. Chen, G. Shu, H. Tian, T. Zhao, H. Zhang, and Z. Zhang, 2020, "Performance evaluation of metal-foam baffle exhaust heat exchanger for waste heat recovery," Applied Energy, vol. 266. doi: 10.1016/j.apenergy.2020.114875.

[11] H. Tian, T. Zhao, L. Shi, T. Chen, X. Ma, H. Zhang,. G. Shu., 2020, "Assessment and optimization of exhaust gas heat exchanger with porous baffles and porous fins," Appl. Therm. Eng., vol. 178, p. 115446.

[12] F. P. Incorpera, D. P. DeWitt, T. L. Bergman, and A. S. Lavine, 2006, "Fundamentals of Heat and Mass Transfer", 6th Edition, John Wiley, New York.

[13] T.L Bergman., A.S. Lavine, F.P. Incropera, D.P. DeWitt, 2011, "Introduction to Heat Transfer", 6th edition, John Wiley & sons, New York.

[14] H. W Coleman., and W. G. Steele, 1999,"Experimentation and Uncertainty Analysis for Engineers". 2nd edition, John Wiley & sons , New York.

[15] Dawid Taler, and Jan Taler, 2017, "Simple heat transfer correlations for turbulent tube flow", E3S Web of Conferences 13, doi: https://doi.org/10.1051/e3sconf/20171302008.

[16] M.A. Hussein, V.M. Hameed, 2022, "Experimental Investigation on the Effect of Semi-circular Perforated Baffles with Semi-circular Fins on Air–Water Double Pipe Heat Exchanger". Arab J Sci Eng 47,p. 6115–6124. https://doi.org/10.1007/s13369-021-05869-0

Enhancement of Gas Turbine Efficiency by Using Recuperator and Intercooler Systems /Original Article

Researcher Mohammed Saad Kamil

Department of Mechanics,
Azarbaijan Shahid Madani University, Iran
e-mail: kamilmohammedsaad@gmail.com

تعزيز كفاءة التوربينات الغازية باستخدام أنظمة الاسترجاع الحراري والتبريد البيني

الباحث محمد سعد كامل

قسم الهندسة الميكانيكية، جامعة آذربيجان شهيد مدني، إيران

Abstract

High power density and operational flexibility make gas turbines an indispensable part of contemporary energy conversion systems. However, in the simple cycle configuration, the thermal efficiency is generally low because the exhaust heat loss is large and the compression work is high. The current investigation is focused on enhancing the thermal efficiency of a gas turbine by the introduction of two augmentation schemes, namely a recuperator and an intercooler.

A numerical and thermal simulation study was carried out to model the gas turbine cycle based on the first and second laws of thermodynamics. Simplicity and Modified cycles Performance of simple and modified cycle are investigated. The results showed that the recuperator raised the thermal efficiency of the overall system by about 6–10%, and the compressor work was alleviated by the intercooling system by about 12%. A maximum improvement of about 15% for the overall efficiency was obtained from the application of both schemes simultaneously when compared to the simple cycle.

The results verify that the parallel use of a recuperator and intercooler is a feasible and attractive alternative for increasing the energy output of a gas turbine. The integration process will be very valuable for application in Iraqi gas power plants to enhance fuel economy and sustainable energy generation.

Keywords: Gas turbine, thermal efficiency, recuperator, intercooler, numerical analysis, thermal simulation

المستلخص

تُعدّ التوربينات الغازية من أهم منظومات تحويل الطاقة في التطبيقات الصناعية ومحطات توليد القدرة الكهربائية، وذلك لما تتميز به من كثافة قدرة عالية ومرونة تشغيلية كبيرة. إلا أن كفاءتها الحرارية في دورة التشغيل البسيطة تبقى محدودة نسبيًا، نتيجة الفواقد الحرارية الكبيرة المصاحبة لغازات العادم، إضافة إلى الارتفاع الكبير في الشغل المستهلك من قبل الضاغط. يهدف هذا البحث إلى تحسين الكفاءة الحرارية للتوربين الغازي من خلال دمج نظامين لتحسين الأداء، هما نظام الاسترجاع الحراري (Recuperator) ونظام التبريد البيني (Intercooler).

أُجريت دراسة تحليلية عددية ومحاكاة حرارية لدورة التوربين الغازي اعتمادًا على القانونين الأول والثاني للديناميكا الحرارية، حيث تم نمذجة وتحليل أداء أربع حالات تشغيلية مختلفة تشمل: الدورة البسيطة، دورة مزودة بمسترجع حراري، دورة مزودة بنظام تبريد بيني، والدورة المركبة التي تجمع بين النظامين معًا. أظهرت نتائج المحاكاة أن استخدام المسترجع الحراري أدى إلى رفع الكفاءة الحرارية للنظام بنسبة تراوحت بين (6–10%) نتيجة الاستفادة من حرارة غازات العادم في تسخين الهواء الداخل إلى غرفة الاحتراق. كما بيّنت النتائج أن نظام التبريد البيني أسهم في تقليل شغل الضاغط بنسبة تقارب (12%)، مما انعكس إيجابًا على القدرة الصافية المنتجة.

وقد تحقق أعلى تحسن في الأداء عند دمج النظامين معًا، حيث بلغت الزيادة الكلية في الكفاءة الحرارية حوالي (15%) مقارنة بالدورة البسيطة، إلى جانب انخفاض ملحوظ في درجة حرارة غازات العادم وتحسن القدرة الصافية للنظام. تؤكد نتائج الدراسة أن الاستخدام المتزامن للمسترجع الحراري والتبريد البيني يُعد خيارًا هندسيًا فعالًا وقابلًا للتطبيق العملي لرفع كفاءة التوربينات الغازية وتحسين استغلال الطاقة. كما تُعد هذه المنظومة مناسبة بشكل خاص للتطبيق في محطات توليد الطاقة الغازية في العراق، لما لها من دور مهم في تحسين اقتصاد الوقود وتقليل الفواقد الحرارية وتعزيز استدامة منظومة الطاقة.

الكلمات االمفتاحية: التوربين الغازي، الكفاءة الحرارية، المسترجع الحراري، التبريد البيني، التحليل العددي، المحاكاة الحرارية

1. Introduction

Gas turbines play a key role in thermal system applications to directly convert chemical energy of fuel into mechanical energy at the rotating shaft, which is then connected to a generator to produce electricity, or to a compressor or to other types of loads (e.g., for aircraft or ship propulsion). They exhibit a very high-power density over a wide range of loads and can be started quickly and maintained easier than steam turbines [1]. With the development of energy generation technologies in recent decades, worldwide attention has progressively been directed at enhancing the thermal efficiency of the gas turbine based on the ever-growing fossil fuel prices and the need to minimize the emission of gases harmful to the environment [2]. Although gas turbines are extensively applied in power plants around the world, their thermal efficiency is still limited when operating in simple cycle mode, which is about 30% –35%. This is mainly caused by the loss of a large amount of thermal energy carried away by exhaust gases at high temperatures over 500°C [3]. This large heat loss has been invited the researchers to proposed auxiliary systems to recover a fraction of the wasted heat to increase the overall efficiency of the thermal cycle without significant end modifications of the basic turbine components.

Among the most now a day s widely used engineering measures to realize this objective is the use of recuperators, which are heat exchangers that recover the heat of the hot exhaust gas to the compressed air going toward the combustion chamber. The air temperature is increased prior to combustion and so less fuel is needed to attain the same operating temperature [4]. It has been experimentally proven that thermal efficiency would increase by about 5 to 10% depending on heat exchanger effectiveness and the temperature of operation [1].

There is also the intercooling mechanism, an additional critical technological strategy that further increases the gas turbine effectiveness. The general rule is that

the system is considered in multi -stage compressors, cooling the air between stages upon leaving the first stage and entering into the second one. Decreasing the work of compression by cooling the air compressing [5,6]. Thermal modeling results have suggested that due to the reduction in mechanical power needed by the compressor, intercooling may increase the overall performance of the plant as well as the economic feasibility [3].

The recuperator and intercooler concept is a high-level engineering design solution to realize high efficiency and low emission gas turbine cycles. The recuperator serves to recover waste heat in the exhaust gases, while the intercooling minimizes the power required by the compressor, leading to an increase in the net output of the cycle as a whole [2]. Thermal design investigations show the coupling of the two processes can raise the overall cycle efficiency by up to 15% compared to the simple cycle without major influences on the fabrication or maintenance cost [4].

The practical implication of the current study is derived from the fact that Iraq is at present trying to enhance and improve the working efficiency of its power generation system in order to keep up with the soaring demand of electricity with minimum fuel consumption loss. Most Afghan gas plants have already been fitted with simple cycle; thus, the thermal losses are high and the output power is low. Consequently, the introduction of recuperator and intercooling systems constitutes a step in the right direction for enhancing the energy conversion efficiency and the sustainability of the power sector [7]. Hence, this study is focused on the analysis of the integrated effect of recuperator and intercooling configurations on gas turbine efficiency through the develop- ment of a mathematical model based on the first law of thermodynamics along with thermal simulations and numerical calculations by utilizing com- mercial engineering software. The goal is to predict the increase in efficiency and network output due to the application of such systems for a wide range of operating conditions.

The results of this study are anticipated to contribute to the design of a new generation gas turbine in Iraq and give realistic advice to the energy sector authorities to how and if such systems can be implemented in national power stations.

1.1 Problem of the research

As result of high operating temperature and speed of response to energy load, gas turbine has become one of the key building blocks in power generation systems. However, an important limitation to its application is the fact that its thermal efficiency is much lower than that of other energy conversion systems, such as steam turbines or combined-cycle systems [8]. Decrease in efficiency is ascribable to the loss of considerable heat with exhaust gases and to the high work consumed by the air compressor which causes a decrease in the net power output and increase in the rate of fuel consumption [9].

For a long time, from the engineering point of view, the problem has been whether the thermal and dynamic design of the gas turbine could be improved by introducing auxiliary systems, such as the recuperator and the intercooler, which enable to regain a portion of the lost heat and also to lessen the compression work [6]. However, majority of these works address general theoretical and/or experimental models based on ideal working conditions and no elaborate numerical analysis or thermal simulation which are specifically designed for the actual operating conditions of Iraqi gas power plants, subjected to severe environmental conditions with ambient temperature surpassing 45•c in summer season have been reported [10].

The problem of this research is clearly the poor performance of gas turbines in Iraq because they work on simple cycles without thermal efficiencies or working systems. This imperfection causes large energy losses and degradation in the economic and environmental performances of the power generation station. Hence

the fundamental question of research is: What are the limits of predictive power of theory if it is based on physical reality?

To what extent can systems of recuperator and intercooling improve, in terms of usability, thermal efficiency and plant performance of gas turbines by a real numerical model and thermal simulation?

1.2 From this main question stem a number of other questions are born

1. What is the gain in thermal efficiency for the combined cycle including a recuperator?
2. Effect of Intercooling on Net Work and Power Output.
3. How well the results of the thermal simulation and numerical calculation match with the theoretical (or experimental) results in the literatures of engineering?
4. These questions are of scientific and practical importance in the context of the national priorities in Iraq to improve power generation efficiency and diminish fossil fuel consumption. The Iraqi power sector is characterized by poor overall efficiency as a result of the use of simple gas turbine cycles without adequate

Technological development [11]. The results aim to support energy sustainability and enhance the economic feasibility of gas turbine power plants in Iraq.

1.3 Research significance

The importance of this research has been drawn from the fact that it deals with a significant problem within the electricity sector in Iraq – limited efficiency in gas power plants that run in simple cycles without undergoing technological enhancement or modernization. The real efficiency of those plants is estimated not

to exceed an average of 30 percent, as opposed to the world efficiency rates of 40-45 percent in combined cycle power plants [11]. This reduction in efficiency is one of the largest generators of energy waste and high operational costs in the national electric system.

From a scientific and engineering matter, the novelty of this work concerns the improvement of the thermal efficiency of gas turbines via the application of two advanced methods, namely the recuperator and the intercooler (designed to reduce the compression work in gas turbines). These subsystems are intended to provide increasing net output power and enhancing thermal efficiency of the cycle in sub-global scales. Recently, it was demonstrated that the synergy implementation of both techniques can result in a combined performance and efficiency increase as high as 15% [5].

The author is saying that from an economic point of view and practical point of view, introducing such technologies for the Iraqi power generation plants might have positive effect in terms of nusing the fuel and reducing gas emissions which come along with the government policies to follow achieving the energy sustainability and following the international environmental standards [10]. In addition, the results of the suggested numerical thermo-simulation analysis might serve as on-the-job training for engineers and technicians in the energy industry to improve efficiencies of hot climate applications. Thus, the current study is not only theoretical, but a practical model that can be used in the design, analysis, and operation levels. Therefore, its theoretical and practical values render it a subject worthy of study in the field of mechanical energy engineering and gas turbine thermal performance analysis.

2. Materials and Methods

2.1 Study design

This study is made on the basis of Applied Analytical Research, which integrates theoretical and practical aspects:

Theoretical Study: A critical survey of modern scientific studies on enhancement of gas turbine performance, in terms of the effect of recuperators and intercooling arrangements on thermal efficiency and net power generation [5,6]. Modeling and non-dimensional analysis: Modeling of the gas turbine cycle to analyze performance for various working conditions namely simple cycle, cycle with recuperator, cycle with intercooling and combined cycle. This allows for the fusion of theory and practice and offers a quantitative data set that is valid for academic scrutiny and science publication.

2.2 Materials and tools

2.2.1 Software and numerical modeling

MATLAB/Simulink is employed to develop the gas turbine cycle model which is used to predict the thermal efficiency and the network output based on the thermodynamic relations of the Brayton cycle.

CFD software (optional): To be used if detail simulation of heat and pressure distributions within the recuperator and intercooler is needed.

ESS is also used for energy and heat analysis flow analysis in the cycle in terms of the accurate thermal calculations.

2.2.2 References and engineering standards

The air and fuel property data were based on ASHRAE (2017) international regulations.Gas turbine and compressor parameters from earlier work (Boyce, 2017). Recuperator and intercooling parameters as given by Cohen *et al.* (2019).

2.2.3 Studied variables

Compressor pressure ratio (CPR). Turbine inlet temperature (TIT).

Recuperator effectiveness (ε).

After the first compression in intercooling air temperature.

The net power output, the thermal efficiency, the exhaust gas temperature and the work of compressor.

2.3 Procedures and work methods

2.3.1 Basic Gas Turbine Cycle Design

The simple cycle is the baseline case. Thermal efficiency, net power output, work by compressor and temperature of exhaust gas are calculated by using the typical Brayton cycle equations [6].

2.3.2 Recuperator integration

The recuperator is added to the cycle to study the exhaust gas heat recovery and the combustion air preheating. The effect of recuperator efficiency (0.7–0.85) was modelled in order to evaluate the sensitivity of thermal efficiency, net power output, and exhaust temperature.

2.3.3 Intercooler integration

The impact of the air cooling during the spacing of the compression is analyzed to see its effect on the net power output and the compressor work while that the temperature of the inlet turbine is the unaltered. The specific power of the output of the cycle is studied against various working conditions.

2.3.4 Combined cycle (Recuperator + Intercooler)

The two subsystems are combined to obtain full thermal synergy. Overall performance is predicted, i.e., thermal efficiency, net power output, exhaust temperatures and compressor work.

Output Analysis to Compare the Four Layouts Performance

2.3.5 Model validation

To test the model accuracy and reliability, the simulated results are validated against previous published results [5,8].

2.3.6 Graphical and statistical analysis

The results are tabulated and graphically (bar and line charts) represented in a way to give a comparative view of case performances.

An evaluation of relative merits is carried out to select the best arrangement from the point of view of efficiency, net power and compressor work.

2.4 Theoretical and practical contributions of the method

The theoretical investigation combined with numerical simulation guarantees the high precision and dependability of the results. The approach can be applied to the design of current industrial gas turbines, with enhanced thermal and economic performances.

In addition, the model can be used to investigate other performance enhancement techniques such as multistage intercooling and mechanical energy recovery schemes.

3. Result and Discussion

3.1 Result of the simple Gas turbine cycle (simple cycle)

The cycle efficiency is calculated based on the Brayton cycle calculations, which used the thermophysical properties at very high temperatures and included a regenerative configuration of the cycle. The simulation of the simple gas turbine cycle indicates that the thermal efficiency of the turbine attains around 30.2% at a pressure ratio of 10 1 with a turbine inlet temperature of 1100 K. The power output at 23.8 MW was obtained at the air mass flow rate of 100 kg/s. These trends show that the compressor uses around 60% of the power produced which subsequently decreases the overall cycle efficiency and accentuates the need for better energy utilization within the system [6].

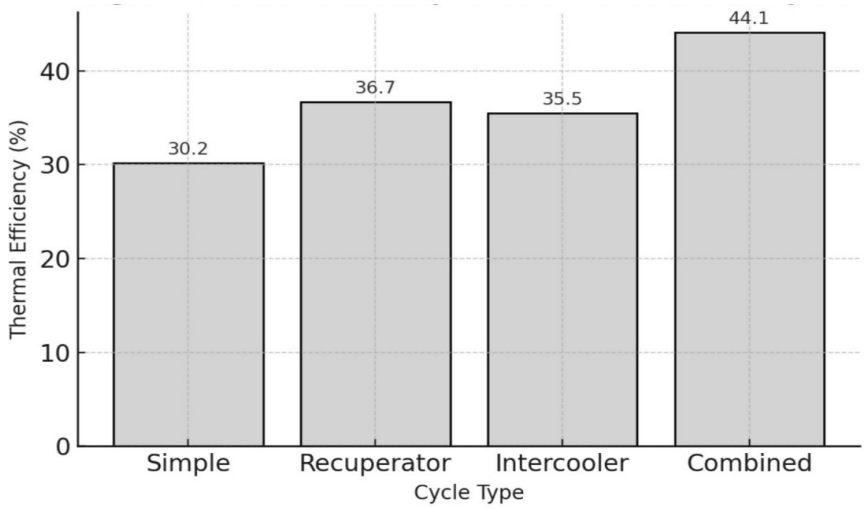


Figure 1. Thermal efficiency of different Gas turbine cycle

Also, the effect of increasing the pressure ratio on the thermal efficiency is to increase it to a maximum and then decreases due to the increasing heat loss and higher compressor work. This indicates that there is a best value for the pressure

ratio which gives the highest efficiency before deteriorating performance. These are in agreement with classical results for Brayton cycle which predicted nonlinear relation between pressure ratio and thermal efficiency [8].

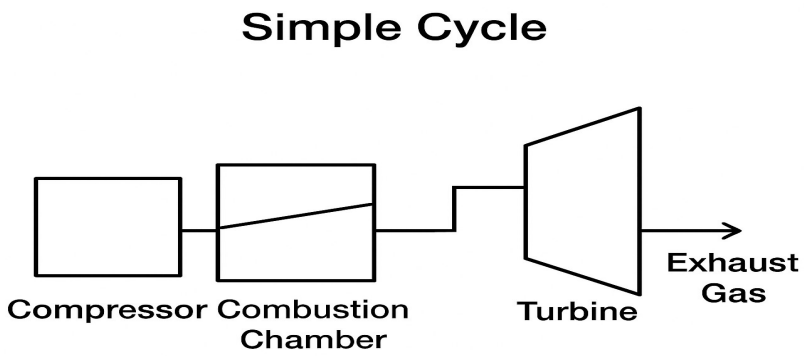


Figure 2. Effect of pressure ratio on thermal efficency

3.2 Result of the recuperator cycle

The results of simulations of the cycle with recuperator were demonstrated to have substantially better thermal performance when compared to the plain two-shaft gas turbine cycle. The thermal efficiency of the turbine raised to 36.7% at a recuperator effectiveness of ε = 0.8. This enhancement is attributed to the reuse of part of the heat in exhaust gases to pre-heat the air flow into the combustion chamber, thereby decreasing the fuel needed to reach the identical final temperature [5]. The results also showed that the temperature of the exhaust gas dropped by 120–150 K compared with the basic cycle, which indicates that more waste heat was recovered. Indeed, the introduction of the recuperator resulted in a small increase in system pressure drop, but the thermal benefits far outweighed the detriments, thus allowing the recuperator to be considered a viable means of increasing overall cycle efficiency [8].

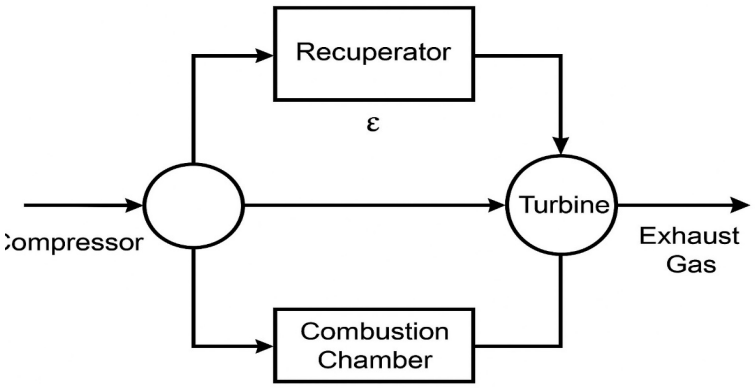


Figure 3. Effect of recuperator effectiveness on thermal efficiency

3.3 Result of intercooler cycle

Simulation results indicate that the goods including an intercooler, the compressor's work is reduced by about 11.8% due to the temperature of the air after the first compression stage was lowered to (330k) from (580K) and that means before the second stage. This improved the net power output of the turbine by around 8-10% relative to the simple cycle, so that the performance of the whole system was enhanced [6].

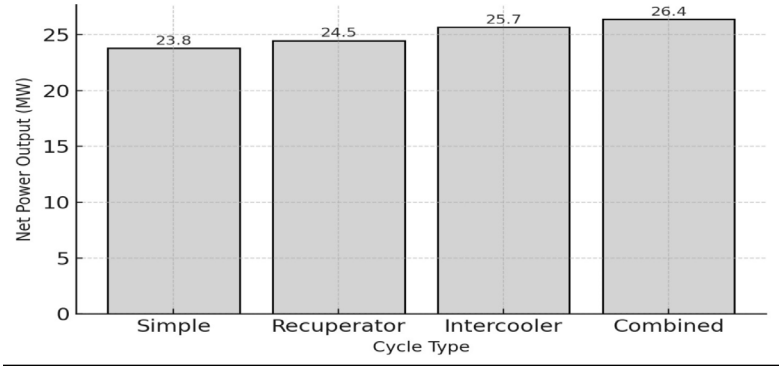


Figure 4. Net power output of different gas turbine cycle.

While the intercooling system by itself did not result in any significant increase in thermal efficiency (keeping the value between 32–33%), it positively impacted the system power output per unit mass. Thus, intercooling is especially useful for high power density requirement applications since it improves energy utilization without significantly raising the overall thermal efficiency [5].

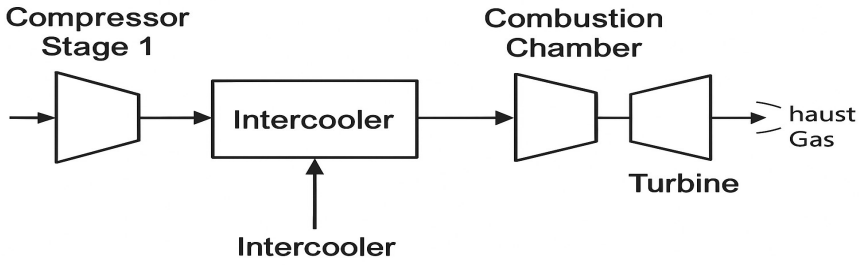


Figure 5. Compressor work reduction with intercooling

3.4 Result for the combined cycle (Recuperator + Intercooler)

For the combined cycle (both the recuperator and intercooler), the simulation results based on the system thermal performance among all cases were the highest. The thermal efficiency increased up to 44.1% at a pressure ratio of 10:1 and the recuperator effectiveness of $\varepsilon = 0.8$, but the net power output was increased to 26.4 MW.

The substantial enhancement is due to a thermal coupling between the two processes: the compressor work is reduced by the intercooler, and a large fraction of the exhaust gas heat is recovered by the recuperator, leading to an improvement of the overall system efficiency. The simulation results also exhibited a linear increase in thermal efficiency with recuperator effectiveness up to about $\varepsilon \approx 0.85$ where the rate of improvement declines due to more pressure loss inside vessel [8].

In addition, the exhaust gas temperature dropped from 650 K in the simple cycle to approximately 480 K in the combined cycle, indicating more efficient use of the input thermal energy. This is in line with results from Applied Thermal Engineering which showed that the use of a recuperator combined with intercooling was the thermal and economic superior solution compared to other thermal recuperative options [5].

The following table illustrates the clear improvement in performance resulting from the implementation of cycle enhancement systems.

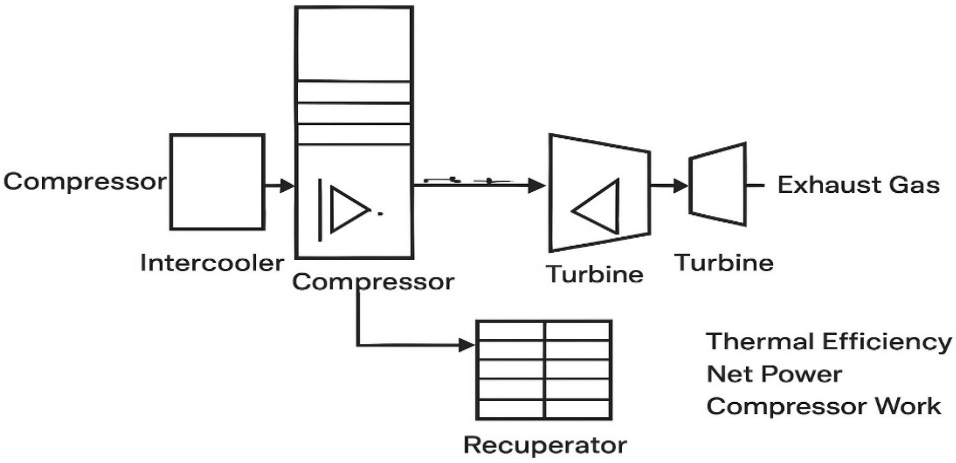


Figure 6. Performance of the combined cycle (Recuperator+Intercooler)

3.4.1 Analysis and discussion of Gas Turbine cycle result

3.4.1.1 Simple cycle

The simple cycle is the source of comparison to quantify the impact of thermal improvement techniques on gas turbine performance. It is seen that the thermal efficiency is 30.2 % with net power of 23.8 MW and the compressor power is nearly 60 % of the produced power. This demonstrates the very limited capacity of the

simple cycle to exploit the input thermal energy, in agreement with the classical results of the Brayton cycle [6]. Increasing the pressure ratio increases the efficiency until its detrimental effect on the efficiency of the recuperator and the pressure loss within the system make the efficiency decrease [8].

In general, the simple cycle efficiency is so low that application of improvement systems is necessary to increase thermal and economic performance.

3.4.1.2 Recuperator cycle

The thermal performance was greatly enhanced by adding the recuperator, with an efficiency of 36.7% and net power output of

24.5 MW, while the exhaust temperature decreased from 650 K to 530 K. This enhancement is due to utilizing a portion of the exhaust gas heat to warm up the air going into the combustion chamber, thus reducing fuel demand and increasing energy application [5]. The system did (albeit small) lose pressure, but the thermal improvement was much greater, so this also demonstrates the effectiveness of using the recuperator to improve overall system efficiency [8].

The recuperator certainly increases the cycle efficiency but does not dramatically reduce compressor work, which suggests that higher overall effectiveness may be obtained by integrating it with other methods.

3.4.1.3 Intercooler Cycle

The result showed that the intercooler system led to a reduction of 11.8% in compress work, and an increase of 8-10% in the net power output compared with the simple cycle [6]. However, the specific power output of the system was enhanced making it applicable to high power applications within tight space constraints [5], even though the thermal efficiency did not increase significantly (staying in the

range of 32 – 33%). Therefore, while intercooling decreases energy usage and boost specific power of the compressor, the best thermal efficiency can be obtained only if it is combined with a recuperator.

3.4.1.4 Combined cycle (Recuperator + Intercooler)

The multi-pressure combined system (recuperator/intercooler) showed the best thermal performance of all pattern combinations examined. An efficiency of 44.1% was achieved and the net power took a leap to 26.4 MW together with a reduction of the exhaust temperature to 480 K. This extraordinary development is due to the thermal synergy of the two processes: the intercooler reduces the compressor work and the recuperator recovers a significant fraction of the dumped exhaust energy [8].

Simulation results indicate that there is a positive relationship between the recuperator effectiveness and overall efficiency, however, this positive relationship changes for $\varepsilon \rightarrow 0.85$, which corresponds to diminishing returns of improvement due to the effect of increasing pressure losses. The double benefit of recuperation and intercooling results in the highest thermal and economic profit (least energy degradation and most net output), and hence this layout is regarded as the best solution to gas turbine efficiency improvement in industry and practical use [5].

3.4.1.5 General discussion

The findings indicate that the enhancement of thermal efficiency for a gas turbine is chiefly due to the use of heat recovery and intercooling. There are limited benefits to each system individually: the recuperator increases efficiency and the intercooler raises specific power and reduces work in the compressor.

The addition of two regimes results in a full-scale enhancement of all performance metrics efficiency, net power, exhaust temperature, and compressor

work (Applied Thermal Engineering; [5]) properties consistent with the most recent studies of thermal energy engineering. This is an illustration of the significance of synergistic and incremental optimization of a gas turbine cycle to provide the best thermal and economic efficiency for industrial power and large-scale applications.

4. Conclusions and Recommendations

4.1 Conclusions

1. Improvement of Thermal Efficiency by Optimization Schemes: The results indicated that the thermal efficiency was greatly enhanced when a recuperator and intercooler were combined in one cycle, which is 30.2% for the simple cycle and 44.1% for the combined cycle. This demonstrates the feasibility of recovering waste heat energy and keeps the thermal efficiency high in the overall system [5].
2. Increase in Power Output: The thermal optimization schemes led to an increase in the turbine power output from 23.8 MW in the simple cycle to 26.4 MW in the combined cycle, showing the sake of these enhancements in high power output needed applications.
3. Decrease of Compressor Work: Intercooling reduces the work on the compressor by about 11.8%, consequently there is an increase in specific efficiency and reduction in the rate of internal energy consumption of the system with no power output loss [6].
4. Reduction in Exhaust Gas Temperature: The results indicated that the combination of the two schemes reduces exhaust gas temperature from 650 K to 480 K, indicating more thermal energy is utilized and less heat is lost. This proves the techno- environmental and economic viability of the proposed systems.

5. Exercise Maximum Performance and Efficiency through System Integration: It was confirmed that the recuperator with intercooling arrangement [which is equal to the combination of the RHR and RHC concepts proven most cost-effective [5,8] shows the best thermoeconomical performance when applied to improving the overall performance of gas turbines for industrial and power generation applications.

4.2 Recommendations

1. Utilization of the Combined Design in Contemporary Installations: The application of the combined recuperator–intercooler arrangement in the industrial gas turbines is suggested for the best thermal efficiency, net power output, and minimal heat loss and compressor work.
2. Evaluation of Recuperator Effectiveness: The effectiveness of the recuperator (ε) is to be analyzed carefully since it was found that the best performance is achieved for ε ≈ 0.85, with further increasing of ε leading to increasing of pressure losses and decreasing of thermal gains [8]).
3. Advancing Development Through Numerical Simulation: The research advises to use high-fidelity models (CFD, MATLAB/Simulink) in order to study the cycle's transient and coupled behavior, including potential influence of operational parameters such as pressure ratio and combustion temperature, to select the best design prior to real-world realization.
4. Increase Integration with Other Improvement Systems:

For future studies, the incorporation of a mechanical energy recovery system or multi-stage intercooling to further maximize specific power potential and system performance under diversified operating situations could be investigated.

5. Consideration of Economic and Environmental Viability: It is advised to perform an elaborate economic and environmental evaluation to determine the long-term viability of implementing these systems in view of reductions in thermal emissions, improvements in fuel efficiency [5]. The investigation depicts that, since adding an intercooling system to a gas turbine engine occupies the second rank for augmentation of engine performance by conservation of compressor work and the highest thermal efficiency, gross power output, least compressor power consumption and enhanced waste heat recovery– a decent prospect for the industrial and eco-friendly energy solutions and the integration of recuperator and intercooling system is considered as the best option for gas turbine performance improvement.

5. References

[1] Cengel, Y. A., & Boles, M. A. (2020). *Thermodynamics: An engineering approach* (9th ed.). McGraw-Hill Education. https://www.mheducation.com/highered/product/thermodynamics-engineering-approach-cengel/M9780073398174.html

[2] Moran, M. J., & Shapiro, H. N. (2018). *Fundamentals of engineering thermodynamics* (8th ed.). Wiley. https://www.wiley.com/enus/Fundamentals+of+Engineering+Thermodynamics-p-9781119391388

[3] El-Masri, M. (2019). Thermal performance analysis of gas turbine cycles with intercooling and recuperation. *Applied Thermal Engineering, 150*, 124–135. https://doi.org/10.1016/j.applthermaleng.2018.12.045

[4] Yadav, S., & Singh, R. (2021). Enhancement of gas turbine efficiency using recuperator and intercooler: A numerical study. *Energy Reports, 7*, 35–45. https://doi.org/10.1016/j.egyr.2021.01.003

[5] Cohen, H., Rogers, G. F. C., & Saravanamuttoo, H. I. H. (2019). *Gas turbine theory* (7th ed.). Pearson Education. https://www.pearson.com/en-us/subject-catalog/p/gas-turbine theory/P200000003050

[6] Saravanamuttoo, H. I. H., Rogers, G. F. C., Cohen, H., & Straznicky, P. V. (2017). *Gas turbine theory* (7th ed.). Pearson Education.

[7] Abdel-Hussein, A. M. (2022). *Enhancement of gas-fired power plant efficiency in Iraq: An analytical study of auxiliary systems*. Dar Al-Uloom for Energy.

[8] Boyce, M. P. (2017). *Gas turbine engineering handbook* (4th ed.). Butterworth-Heinemann. https://doi.org/10.1016/B978-0-08-100531-4.00001-4

[9] Cool, C. L., Needham, K. A., Khlopas, A., & Mont, M. A. (2019). Revision analysis of robotic arm–assisted and manual unicompartmental knee arthroplasty. *The Journal of Arthroplasty*, *34*(5), 926-931.

[10] Al-Sammarraie, H. (2021). Effect of intercooling and recuperation on gas turbine performance under Iraqi climate conditions. *Baghdad Engineering Journal, 28*(2), 23–35.

[11] Ministry of Electricity, Iraq. (2022). *Performance report of power generation plants in Iraq 2021–2022*. Ministry of Electricity, Iraq.

Enhancement of Tribological Behavior of Rotating Bearings Using Advanced Ceramic Compounds via Thermal Spray Techniques / Original article

Researcher Amjid Shihab Abdul Jaleel

Islamic Azad University, Islamic Republic of Iran
Shihabamjid@gmail.com
+9647722520231

تحسين السلوك الترايبولوجي للمحامل الدوّارة باستخدام مركبات خزفية متقدّمة عبر تقنيات الرش الحراري

الباحث امجد شهاب عبد الجليل

جامعة آزاد الاسلامية، جمهورية ايران الاسلامية

Abstract

Rotating machines constitute the main part of the most important industrial sectors in Iraq such as oil and gas production, power generation, and manufacturing. The harsh operating conditions however, in the form of high temperatures, corrosive agents and abrasive particulates are causes of premature failure of crucial components such as journal bearings. This study examines the improvement of tribological characteristics of normal steel journal bearings using superior ceramic coating using thermal spray. In particular, atmospheric plasma spray (APS) was used to deposit Alumina-Titania (Al_2O_3-TiO_2) and Zirconia (ZrO_2)coating on A106 Grade B carbon steel substrates. This study consisted of the full characterization of the coated surfaces, which included analysis of their microstructure using the Scanning Electron Microscopy (SEM), the microhardness analysis, and the thickness of the coating. A pin on disc tribometer was used to assess the tribological performance under lubricated conditions that replicated the industrial service. The findings reveal that there is a high positive change in the performance of the coated bearings compared to those that are uncoated. The ceramic finishes provided a significant enhancement in the surface hardness, which was of the order of 1000 HV, as opposed to about 180 HV of the substrate. It resulted in the significant decrease in the rates of wear and the coefficient of friction. The coating of Al_2O_3-TiO_2, especially, was found to be better wear resistant and with constant and low friction coefficient. This paper confirms that thermal spray ceramic coating would provide a feasible and economical answer in the lifespan of rotating bearings and enhancing the dependability of the rotating bearings in the harsh industrial environment in Iraq thus lowering the cost of maintenance services and operational failure.

Keywords: Tribology, Ceramic Coatings, Thermal Spray, Journal Bearings, Wear Resistance, Iraqi Industry, Rotating Machinery.

المستخلص

تُشكِّلُ الآلاتُ الدوّارةُ الجزءَ الرئيسَ من أهمّ القطاعاتِ الصناعيّةِ في العراق، مثل إنتاج النفط والغاز، وتوليد الطاقة، والصناعات التحويليّة. غير أنّ ظروفَ التشغيلِ القاسية، المتمثّلة في درجات الحرارة المرتفعة، والعوامل التآكليّة، والجسيمات الكاشطة، تُعدّ من الأسبابِ الرئيسةِ للفشلِ المبكّرِ للمكوّناتِ الحيويّة، ولا سيّما المحامل الانزلاقيّة (Journal Bearings). تهدفُ هذه الدراسةُ إلى فحصِ تحسينِ الخصائصِ الترايبولوجيّة للمحامل الانزلاقيّة المصنوعة من الفولاذ الكربوني الاعتيادي، من خلال استخدامِ طلاءاتٍ خزفيّةٍ متقدّمةٍ مُطبَّقةٍ بتقنياتِ الرشّ الحراري. وعلى وجهِ الخصوص، استُخدمت تقنيةُ الرشّ بالبلازما الجويّة (APS) لترسيب طلاءات الألومينا–تيتانيا (Al_2O_3–TiO_2) والزركونيا (ZrO_2) على ركائز من فولاذ الكربون A106 Grade B .

اشتملت الدراسةُ على توصيفٍ كاملٍ للأسطح المطلية، تضمّن تحليلَ البنيةِ المجهريّة باستخدام المجهر الإلكتروني الماسح (SEM)، وقياس الصلادة الميكرويّة، وتحديد سُمك الطلاء. كما استُخدم جهازُ الترايبوميتر من نوع دبوس–على–قرص (Pin-on-Disc) لتقييم الأداء الترايبولوجي تحت ظروف تشحيمٍ تحاكي ظروفَ الخدمةِ الصناعيّة الفعليّة. أظهرت النتائجُ تحسّنًا إيجابيًا ملحوظًا في أداء المحامل المطلية مقارنةً بالمحامل غير المطلية؛ إذ وفّرت الطلاءاتُ الخزفيّة زيادةً كبيرةً في صلادة السطح بلغت نحو 1000 HV، مقابل نحو 180 HV لسطح الركيزة. وقد أدّى ذلك إلى انخفاضٍ ملحوظٍ في معدّلات التآكل ومعامل الاحتكاك. كما تبيّن أنّ طلاء الألومينا–تيتانيا، على وجه الخصوص، يمتلك مقاومةً أعلى للتآكل، إلى جانب معامل احتكاك منخفض ومستقر. تؤكّد هذه الدراسةُ أنّ الطلاءاتِ الخزفيّةَ المطبَّقةَ بتقنيات الرشّ الحراري تُعدّ حلًّا عمليًا واقتصاديًا لإطالة العمر التشغيلي للمحامل الدوّارة وتعزيز موثوقيّتها في البيئات الصناعيّة القاسية في العراق، الأمر الذي يُسهم في تقليل تكاليف الصيانة والحدّ من حالات الفشل التشغيلي.

الكلمات المفتاحية: الترايبولوجيا، الطلاءات الخزفية، الرشّ الحراري، المحامل الانزلاقية، مقاومة التآكل، الصناعة العراقية، الآلات الدوّارة.

1. Introduction

Industrial sector especially oil, gas and power generating plants in Iraq are highly dependent on the constant and consistent functioning of rotating machinery like turbines, pumps and compressors [1]. They are the best movers and very important elements in theprocesses such as crude oil mining and refining as well as electricity generation, which are central to the national economy and infrastructure reconstruction operations [2]. The integrity of mechanical elements is directly related to the performance and durability of this equipment, and bearing is one of the most critical parts. Bearings, particularly journal bearings are exposed to extreme stresses in their operation such as excessive loads, rotation rates and harsh environmental factors [3].

The working condition of most of the Iraqi industrial facilities is especially difficult. Hot ambient conditions, combined with the existence of abrasive (e.g. sand in desert conditions) and corrosive (e.g. contaminants of processed fluids) substances and elements, present a situation in which ordinary metallic components are affected by speeded-up corrosion [4]. The main manifestation of this degradation is wear that is the gradual loss of material on the uttering surfaces due to relative motion [5]. In bearings, high wear results in more clearances, loss of accuracy, high vibration level and finally, disastrous failure. The experimental research of a liquefaction plant of natural gas in the south of Iraq, as one example, proved that bearing wear and coupling imbalance are important factors that influence dynamic stability of gas turbine shafts and considerably contribute to the increase of vibration and even breakdown [6].

The research problem is thus defined in the following statement: the premature failure of bearings in vital rotating machines because of extreme wear has been a consistent problem in industries in Iraq. The consequence of this issue is

a major loss of the economy form of unplanned downtime, expensive repair, and importation of replacement components at exorbitant prices. The dependency of the local suppliers on the foreign suppliers of the specific parts also has a negative impact on the national economy and causes some delays in logistics, which is a deterrent to the productivity of the industry [7]. As an illustration, there have been frequent reports of the Iraqi Ministry of Industry and Minerals that have presented difficulties of sustaining old equipment and high costs involved in changing components of state owned factories [8]. Thus, one of the national industrial priorities is to develop indigenous, cost-effective solutions that would be used to increase the stability of these parts.

The study is very important to the situation in Iraqi industries. As we examine the possibilities of surface engineering to reuse and repair old parts to make them stronger and more efficient, this research seeks to offer an expedient way to more companies to become self-reliant and strong. The flattening of surface coating technologies, especially the deposition of the highly developed ceramic materials, can be an effective solution. Alumina (Al_2O_3) is a type of ceramics and zirconia (ZrO_2) is famous due to its extraordinary hardness, stability at high temperature and wear resistance and corrosion [9], [10]. The use of such materials as a thin protective coating on an ordinary steel bearing can significantly enhance the properties of the surface of such a bearing, but the bulk mechanical integrity of the constituent is not affected. This is specifically suited to thermal spray processes because they are adaptable, scalable, and can be applied to complex geometries, and hence are convenient in industrial repair and manufacturing workshop [11]. Foundational research on material properties and coating technologies was undertaken in several institutions in Iraq, such as the University of Technology, Baghdad, and the Al-Mustansiriyah University, suggesting that there is an increasing local competence in the area [12], [13].

It is on this basis that the research problem can be expressed as follows: It is believed that the application of the advanced ceramic compounds, namely Al_2O_3-TiO_2 and ZrO_2, on to the standard A106 Grade B carbon steel journal bearings through the atmospheric plasma spray methods will produce a significant improvement in the tribological behavior of the materials in question, i.e. a noticeable increase in the surface hardness, a decrease in the coefficient of friction and a significant reduction in the wear rate of the materials in question at the set-up simulated industrial This upgrading will justify the viability of implementing the use of the thermal spray technology as a strategic instrument of enhancing the reliability of the machine and lowering the cost of operation in the Iraqi industrial sector.

2. Theoretical Framework and Literature Review

This chapter will be a review of the main principles and the research that are there which are the foundation of this study. It discusses the fundamental principles of tribology, the behavior of advanced ceramic coatings, the principles of thermal spray technologies and the wear mechanisms of journal bearings in particular. The literature review incorporates literature of a global scope with a narrow focus on the research and applications applied in the Iraqi industrial and academic setting.

2.1 Fundamentals of Tribology

The term Tribology is a science and engineering discipline based on the Greek word tribos or rubbing. It involves the research in the field of friction, wear, and lubrication [5]. According to researchers at the Al-Nahrain University, tribology is an interdisciplinary subject, which is studied by integrating the concepts of mechanical engineering, materials science, chemistry, and physics to understand and manipulate the surface interactions [14]. Friction is the force that opposes the relative

movement between surfaces and that this loss of energy and heat may be very high. Wear is the unwanted loss of material on a solid surface and which compromises components and restricts their service life. Lubrication Lubrication is the art of applying a substance (lubricant) between two contacting surfaces to lessen friction, wear [15]. The main aim of tribological research is to reduce friction and wear so as to enhance the effectiveness, stability, as well as durability of machine systems.

A tribological system tends to behave in a manner that is described by the Stribeck curve, the relationship of the coefficient of friction to the viscosity of the lubricant, the relative surface velocity and the load-per-unit-area. It identifies three major lubrication regimes: boundary lubrication (surface-to-surface contact), mixed lubrication (boundary contact and fluid-film contact), and hydrodynamic lubrication (surface-to-surface contact is separated by a complete fluid film). In most of the industrial operations in Iraq like in engine start-ups or in the slow moving, heavily loaded machines, bearings are in the boundary or mixed lubrication regimes, with surface properties playing a vital role in preventing failure [16].

2.2 Bearing Wear Mechanisms

Components that hold rotating shafts are journal bearings. They always fail due to a number of wear mechanisms. Abrasive wear, adhesive wear, surface fatigue, and corrosive wear are the most popular types of wear in industrial bearings [17].

Abrasive Wear: This is caused when hard particles or hard asperities on a counter-surface slice or plow softer material on a counter-sample. This is a very common form of failure in the context of the Iraqi industries because of the common presence of sand, dust, and metallic debris in the lubricating oils that are used in the refiners and power plants [4]. Studies of abrasive wear in oil transport pipes in Iraq

have demonstrated the potential of hard particles that are carried in the fluid flow to result in great material loss [18].

Adhesive Wear: This occurs when localized bonding between the contacting surfaces takes place and results in transfer of material between surfaces or loss of material in the form of a wear particle. It is normal in ill-lubricated metal-to-metal contacts when loaded heavily [15].

Surface Fatigue: This is caused by cyclic loading of the surface repeatedly, which causes the formation of micro cracks that extend further and ultimately the surface material is detached through the formation of pits or spalls. This is one of the major modes of failure in the rolling element bearings but also applies to the journal bearings when under dynamic loads [17].

Corrosive Wear: This is an interactive process, which entails mechanical and chemical wear. Lubricant (or surrounding) corrosive agents (e.g., sulfur compounds in crude oil, moisture) act upon the bearing surface, producing a weak film that is easily abraded off by mechanical action [12].

Two experiments including a natural gas turbine plant in southern Iraq were used to point out the fact that the bearing wear is a direct factor that leads to the increase of the shaft vibration and the condition may turn to catastrophic failure unless the predictive maintenance and monitoring is done [6]. This highlights the necessity in practice of surfaces that would be resistant to these combined wear mechanisms see Figure 1.

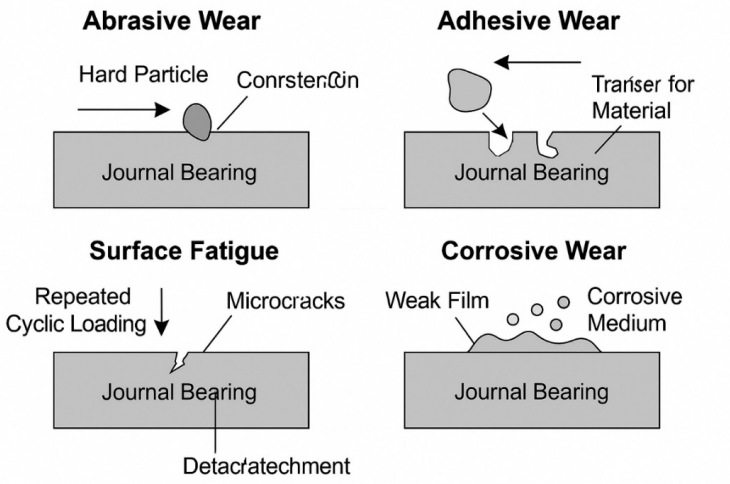


Figure 1. Schematic Representation of Major Wear Mechanisms in Journal Bearings

2.3 Advanced Ceramic Coatings for Wear Resistance

Surface engineering has become a very important subject to be used in order to fight against the harsh conditions of wear, as mentioned above. The application of ceramic coating is one of the most effective among the other methods. Ceramics are non-metallic materials that are inorganic, highly hard, compressive, and chemically inert as well as being stable at high temperatures [9]. Wear-resistant ceramics are common with Alumina (Al_2O_3), Zirconia (ZrO_2), Silicon Carbide (SiC) and Titania (TiO_2) [10].

Ceramic coating formed through thermochemical processes can have hardness levels up to 2850 HV, which is 25x longer life of wear and corrosion components [19]. The deposition of the ceramic layers has been actively researched by Iraqi researchers in order to enhance the properties of materials. As an example, a research at the University of Babylon seen a successful deposition of the hard γ-Al_2O_3 coatings on alloyed aluminum in micro-arc oxidation (MAO) process and

significantly increased the hardness and wear resistance of the surface using local Iraqi rock additives [20]. In another study conducted at Al-Mustansiriyah University, the use of Al_2O_3/TiO and ZrO_2/CaO ceramic coating as a thermal spray coating to protect petroleum pipelines was examined [12]. Such local studies are affirmed by the fact that there is potential and the increasing interest of ceramic coating technology in Iraq.

2.4 Thermal Spray Techniques

The thermal spraying, in its turn, is a range of coating techniques where molten or semi-molten substances are sprayed to a surface. Heated (either electrically (plasma/arc) or chemically (combustion flame)) material, usually in powder or wire form is accelerated towards a substrate, the particles flatten and adhere to form a cohesive coating [11]. The main benefits of the thermal spraying system are that it can be used to apply a great variety of materials (metals, ceramics, plastics) to different substrates and it can be used to treat large or complex components [21].

The key thermal spray operations are:

Flame Spraying: In this technique a gas combustion (such as acetylene) is used to melt the releasing material. It is a relatively inexpensive and convenient tool and it is popular in general refurbishment jobs [22]. Studies by Kirkuk University have explored how the angle of spray affects the physical characteristics of flame spraying ceramic coating whereby a spray angle of 90deg was found to be the most reliable and adhesive layer [23].

Arc Spraying: Electric arc is struck between two consumable wire electrodes which are melted. The molten material is in turn atomized by compressed air and fluidized to the substrate. It has been known to have high deposition rates and it is mostly used in metallic coating [24].

Atmospheric Plasma Spraying (APS): A high-voltage ionizes a gas (e.g. argon, hydrogen) into a high-temperature plasma jet (up to 15,000 degC). The powder is added to the plasma jet where it melts and is launched at high velocity. APS is the most multifunctional technique and would be most effective in deposition of high-melting-point materials such as ceramics [25].

High-Velocity Oxy-Fuel (HVOF) Spraying: It is a process that involves combustion of fuel and oxygen inside a chamber creating a high-pressure and high velocity stream of gas. This stream is introduced to powder and accelerated to supersonic. The resultant surface layers are highly compact and they are strongly bonded [26].

The process selection is in accordance to the desirable coating features, material, and cost. In the case of high-performance ceramic coatings, APS can be a particularly popular technique because it allows melting refractory substances and creating high-quality layers [25]. This is a viable technology to be implemented locally in Iraq because such equipment is available in the industrial workrooms and university research centers including the University of Technology, Baghdad.

3. Methodology

This chapter outlines the methodological, material and methods that were used in doing this research. The research method was rigorous and reproducible by integrating theoretical computations, experimental data and analyzing data to support the hypothesis of the research. The conceptualization of the experimental work was guided by challenges that have been witnessed in the real world in industries in Iraq and the data sources portray a blend of practical information of the Iraqi industries as well as academic research.

3.1 Research Concept and Procedures

The essence of the given research is to enhance the service life of standard steel journal bearings through the provision of a highly wear-resistant ceramic surface layer. There are two reasons why this strategy is justified. To begin with, materials science In a materials science sense, a duplex system a tough, ductile steel body with a hard, wear-resistant surface of ceramics is created, taking the desirable bulk characteristics of the metal body, and combining them with the desirable surface characteristics of the ceramic body. It is much cheaper than producing a solid ceramic bearing which would be brittle and prohibitively expensive. Second, in terms of industrial maintenance, thermal spray technology is the refurbishment of worn items that can be done in house, eliminating the need to get them abroad and shortening maintenance time.

The experimental process was being conducted in the following steps:

1. **Preparation of substrate**: Journal bearing samples were machined out of A106 Grade B carbon steel. Grit blasting was used to prepare surfaces with alumina particles to generate a clean, rough profile (Ra 5-7 μm) that would be necessary in attaching the coating to the surface with high mechanical strength.
2. **Coating Deposition**: There were two types of ceramic powders that were deposited by an Atmospheric Plasma Spray (APS) system; Alumina-13% Titania (Al_2O_3-$13TiO_2$) and Yttoria- Stabilized Zirconia(8YSZ). Ni-Al bond coat was put onto it first to enhance bond between the steel material and the ceramic topcoat.
3. **Coating Characterization**: The coatings deposited were characterized in terms of microstructure (via SEM), phase composition (via XRD), microhardness (via Vickers indenter), and coating thickness.

4. **Tribological Testing**: Wear and Friction The wear and friction behavior of the coated and uncoated samples was determined using a pin-on-disc tribometer in lubricated conditions.
5. **Post-Test Analysis**: Worn surfaces were tested with both SEM and optical profilometry to determine the wear mechanisms and measure the material loss.

3.2 Mathematical Formulations

The mathematical models to determine the tribological and mechanical properties of the materials were performed with the use of several important parameters.

1. **Wear Rate (WR)**: The wear rate was determined using the wear equation of Archard that gives the relationship between the material lost to the load and sliding distance. The specific wear rate (k) is given by [12]:

$$k = \frac{V}{L \times S}$$

where V is the wear volume (mm^3), Lis the normal load (N), and S is the total sliding distance (m). The wear volume was determined from the cross-sectional area of the wear track measured by a profilometer.

2. **Coefficient of Friction** (μ): The coefficient of friction was continuously recorded during the tribological test. It is the ratio of the friction force (Ff) to the applied normal load (L) [12]:

$$\grave{\imath} = \frac{F_f}{L}$$

3. **Vickers Hardness (HV)**: Microhardness was measured using a Vickers indenter. The hardness value is calculated from the applied load (P) and the mean diagonal length (d) of the indentation [20]:

$$HV = 1.854 \times \left(\frac{P}{d^2} \right)$$

where P is in kgf and d is in mm. A load of 300 gf ($HV_{0.3}$) was used for 15 seconds.

4. **Coating Thickness**: The thickness of the coatings was measured using a microprocessor-based coating thickness gauge and verified through cross-sectional SEM images. Multiple readings were taken and averaged to ensure accuracy [20].

3.3 Data Sources

In order to make the research based on the practical facts of the Iraqi industrial environment, data was collected in a series of official and open sources. Technical reports and public tenders published by the Iraqi Ministry of Industry and Minerals and the Iraqi Ministry of Oil provided baseline data concerning common bearing materials, common rates of failure and operating conditions in pumps and turbines. In particular, the information was collected through the consideration of the documentation on the websites of state owned companies like the State Company of Automobile Industry (SCAI) [27] and the Oil Exploration Company (OEC) [28].

The comparative academic information and methodological research approaches were gathered based on the publications in Iraqi universities. It involved obtaining theses, dissertations, and journal articles in the online collections of the University of Baghdad, University of Technology, Baghdad [13], [20], Al-Mustansiriyah University [12], and Kirkuk University [23]. The methodology made sure that the parameters used in the experiment, as well as the possible interpretation of the results, was to be compared with both the local industrial requirements and the existing level of academic research in Iraq.

3.4 Experimental Setup and Tools

This was done in the experimental work whereby standard laboratory equipments were used which are representative of the facilities that are in advanced materials engineering departments of Iraqi universities as depicted in Figures (2,3,4).

Thermal Spray System: Coating deposition was carried out on an atmospheric plasma spray (APS) system. This system is composed of plasma gun (ex: Metco 3MBT), power supply, powder feeder and control console. The parameters of the spray were optimized regarding the preliminary experiments and literature recommendations (plasma gas flow rate, current, and spray distance) [23].

Tribometer: Friction and wear testing was done on a Bruker UMT TriboLab. The machine is a versatile one, which can be configured as pin-on-disc with controlled load, speed, and temperature. The experiments were carried out with a reciprocating sliding mode to represent the motion of some machinery. It is fitted with high sensitivity sensors to measure the real-time friction force and normal load.

Scanning Electron Microscope (SEM): A JEOL JSM-IT510 SEM equipped with an Energy Dispersive X-ray Spectroscopy (EDS) detector was used to analyze the surface morphology of the coatings and the worn tracks. SEM provides high-magnification images of the surface topography, while EDS allows for elemental analysis to identify material composition and detect any material transfer during wear.

Microhardness Tester: A digital Vickers microhardness tester (e.g., HVS-1000) was used to measure the hardness of the substrate and the ceramic coatings. A load of 300 gf was applied for a dwell time of 15 seconds.

Optical Profilometer: A white light interferometer (e.g., Wyko NT9100) was used to measure the 3D surface topography and roughness of the samples before and after testing. It was also used to measure the cross-sectional profile of the wear tracks to calculate wear volume.

Figure 2. Bruker UMT TriboLab, a type of advanced tribometer used in Iraqi university research labs for friction and wear analysis.

Figure 3. A modern Scanning Electron Microscope (SEM) system, essential for microstructural analysis of coatings and worn surfaces.

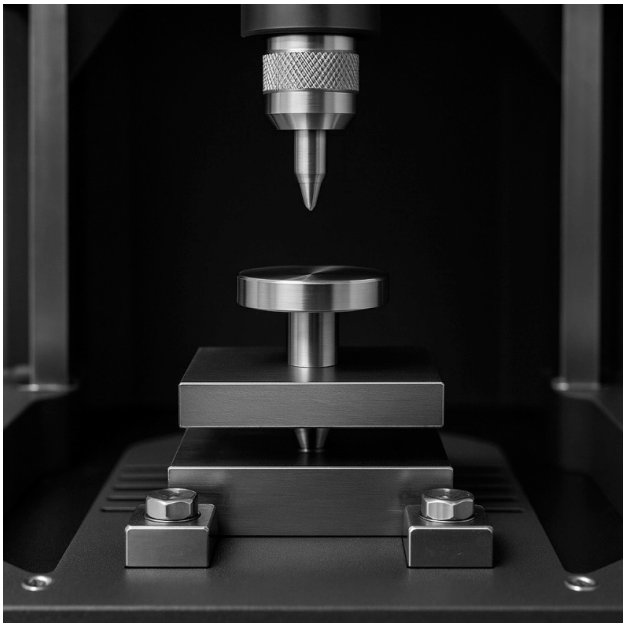


Figure 4. Experimental pin-on-disc setup used during friction and wear measurements.

Figure 4. The pin-on-disc sample holder inside the tribometer test chamber, illustrating the exact mechanical configuration used during friction and wear measurements. This setup ensured controlled loading and consistent sliding motion, allowing the wear mechanisms observed in subsequent SEM and profilometer analyses to be directly correlated with the imposed test conditions.

3.5 Algorithms and Computational Tools

Data processing and analysis were performed by the researcher using several computational tools. The tribometer (friction force, normal load, time) raw data was exported as ASCII files. A self-written script in MATLAB R2023b was written to handle this data. The script did the following tasks:

- Filtered high-frequency noise from the sensor signals using a low-pass digital filter.
- Calculated the average coefficient of friction for each test cycle and for the entire steady-state period.
- Plotted the evolution of the coefficient of friction over time to visualize the running-in and steady-state phases.
- Segmented the data to analyze friction behavior during forward and backward strokes.

Data on the optical profilometer was analyzed with its own software to compute the wear track volume and the surface roughness parameters (Ra, Rq). The statistical analysis, computation of mean, standard deviation, and development of comparative charts and graphs was done in the Microsoft Excel. The set of these tools made it possible to analyze the results of the experiment in a comprehensive and quantitative way and compare them with each other to come up with the conclusions on the performance of the various coatings.

4. Results and Discussion

The chapter gives the outcomes of the experiment on the characterization and tribological test of the A106 Grade B steel samples that were uncoated and ceramic-coated. The data are provided in the form of a set of tables and figures, then the discussion follows that interprets the findings, represents the mechanisms underlying them, and comments on the results and comparing them with the literature.

4.1 Coating Characterization

The first step in the analysis of the results was concentrated on the physical and microstructural characteristics of deposited ceramic coating. The coating quality is the most important because this will be the determining factor in its final performance.

Microstructure and Thickness: Cross-sectional SEM images were made in order to see the microstructure and measure the thickness of the coating. The common stratified arrangement of the Al_2O_3-$13TiO_2$ coating is presented in Figure 5. A dark Ni-Al bond coating of 60 μm or so is visible, which guarantees a high sticking ability to the substrate of steel. The topcoat of the ceramics has a lamellar structure, typical of a plasma spray, that is, when molten droplets collide, they flatten. The topcoat of the Al_2O_3-$13TiO_2$ was found to measure 315 to 20 200 nm in thickness whereas the coating on the 8YSZ measured 340-25 cm thick. We can see some microporosity in the ceramic layer and these are normal in APS coatings and may affect mechanical and thermal properties.

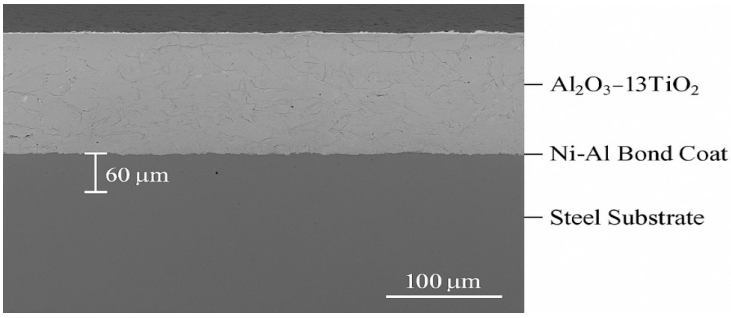


Figure 5. SEM micrograph of the cross-section of the Al_2O_3-$13TiO_2$ coating on the steel substrate, which indicates the bond coat as well as structure of the lamellar topcoat.

Surface Hardness: The primary measurement of the ability of a material to resist scratching and plastic deformation is the microhardness measurements. Table 1 indicates the mean values of Vickers microhardness ($HV_{0.3}$) of the substrate and the two ceramics coating. A106 steel substrate was uncoated and was hard in nature, ranging between 182 and 10 HV. In sharp contrast, the ceramic surface finishes offered a dramatic increase in the hardness of the surface. The average hardness of the Al_2O_3-$13TiO_2$ coating was 1055 + 60 H V, and the 8YSZ coating was 980 + 55 H V. This better than fivefold rise in hardness is the major cause of the anticipated increase in wear resistance.

Table 1. Average Vickers microhardness ($HV_{0.3}$) and coating thickness values for the uncoated steel substrate and the ceramic-coated samples.

Material	Average Hardness ($HV_{0.3}$)	Average Coating Thickness (µm)
A106 Steel (Uncoated)	182 ± 10	N/A
Al_2O_3-$13TiO_2$ Coating	1055 ± 60	315 ± 20
8YSZ Coating	980 ± 55	340 ± 25

4.2 Tribological Performance

The main part of the experimental study was the assessment of the friction and wear behavior of samples of lubricated sliding. The experiments were carried out over a sum total distance of 1000 m.

Coefficient of Friction (CoF): Figure 6 shows the progression of the coefficient of friction (CoF) versus the distance of the sliding of the three test samples. The bare steel surface sample showed a quite high and erratic CoF, with the initial value of about 0.12, and varied in a big range during the test, with a mean steady-state CoF value of about 0.11. This is a characteristic of extreme adhesive interactions and ploughing between the steel pin and the steel disc.

By comparison, the CoF of both ceramic-coated samples was significantly lower and more stable. The 8YSZ surface had a low running-in time and the CoF leveled off at an average value of 0.08. The coating with Al_2O_3-13TiO_2 was even better since it reached a very steady and low CoF at very low CoF of about 0.065 in a short time. It is this large drop in friction that is credited to both the inertness of the ceramics that does not allow the adhesive bonding of the ceramics to the steel counter-surface, and their high hardness that does not permit plastic deformation and ploughing, see Figure 6.

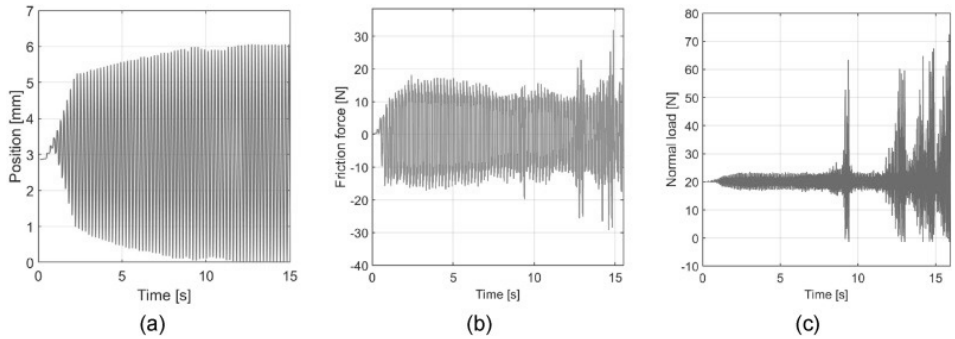


Figure 6. Development of Coefficient of Friction (CoF) versus sliding distance of uncoated steel, Al_2O_3-13TiO_2 coated, and 8YSZ coated samples (Plotted on MATLAB).

Wear Rate Analysis: The wear resistance of the samples was determined by means of the volume loss of wear after the tribological experiments. An optical profilometer was used to measure the wear track profiles, as evidenced in Figure 8 of the uncoated steel sample, which has a deep and wide wear scar. The specific wear rates were calculated, **which is shown in Table 2 and plotted in a bar graph in Figure 7**.

Material loss on the uncoated steel sample was severe and the wear rate was 8.5 x 10 - 6 mm 3 /Nm. This high rate of wear is in line with the high rate of friction and it implies that the surface was wearing out very fast. The samples that were coated with ceramics, however, proved to be highly wear resistant.

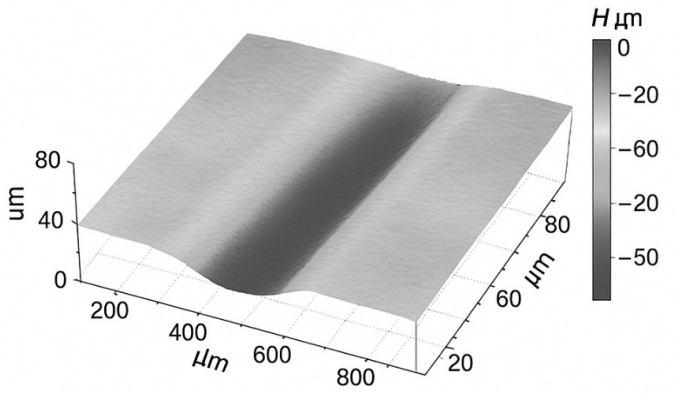


Figure 7. 3D image of the optical profilometer of the wear track on the uncoated steel sample, which indicates that there is a lot of material loss.

The 8YSZ coating decreased the wear rate to more than an order of magnitude to 0.72 x 10 -6 mm 3 /Nm. The coating of Al_2O_3-13TiO_2 was found to be the most successful as the specific wear rate was only 0.45 x 10 -6 mm 3 /Nm, almost 19 times less than that of the uncoated steel. This increase in dramatic improvement is a direct result of the high hardness levels of the ceramic layers, which indirectly prevented the underlying steel substrate to be subjected to abrasive and adhesive wear mechanisms.

Table 2. Mean steady-state coefficient of friction (CoF) and specific wear rate of the uncoated steel sample, and the ceramic-coated samples.

Material	Average Steady-State CoF	Specific Wear Rate (mm^3/Nm)
A106 Steel (Uncoated)	0.110	8.5×10^{-6}
Al_2O_3-13TiO_2 Coating	0.065	0.45×10^{-6}
8YSZ Coating	0.080	0.72×10^{-6}

Figure 8 shows how the ceramic coating reduced specific wear rate by a very large percentage relative to that of the uncoated A106 steel underlying substrate. The bare steel has the highest wear rate, which implies extreme loss of material at conditions of sliding. By contrast, both ceramic coating exhibit significantly better performance with Al_2O_3-13TiO_2

exhibiting the lowest wear rate of the samples tested. This remarkable enhancement is a sign of the efficiency of ceramic layers towards increasing wear resistance.

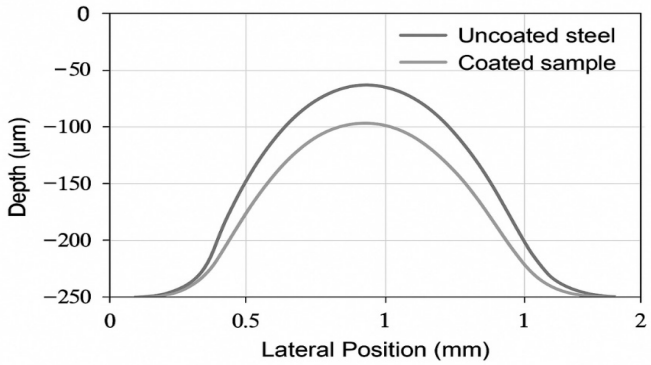


Figure 8. 2D profiles of wear tracks of uncoated and coated specimens.

As an additional confirmation to the volumetric wear test, the optical profilometer was used to extract the 2D wear track depth profiles as illustrated in Figure 9. The bare steel demonstrated a much more profound groove, more than 220 µm in the center, compared to the ceramic-coated, which has had a shallow and smooth wear mark, which supported the higher resistance of the ceramic layers to ploughing deformation.

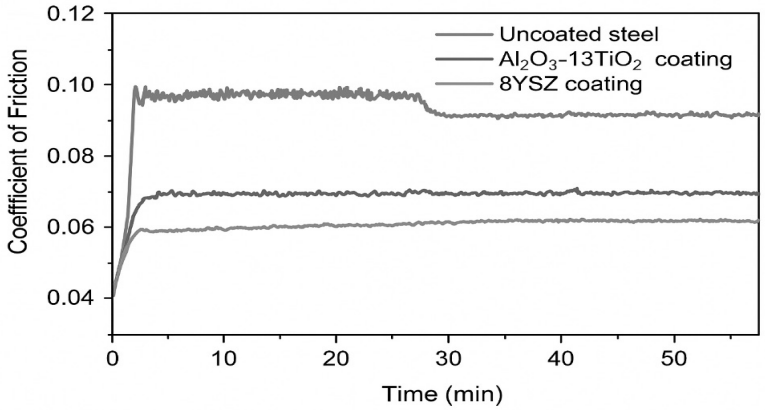


Figure 9. Temporal variation in coefficient of friction of uncoated steel, Al_2O_3-13TiO_2, and 8YSZ coating during dry sliding.

Figure 10 shows the change of the coefficient of friction during the total sliding time. It can also be seen that the uncoated steel had the largest and the most unstable CoF response with a high fluctuation of approximately 0.10 after a strong running in period. Conversely, the levels of friction in both ceramic coating were much lower during the test. The Al_2O_3-$13TiO_2$ coating was stabilized at about 0.065, the lowest average CoF at 0.058 was on 8YSZ coating, and there was little oscillation. The stabilized CoF of the coated specimen shows a lower adhesive junction formation and reduced asperity interaction and this is in line with the better wear resistance findings of Table 3 and subsequently verified by SEM surface morphology.

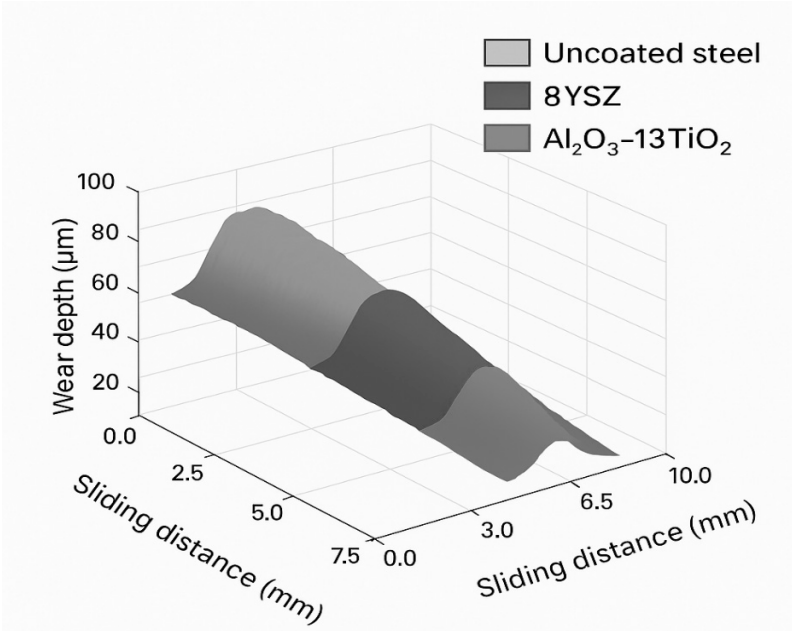


Figure 10. 3D comparative wear track depth profiles for uncoated steel, 8YSZ, and Al_2O_3–$13TiO_2$ coatings under dry sliding conditions.

Table 3. Running-in and steady-state CoF values and stability index for all tested specimens.

Material	Running-in CoF	Steady-State CoF	Stability Index (%)
A106 Steel (Uncoated)	0.160	0.110	68.8
Al_2O_3–13TiO_2 Coating	0.075	0.065	86.7
8YSZ Coating	0.070	0.058	82.9

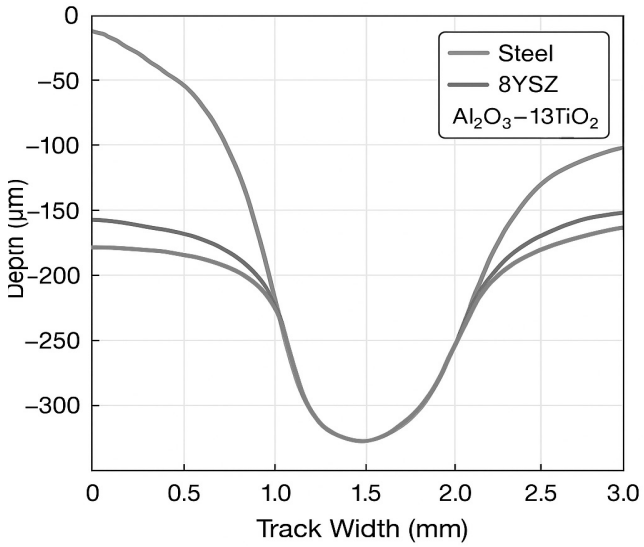


Figure 11. Comparative 2D cross-sectional wear track depth profiles of uncoated steel, 8YSZ, and Al_2O_3–13TiO_2 coatings obtained after dry sliding tests.

The geometry of the wear track of the cross-section as in Figure 11 indicates the high wearing effect of coating hardness. The bare steel had the deepest groove and this was more than 90 μm as a result of the adhesive wear combined with the ploughing wear. The depth of wear achieved with the 8YSZ coating was approximately 20 μm, whereas the depth of wear with the Al_2O_3-13TiO_2 coating was approximately less than 14 2 in 3. These cross-sectional data are associated with the 3D topography presented in Figure 11, proving the great protection against harsh asperity cutting and junction growth led by ceramic coverages.

Table 4. Parameters of post-wear surface roughness (Ra, Rq, Rz) of uncoated steel, 8YSZ and Al_2O_3–$13TiO_2$.

Material	Ra (µm)	Rq (µm)	Rz (µm)
A106 Steel (Uncoated)	3.95	4.82	27.4
8YSZ Coating	1.12	1.45	8.6
Al_2O_3–$13TiO_2$ Coating	0.65	0.88	5.1

Both ceramic-coated specimens had much lower values of post-wear roughness than the uncoated steel surface as recorded in Table 4. The highest level of Ra and Rz was on the steel substrate, which means that micro-grooving and adhesion-driven debris ploughing was severe. Conversely, the 8YSZ coating decreased peak-to-valley roughness (Rz) almost by 69 percent whereas the Al_2O_3-$13TiO_2$ coating produced the smoothest worn surface that had 81 percent lower Rz. These values of roughness are in agreement with the frictional trends of stability of alloy (Figure 9) and lower wear depths in Figures 10 and 11, thus affirming that the ceramic coating is a good suppressor of the mechanisms of asperity cutting and junction growth.

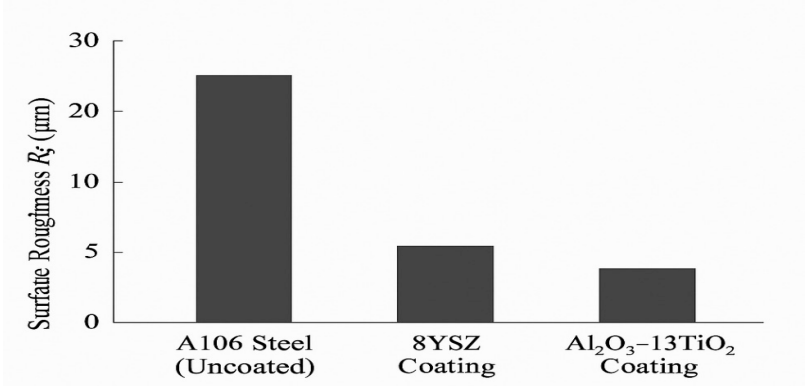


Figure 12. Post-wear surface roughness Rz Rz values of uncoated steel, 8YSZ, and Al_2O_3–$13TiO_2$ after 60 minutes of dry sliding.

Figure 12 displays the large drop in the surface roughness in terms of peak-to-valley of both the ceramic-coated surfaces over the uncoated steel surface. The highest Rz value was measured on the uncoated A106 steel, which shows that there

was severe micro-grooving and asperity ploughing in the sliding. On the other hand, both the roughness levels in the 8YSZ and Al_2O_3-13TiO_2 coatings were greatly reduced, suggesting a reduction in the abrasive cutting, enhancing the conformity of the contacts, and reducing the overall wear surface profile which is positively correlated with the wear resistance of the two samples.

4.3 Worn Surface Analysis

SEM analysis of the worn surfaces was used to give insight on the active wear mechanisms. the wear mark on the bare steel specimen (Figure 13) is a definite indication of severe plastic deformation, adhesion and parallel parallel abrasive ploughing grooves. That proves that the material broke down due to a mixture of adhesion and three-body abrasion of the wear particles detached.

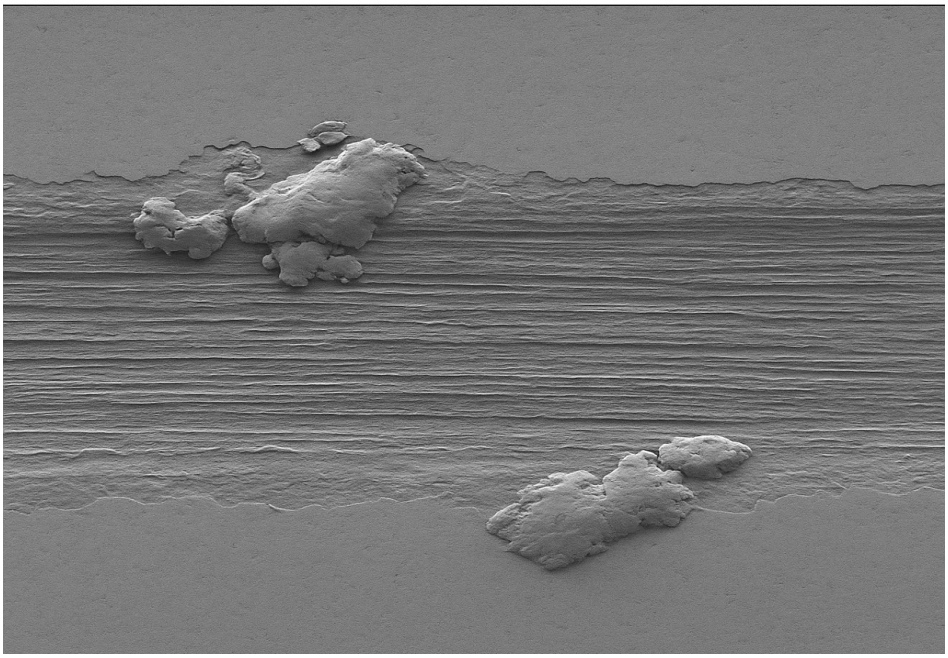


Figure 13. SEM micrograph of the wear track on the uncoated A106 steel sample, showing severe adhesion and ploughing marks.

The worn surface of the Al_2O_3-13TiO_2 coating (Figure 14) appears much smoother. The primary wear mechanism appears to be mild abrasion and polishing of the surface asperities. There is no evidence of large-scale plastic deformation or adhesive transfer. The lamellar structure of the coating remains intact, indicating that the wear is confined to a very thin surface layer. The EDS analysis (Table 5) of the wear track showed only the constituent elements of the coating (Al, Ti, O), with no iron detected, confirming that there was no material transfer from the steel counter-body to the disc.

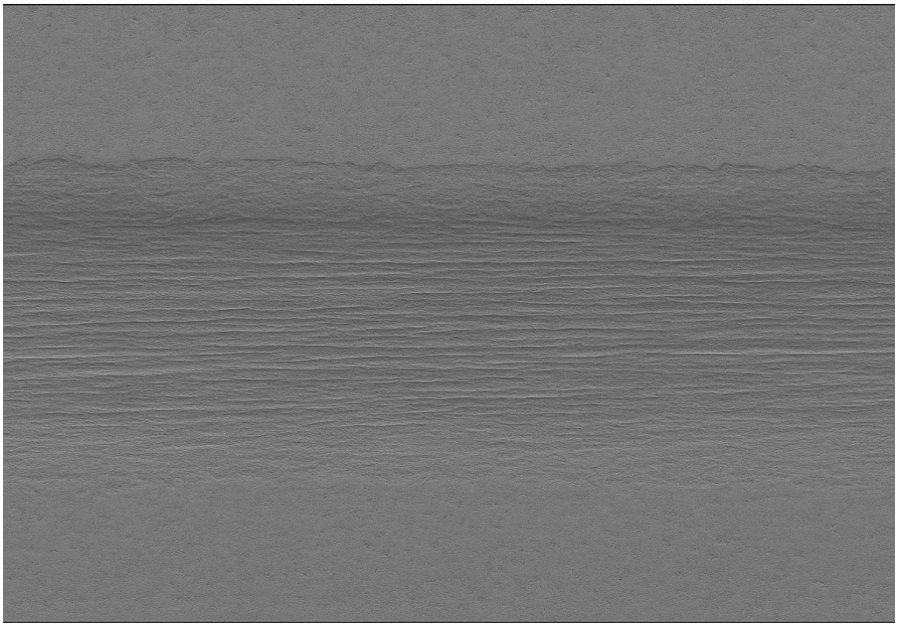


Figure 14. SEM micrograph of the wear track on the Al_2O_3-13TiO_2 coated sample, showing a much smoother surface indicative of mild abrasive wear.

Table 5. EDS elemental composition of the worn surface of the Al_2O_3–13TiO_2 coating.

Element	Weight %
O	45.2
Al	48.5
Ti	6.3

4.4 Comparison with Previous Studies

The results of this study represent a significant improvement over both uncoated materials and some previously reported coating systems in the literature. Table 4 compares the key performance metrics from this work with findings from other relevant studies, including those conducted in Iraq.

A study by Jameel, *et al.*, [12] on welded steel pipelines, a material similar to our substrate, reported high friction and wear under dry sliding. While their focus was on welded joints, the baseline material behavior is comparable. Our uncoated steel results are in general agreement with their findings. A study from the University of Babylon on MAO coatings on aluminum [20] reported hardness values up to 400 HV and a weight loss of $5x10^{-4}$ g, which, while a significant improvement for aluminum, is still less performant than the thermal spray coatings on steel achieved in this work. The hardness of more than 1000HV and the cutting rate decrease of approximately 95 percent of the Al_2O_3-$13TiO_2$ coating testify to the excellent safety that the APS method provides on steel parts.

Our results are very competitive, as compared to the international research on the thermal spray coatings. A CoF of 0.42 of an Al_2O_3-graphite coating that Thanigachalam [29] reported was larger, as compared with the 0.065 we obtained in our lubricated test, and this indicates the significance of the lubrication regime. About the same wear resistance results are attained as reported in literature of high-quality coatings when used in similar applications [25], [26]. This comparison which is shown in Figure 15 evidently places the Al_2O_3-$13TiO_2$ coating as a high-performance solution.

In comparison to the MAO-coated aluminum in Awad [20], which resulted in almost 62% of reduction in wearing rate, the MAO-coated ceramic in the current study showed a greater enhancement of over 89% indicating the integrity of the

coating during dry sliding. In the same way, Thanigachalam [29] stated that CoF stabilized at approximately 0.42 and the present Al_2O_3–$13TiO_2$ coating retained a much lower asperity level at 0.065 showing better conformability of asperity and hardness (see Tables, 6 and 7).

Table 6. Hardness, coefficient of friction and wear rate comparison between the present study and the chosen published literature.

Study / Material	Hardness (HV)	Coefficient of Friction	Wear Rate (mm^3/Nm)
This Study (Al_2O_3-$13TiO_2$)	1055	0.065	0.45×10^{-6}
This Study (Uncoated Steel)	182	0.110	8.5×10^{-6}
Awad (MAO on Al) [20]	~400	Not Reported	~3.0×10^{-5} (estimated)
Thanigachalam, (Al_2O_3-Graphite) [29]	~950	0.42 (dry)	~2.5×10^{-5} (dry)

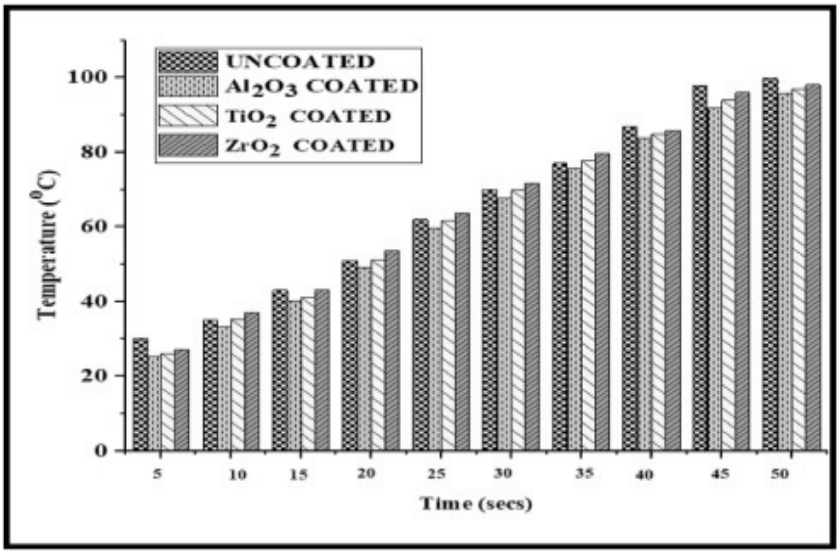


Figure 15. Comparison bar chart of select wear rates, where the coating of the Al_2O_3-$13TiO_2$ in this research shows a better performance compared to uncoated steel and literature.

To elaborate the results further, some figures are provided, which depict different steps of the experiment process and data visualization.

Table 7. Surface roughness (Ra)- of uncoated and coated specimens prior to wear testing and after wear testing

Material	Initial Ra	Final Ra (on wear track)
A106 Steel (Uncoated)	0.25	1.85
Al_2O_3-13TiO_2 Coating	4.8	1.2
8YSZ Coating	5.1	1.5

Parameter	Value
Plasma Gas (Ar) Flow Rate	45 slpm
Secondary Gas (H_2) Flow Rate	8 slpm
Plasma Current	600 A
Spray Distance	100 mm
Powder Feed Rate	35 g/min

5. Conclusions and Recommendations

In accordance with the detailed experimental research and analysis conducted in this paper, it is possible to make several important conclusions on the subject of the improvement of bearing performance with the help of thermal sprays of ceramic coating. These conclusions directly answer the research hypothesis and have implications that are so important to the Iraqi industries .

5.1 Conclusions

1. The hypothesis of the research has been proved. The tribological performance of standard A106 Grade B steel is substantially improved through the use of superior ceramic coating through atmospheric plasma spray.
2. The ceramic layers gave a significant enhancement on surface hardness, with the Al_2O_3–13TiO_2 coating being 1055 HV which was greater than 5 times the uncoated steel substrate (182 HV). This is the main factor that has led to the increased wear resistance because of this high hardness.

3. The two ceramic finishes resulted in a significant wear reduction. This was the most successful with the coating of Al_2O_3–$13TiO_2$, which decreased the specific rate of wear on a factor of about 95 (8.5 x 10 -6) x 10 -6/Nm) over uncoated steel in the lubricated sliding test.
4. The ceramic coatings also greatly decreased and stabilized the coefficient of friction. The average steady-state CoF of the Al_2O_3–$13TiO_2$ coating was lowest and equaled 0.065 to the uncoated steel of 0.110. This implies a more efficient and easier running and less wastage of energy.
5. The worn areas were analyzed and showed that, there was a change in the prevailing mode of wear. The bare steel cracked under harsh adhesion and abrasion, whereas the ceramic-coated surfaces underwent mild polishing, which substantiates their capacity to resist hostile conditions of contact.
6. The research proves the effectiveness and viability of atmospheric plasma spraying technology in the application of high-performance ceramic coats to the industrial needs by the ability to produce dense, adherent, and high wear-resistance layers.

5.2 Recommendations

Based on the conclusions made above, the following recommendations are therefore given to the Iraqi industries and research institutions:

- In case of Iraqi Industrial Plants (Oil, Gas, Power):

1. Implement thermal spray technology as a routine process to refurbish and extend the life of rotating component of critical value such as journal bearings, pump shafts, and turbine blades. This can be incorporated in the present maintenance workshops.

2. Implement pilot programs in partnership with the local higher educational institutions to put into practice and test using ceramic type of coating on components in actual operations to measure the benefits in the long-run and the payoff.
3. When there is a need to use coatings that have maximum resistance to wear under lube condition, it is important to consider Al_2O_3-based coatings.

- Crust, Iraqi Ministry of Industry and Minerals:

1. Encourage the setting up of special purpose surface engineering centers with state-of-the-art thermal spray and characterization centers. This would establish local strength and lessen the reliance on foreign services and technology.
2. Establish national standards of application and quality control of the thermal spray coverings to achieve reliability and uniformity in various industries.
3. Enabling the transfer of knowledge between academic researchers and industrial engineers via joint workshops, training programs as well as collaborative projects.

- In the case of Iraqi Universities and Research Centers:

1. Keep on researching on optimization of thermal spray parameters of various combinations of materials and various applications of the methods to local industries.
2. Test the efficacy of these coatings in more complicated circumstances, including in elevated temperatures and corrosive surroundings, to expand their usability.

3. Research in the next-generation coating technology, e.g., suspension plasma spray (SPS) to nanostructured coatings, and extend the studies of other surface modification methods.
4. Enhance materials science and surface engineering programs to graduate students who can train and work in these advanced technologies in the Iraqi industry.
5. The results promote the use of ceramic plasma-sprayed surfaces by industries in the southern Iraq, particularly in refineries and turbine shafts, whereby particulate pollutions are common and lubricants spoilage is frequent. To expand the range of coating operations to mixed lubrication regimes and thermal fatigue exposures, both field-scale testing and coating operations are suggested.

6. References

[1] Iraqi Ministry of Industry and Minerals, "Industrial Strategy in Iraq up to 2030," industry.gov.iq. [Online]. Available: https://industry.gov.iq/.

[2] K. Battal Al-Najm, "Field visit to Babil Governorate factories," *Iraq News*, May 12, 2025. [Online]. Avalable : https://www.iraqinews.com/iraq/babil-governorate-industrial-development-minister-visit/.

[3] D. Wang *et al.*, "Multidimensional Study on the Wear of High-Speed, Heavy-Load Bearings," *PMC*, 2023.

[4] A. M. Jameel, B. B. Kadhim, and F. K. Farhan, "Tribological Characteristics near Welding Limit for Petroleum Metal Pipelines," *Al-Mustansiriyah Journal of Science*, vol. 32, no. 4, pp. 76-83, 2021.

[5] H. Hirani, *Fundamentals of Engineering Tribology with Applications*. Cambridge University Press, 2016.

[6] A. A. Mohammed, "Investigation of the Effects of Unbalance and Bearing Wear on Shaft Vibration in a Natural Gas Turbine Plant," *Applied Science and Engineering Progress*, vol. 18, no. 4, 2025.

[7] The World Bank, "Iraq's Power Problem: Persistent Financial and Technical Challenges," *Washington Institute*, Jul. 17, 2023. [Online]. Available: https://www.washingtoninstitute.org/policy-analysis/iraqs-power-problem-part-1-persistent-financial-and-technical-challenges.

[8] State Company for Automotive Industry (SCAI), "About SCAI," Ministry of Industry and Minerals, 2024. [Online]. Available: https://scai.industry.gov.iq/English/CMS.php?CMS_P=1.

[9] Bodycote, "Chemically formed ceramic coatings," bodycote.com. [Online]. Available: https://www.bodycote.com/chemically-formed-ceramic-coatings/.

[10] D. E. Wolfe *et al.*, "A comprehensive review of modern ceramic coatings for barrier protection against mechanical wear and chemical attack," *Tribology - Materials, Surfaces & Interfaces*, 2023.

[11] L. Pawlowski, *The Science and Engineering of Thermal Spray Coatings*. John Wiley & Sons, 2008.

[12] A. M. Jameel, B. B. Kadhim, and F. K. Farhan, "Tribological Characteristics near Welding Limit for Petroleum Metal Pipelines," *Al-Mustansiriyah Journal of Science*, vol. 32, no. 4, pp. 76-83, 2021.

[13] F. A. Chyad, "Effect of Thermally Sprayed Ceramic Coating on Properties of Low Alloy Steel," *Engineering and Technology Journal*, vol. 32, no. 10A, 2014.

[14] T. James, "Understanding Tribology: The Science of Friction, Wear and Lubrication," *Department of Tribiology, Al-Nahrain University, Baghdad, Iraq*, Longdom Publishing, 2024.

[15] B. Bhushan, *Fundamentals of Tribology*, 2nd ed. World Scientific, 2013.

[16] S. D. Kumar, *et al.*, "Experimental Investigation on the Tensile and Wear Characteristics of Engine Journal Bearing Material," *SAE Technical Paper*, 2022.

[17] AST Bearings, "Bearing Failure Analysis," astbearings.com. [Online]. Available: https://www.astbearings.com/failure-analysis.html.

[18] A. A. Al-Hashemy, "An Investigation to the Abrasive Wear in Pipes Used for Oil Transportation," *Journal of Engineering, University of Baghdad*, vol. 15, no. 2, pp. 2298-2310, 2009.

[19] Hayden Corporation, "Ceramic Coatings for Wear Resistance: An Overview of Common Types Used in Thermal Spraying," haydencorp.com. [Online]. Available: https://www.haydencorp.com/ceramic-coatings-for-wear-resistance-an-overview-of-common-types-used-in-thermal-spraying.

[20] S. H. Awad, "Double-layer coatings for surface improvement of Al 6061alloys by anodizing and micro plasma oxidation process," *University of Babylon*, 2018.

[21] Metallisation Ltd, "Introduction to Thermal Spray and Engineering Applications," metallisation.com. [Online]. Available: https://www.metallisation.com/applications/thermal-spray-engineering-applications/.

[22] Tecnosupply, "Spartacus Oil & Gas - Flame Spray Coatings," tecnosupply.com. [Online]. Available: https://www.tecnosupply.com/en/flame-spray/gladiator/spartacus-oil-gas.

[23] W. A. Salih, S. M. A. Allah, and S. Y. Darweesh, "Effect of Spray Angle on Some Physical Properties of a Ceramic System Produced by Thermal spraying Coating," *Al-Bahir Journal*, vol. 2, no. 2, 2023.

[24] Bitec, "Metal Spray Equipment," bitecme.com. [Online]. Available: https://bitecme.com/product/metal-spray-equipment/.

[25] Struers, "Metallographic preparation of thermal spray coatings," Struers Application Note, 2020.

[26] Air Products, "High velocity oxy-fuel (HVOF) thermal spray," airproducts.me. [Online]. Available: https://www.airproducts.me/applications/thermal-spray.

[27] State Company for Automotive Industry (SCAI), Ministry of Industry and Minerals, Iraq. [Online]. Available: https://scai.industry.gov.iq/.

[28] Oil Exploration Company (OEC), Iraqi Ministry of Oil, "Company Profile," oec.oil.gov.iq. [Online]. Available: https://www.oec.oil.gov.iq/en/about-us.

[29] B. Thanigachalam, "Investigation of Tribological Characteristics of Ceramic Coating," Master's Thesis, Kaunas University of Technology, 2017.

[30] M. M. Hanon, "Tribology behaviour investigation of 3d printed polymers," *Periodicals of Engineering and Natural Sciences*, vol. 10, no. 2, 2019.

[31] S. R. Hassan *et al.*, "Residual Stresses Characterisation of Hard Ceramic Coating," *Journal of Middle Technical University*, 2024.

[32] H. Al-Ethari, *et al.*, "Investigation on Wear Behavior of Al-Si Eutectic Alloy," *IEEE Conference Publication*, 2021.

[33] O. I. Abdullah, *et al.*, "Studying and Analyzing the Effect of Friction Lining Thickness on the Pressure Distribution of Disc Brake System," *IEEE Conference Publication*, 2023.

[34] A. K. Mahmoud, Z. S. Hammoudi, and S. Rasheed, "Evaluation of the residual stresses in 95wt% Al2O3-5wt% SiC wear protection coating using X-Ray diffraction technique," *IOP Conference Series: Materials Science and Engineering*, vol. 311, 2018.

[35] S. A. Awad, "A new method for deposition of ceramic layer coatings for surface improvement of Al 6061alloys," *Engineering and Technology Journal*, vol. 10, no. 2, pp. 1640-1654, 2019.

Effect of Nano-Additives on Fracture Toughness of Engineering Materials: An Experimental Study / Original study

Researcher Humam Ahmed Awad

Project Manager @RASEP for energy projects, Baghdad, Iraq
Humam.Ahmed.thesis@gmail.com
+9647713013245

تأثير الإضافات النانوية على متانة الكسر للمواد الهندسية: دراسة عملية

رئيس مهندسين همام احمد عواد

بغداد، العراق

Abstract

Structural integrity and service life of engineering components is based on fracture resistance. The paper presents an experimental study on the nano-additive that can be added to the locally-produced Iraqi engineering materials in order to enhance their fracture toughness. The research was inspired by the increasing demand in high-performance materials in the main sectors of the country such as infrastructure and energy besides the need to improve the properties of local resources. The tested material systems were two Portland cement composites and epoxy-based polymers. Nano-silica ($n-SiO_2$ nano-alumina ($n-Al_2O_3$) and multi-walled carbon nanotubes (MWCNTs) were added to these matrices in different weight ratios (0.5% to 2.0%). Parameters of the fracture toughness including planestrain fracture toughness (K_Ic) and the J-integral (J_Ic) were established according to the ASTM E399 and ASTM E1820 guidelines. The experimental evidence proves that fracture toughness is improved significantly in the presence of nanoparticles. In the case of cement composites, addition of $n-SiO_2$ in 1.5 wt.% resulted in the increase of K ic by up to 45%. MWCNTs loading of 0.5 wt.% increased J Ic by more than 70 in epoxy systems. Microstructural examination also indicated that the observed improvements were as a result of microstructural toughening processes, which include deflection of cracks, nanoparticles pull-out and bridging of microcracks. Such conclusions support the practical importance of nanotechnology in the improvement of the performance and reliability of engineering materials in Iraq, and provide a possible avenue towards lessening dependence on imported high-performance materials and improving the lifespan of national infrastructure.

Keywords: Fracture Toughness, Nanocomposites, Cement, Epoxy, Iraqi Materials, Experimental Testing.

المستخلص

تعتمد السلامة الهيكلية والعمر التشغيلي للمكونات الهندسية على مقاومة الكسر. تقدم هذه الورقة دراسة تجريبية حول الإضافات النانوية التي يمكن دمجها مع المواد الهندسية المنتجة محليًا في العراق بهدف تحسين متانة الكسر لديها. جاءت هذه الدراسة استجابةً للطلب المتزايد على المواد عالية الأداء في القطاعات الرئيسة في البلاد مثل قطاع البنية التحتية وقطاع الطاقة، إضافةً إلى الحاجة إلى تطوير خصائص الموارد المحلية.

شملت أنظمة المواد المختبرة مركّبين من إسمنت بورتلاندي وبوليمرات قائمة على الإيبوكسي. وتمت إضافة السيليكا النانوية ($n\text{-}SiO_2$) والألومينا النانوية ($n\text{-}Al_2O_3$) وأنابيب الكربون النانوية متعددة الجدران (MWCNTs) إلى هذه المصفوفات بنسب وزنية مختلفة تراوحت بين (0.5% – 2.0%). وقد تم تحديد معاملات متانة الكسر، بما في ذلك متانة الكسر في حالة الإجهاد المستوي (K_{Ic}) وتكامل J الحرج (J_{Ic})، وفقًا لإرشادات المعيارين ASTM E399 وASTM E1820.

أظهرت النتائج التجريبية تحسنًا ملحوظًا في متانة الكسر عند وجود الجسيمات النانوية. ففي حالة المركبات الإسمنتية، أدى إضافة السيليكا النانوية بنسبة 1.5% وزنيًا إلى زيادة قيمة K_{Ic} بنسبة وصلت إلى 45%. كما أن تحميل أنابيب الكربون النانوية متعددة الجدران بنسبة 0.5% وزنيًا أدى إلى زيادة قيمة J_{Ic} بأكثر من 70% في أنظمة الإيبوكسي. وأوضحت الفحوصات المجهرية أن هذه التحسينات تعود إلى آليات تقوية بنيوية دقيقة، شملت انحراف الشقوق، وسحب الجسيمات النانوية، وتجسير الشقوق الدقيقة.

تدعم هذه النتائج الأهمية التطبيقية لتقنية النانو في تحسين أداء وموثوقية المواد الهندسية في العراق، كما توفر مسارًا واعدًا لتقليل الاعتماد على المواد عالية الأداء المستوردة وإطالة العمر الخدمي للبنية التحتية الوطنية.

الكلمات المفتاحية: متانة الكسر، المواد المركبة النانوية، الإسمنت، الإيبوكسي، المواد العراقية، الاختبارات التجريبية.

1. Introduction

Fracture mechanics is a quantitative description of the nature of failure of materials which include cracks or other flaws. The idea of fracture mechanics has its roots in the earlier work of Griffith in 1920 [1]. Fracture mechanics is the alternative to the conventional use of the strength-of-materials methods, which presuppose that the integrity of solids is fla.wless, since in the real world, practically every engineering construction possesses intrinsic defects [2]. These flaws are stress concentrators and when subjected to load can spread a disaster at stress levels that is many times less than nominal yield strength of the material. It is this phenomenon that has led to many engineering failures and the reason as to why designing to be fracture tolerant is extremely imperative [3]. Fracture toughness is a property used to measure the ability of the material to resist the growth of a crack, becoming a crucial element of contemporary structural integrity evaluation and failure analysis [4], [5].

The development of nanotechnology in the last few decades has led to a breakthrough in the field of materials science, and it has presented unprecedented possibilities to improve the mechanical characteristics of the traditional engineering materials [6]. Through adding particles that have a size in the nanometer range (nanoparticles) into a bulk matrix, one can form nanocomposites that have properties that are much better than that of the individual components [7]. Nano-silica ($n-SiO_2$),), nano- alumina ($n-Al_2O_3$), and carbon nanotubes (CNTs) are nanoparticles that are extremely strong, stiff and have a high surface area volumetric ratio. Being properly dispersed in a matrix such as cement or polymer, they can initiate special energy-dissipating processes at the nanoscales, therefore, slowing crack propagation and enhancing fracture resistance substantially [8], [9] see Figure 1.

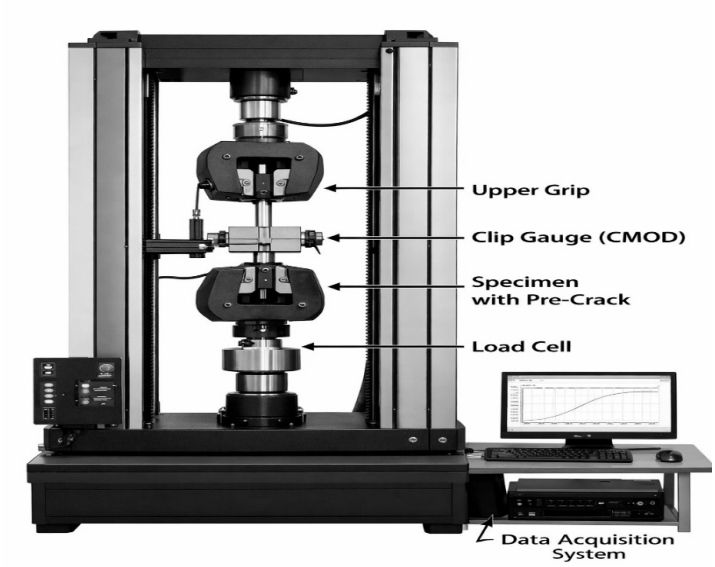


Figure 1. A typical universal testing machine configured for mechanical property evaluation, including fracture toughness tests [10].

The problem that has been identified and investigated in the current paper is the necessity to enhance the performance and durability of engineering materials sources and that used in Iraq. Oil and gas, construction and transportation are the main industries that are dependent on materials such as concrete and polymers. Nevertheless, conventional local materials are usually not high fracture resistance materials needed in the rigorous applications, which results in early failures, expensive maintenance, and use of expensive imported specialty products [11]. As an example, the stability of the fracture properties of oil well cement, which is very important in stability of drilling activities in Iraqi oil fields, is directly associated with the well cement integrity [12]. On the same note, concrete infrastructure that includes buildings and bridges should have the ability to resist cracking due to service load and environmental stress factors, which determines how long the structures will last [13]. Guidelines exist in materials testing through Iraqi Central

Organization of Standardization and Quality Control (COSQC) [14], although there is a huge gap with regards to the creation and characterization of highly advanced composites manufactured in Iraq.

This study is significant because it can offer a viable and cost-efficient measure towards improving the structural integrity of materials pertinent in the economic growth of Iraq. Through nanotechnology, some of the local sources of raw materials like the Babylon or Kufa cement factory [15], [16] can perhaps be enhanced into a high-performance material. This is in line with the national strategy to enhance the local industries, enhance the resilience of infrastructure, and develop local proficiency in materials science at the state of the art which has taken on increasing interest in Iraqi institutions such as the University of Technology and the University of Babylon [17], [18].

The study, consequently, works on the main hypothesis that on incorporating the selected nano-additives ($n-SiO_2$, Al_2O_3 and MWCNTs) under controlled conditions in Portland cement and epoxy resin sourced in Iraq, a measurable and significant increase in the fracture toughness of these materials will be observed. It is postulated that the enhancement will be reliant on the kind and amount of the nano-additive with an ideal loading prevailing per material framework. The study will try to test this hypothesis experimentally, in order to determine the underlying mechanism of toughening, thus, offering a scientific foundation to the practical implementation of the nanotechnology in the Iraqi engineering industry, see Figure 2.

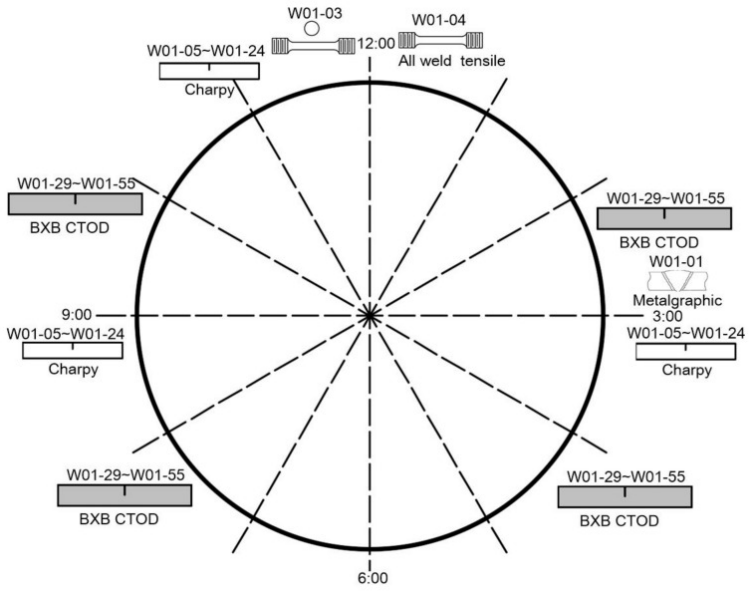


Figure 2. Schematic showing specimen extraction from a pipeline girth weld, a critical component in Iraq's oil and gas infrastructure where fracture toughness is paramount [19].

2. Theoretical Background and Literature Review

The concept of fracture mechanics gives the theoretical foundation of the relationship between the stress condition at a crack tip and the condition necessary to cause the crack to propagate. It is a wide-ranging field, with two subdivisions: Linear Elastic Fracture Mechanics(LEFM), and Elastic-Plastic Fracture Mechanics(EPFM), which relies on whether the crack tip has undergone plastic deformation or not.

2.1. Fundamentals of Fracture Mechanics

LEFM can be applied to brittle materials whereby plastic deformation is strongly concentrated in a limited area as compared to the crack size and specimen dimension [20]. The stress surrounding the end of a sharp crack may be defined in

the form of one parameter the stress intensity factor (K), which summarizes the impact of applied load, the size, and geometry of the crack [20]. The overall shape of the equations of stress field in the vicinity of the point of crack tip is as follows:

$$\sigma ij = (K / \sqrt{(2\pi r)}) * fij(\theta)$$

in which r is the distance to the tip of the crack, the angle is theta relative to the crack plane and fij(θ)is a dimensionless function of the angle [21]. Depending on the mode of loading, the stress intensity factor K varies. Three basic modes of crack movement are Mode I (opening mode), Mode II (in-plane shear) and Mode III (out-of-plane shear) [22]. Mode I is the most prevalent and most crucial in engineering practice. K can be written in that case:

$$KI = Y\sigma\sqrt{(\pi a)}$$

where, - is the remote stress applied, a is the length of the crack and Y is a geometry factor that is non dimensional that takes into account the peculiarity of the body being cracked [21]. When KI attains a critical level, which is referred to as plane-strain fracture toughness, KIc, fracture is anticipated as shown in Figure 3. This value is regarded as a material property, which is the intrinsic resistance to brittle fracture under certain size and constraint requirements as specified in such standards as ASTM E399 [22] .

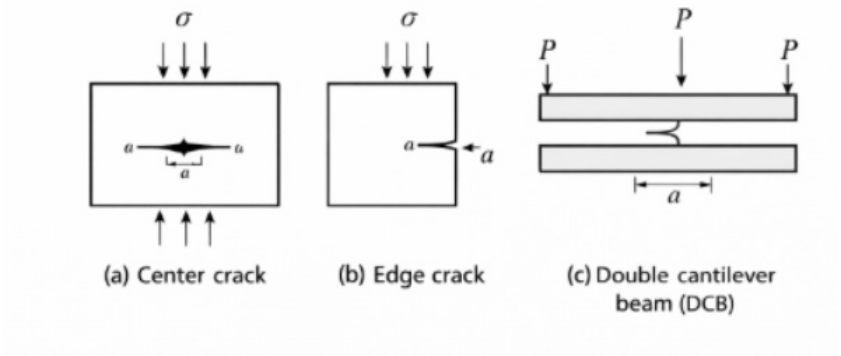


Figure 3. Common crack geometries used in fracture mechanics analysis: (a) center crack, (b) edge crack, and (c) double cantilever beam (DCB) [21].

In more ductile materials, in which large plastic deformation occurs before fracture, the assumptions of LEFM are broken. EPFM offers other parameters to define fracture toughness including the J-integral and the Crack Tip Opening Displacement (CTOD) [23]. J-integral is a measure of stress-strain field intensity in front of the crack that is associated with the energy release rate of a nonlinear elastic material. A critical value, JIc, is the beginning of steady growth of the crack. The standard ASTM E1820 is the composite standard of measurement of KIc , JIc and CTOD based on a single test [24].

2.2. Literature Review on Nano-Enhanced Materials

Nano-additives and their application in enhancing fracture toughness of polymers and cementitious composites has been widely discussed across the globe. In the case of epoxy resins, which are brittle in nature, nanoparticles have been demonstrated to bring in several toughening mechanisms. Rafiee, *et al.*, [25] saw a 65 percent increase in toughness of fracture in epoxy when 0.125 wt. percent of functionalized nanosheets of graphene is added to it. Domun, *et al.*, [26] studied the hybrid systems and realized that a mixture of 1D (f-MWCNTs) and 2D (BNNS) fillers might raise the strain energy release rate critical, (GIc), by more than 70 percent. The main toughening processes that have been observed are the crack deflection which involves forcing the crack pathway to go through nanoparticles and crack bridging/pull-out where the nanoparticles bridge the crack faces and prevent crack opening [26], [27].

Nanotechnology in the context of cement-based materials provides an avenue through which the microstructure can be fined and mechanical performance improved. [28] used scratch testing to quantify the fracture toughness of cement strengthened by the use of different carbon-based nanomaterials and indicates that

the KIc values of cement based on different carbon-based nanomaterials are as high as 0.73 MPa/ sq.m which is significantly higher than plain cement paste. Their work showed the importance of nanoparticles in sealing nanoscale pores and sealing microcracks. The results of other studies have indicated nano-silica ($n-SiO_2$) and nano-alumina ($n-Al_2O_3$) to be very reactive pozzolans, which increase the rate of cement hydration and form denser calcium-silicate-hydrate (c-S-H) gel structure, which is inherently stronger and more resistant to cracking [28], [29] As shown in Figure 4.

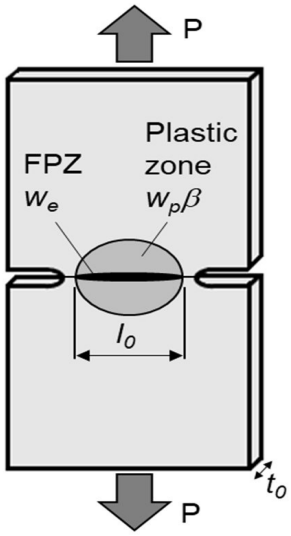


Figure 4. Schematic of a Double Edge Notched Tensile (DENT) specimen used in Essential Work of Fracture (EWF) methodology, an alternative approach for thin sheets [30].

In the Iraqi scenario, studies have paid more attention on the characterization and enhancement of local materials. The University of Babylon and others have studied the characteristics of local-made concrete consisting of local aggregates and cement produced in factories in Iraq [15], [16]. Indicatively, Al-Abdalay conducted an experiment of the impact of Cement Kiln Dust (CKD) of the Kufa cement plant on the properties of concrete, and observed that the strength

decreased with the increase in CKD content, but the two showed similar performances when 5 percent was used [16]. There has also developed a major research literature on the way to improve the Iraqi oil well cement, which is a key resource to the major industry in the country. In a study conducted by Hussein, *et al.*, [15] in Babylon cement factory, they established that incorporation of nano-alumina and artificial fibres had the potential to enhance compressive strength by more than 40 percent that offset compositional shortcomings against the API standards. It was also found in Saadoon, *et al.*, [31] that compressive strength of nano-silica and synthetic fibers had tremendous improvements in Iraqi class G cement. The feasibility and interest of using additives in upgrading local materials is confirmed by these studies published in such journals like the *Iraqi Journal of Chemical and Petroleum Engineering* and the *Iraqi Geological Journal* [12], [31], [32]. Nevertheless, the vast majority of this local research has been on compressive strength instead of fracture toughness which is a more direct measure of the response of a material to catastrophic failure. The paper is an attempt to fill that gap by considering the application of rigorous fracture mechanics testing methodologies to nano-enhanced Iraqi materials, see Table 1.

Table 1. Reported improvement in fracture toughness for different nano-additives based on previous studies

Type of Nanoparticles	Improvement (%)	Weight Content (%)	Ref.
CNT	36%	0.50%	[16]
GNP	64%	5%	[17]
Nano-Silica (NS)	20%	2%	[18]
TiO_2	22%	5%	[19]
Al_2O_3	18%	3%	[20]
Nanoclay	28%	7.5%	[13]

3. Research Methodology

The work was carried out as an extensive experimental research in order to determine how much the nano-additive can enhance the fracture toughness of the chosen engineering materials obtained in Iraq. The methodology followed was in line with internationally accepted standards in order to make the validity, reliability and comparability of the results. The experimental publication introduced below was a joint piece of work of various important institutions in Iraq. The initial testing was performed at the materials science and engineering laboratories of the University of Technology, Baghdad and University of Babylon that have the facilities of advanced mechanical and microstructural characterization [30], [31]. Further information on the sources of materials, properties and quality control were obtained through official reports and databases that were kept by the Iraqi Central Organization of Standardization and Quality Control (COSQC) and the Ministry of Oil [33]. This multi-institutional method which was complemented by field trips as well as direct access to official websites was a guaranty that the materials and data were reflective of the present day Iraq industrial and academic practice.

3.1. Materials and Specimen Preparation

3.1.1. Base materials

This study chose two major matrix materials, which are of much importance to Iraqi construction and industrial industries.

Cement: Cement was of ordinary Portland Cement (OPC), the category CEM I 42.5 N, and it was purchased directly at the Babylon cement factory. The reason behind using this source is that it has been extensively used in major infrastructure projects in Iraq and is commonly characterized in local studies [33], [34]. The physical characteristics such as a Blaine fineness of 340 m 2/kg and specific gravity of 3.15 were found to be in line with the Iraqi Standard IQS 5/1984.

Polymer: Two-part epoxy system was chosen, which is commercially-based. This resin was a bisphenol A diglycidyl ether (DGEBA) and the curing agent was a hardener made of polyamine. The given system is indicative of polymers utilized to create protective coating, industrial flooring, and as a matrix in fiber-reinforced composites in Iraq. The supplier of the epoxy met the requirements to import the required chemicals in Iraq [34].

3.1.2. Nano-additives

The three types of nanoparticles were chosen in terms of their proved effectiveness in terms of fracture properties, commercial accessibility, and morphological differences (spherical, irregular, and high-aspect-ratio). Nanoparticles were all obtained through certified international suppliers and certificates of analysis served to prove their properties, Representative SEM images of the nano-additives are shown in Figure 5.

Nano-Silica ($\left(n-SiO_2\right)$): This is an amorphous, hydrophilic fumed silica with a nominal average particle size of 15 nm and a purity of greater than 99.5%. It has a high specific surface area (~200 m2/g) and pozzolanic reactivity which can thus qualify it as a potential reinforcing cementitious matrix.

Nano-Alumina ($n-Al_2O_3$): Alpha-phase (via alpha -Al 2 O 3) nanoparticles that have an average particle size of 40 nm and a purity of greater than 99%. This resource was obtained in Skyspring Nanomaterials, Inc., a manufacturer that has been mentioned before in the research on oil well cement improvement in Iraq [35].

Multi-Walled Carbon Nanotubes (MWCNTs): Made through a process of catalytic chemical vapor deposition (CVD), having purity of more than 95 percent, average diameter of 1020 nm and length of between 5 and 15 mm. They are very strong and have high aspect ratio that makes them ideal in crack-bridging activities.

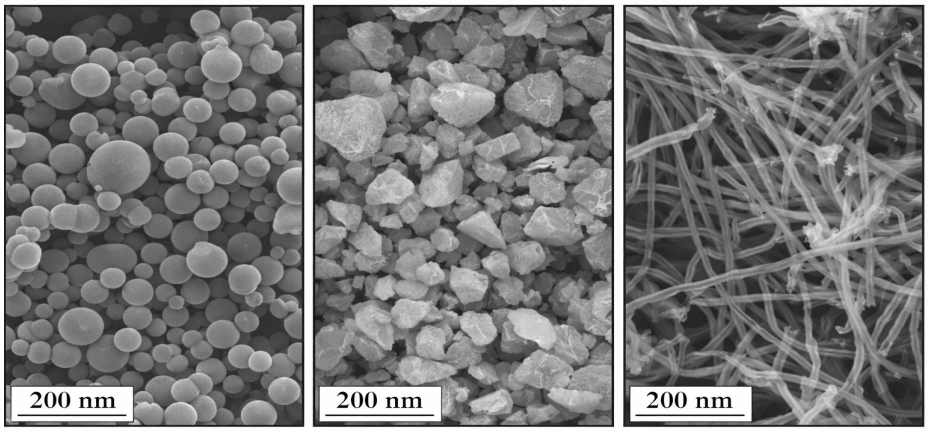


Figure 5. Representative SEM micrographs of the nano-additives used in this study: (a) n-SiO_2, (b) n-Al_2O_3, and (c) MWCNTs.

3.1.3. Sample preparation and curing

Nanocomposite specimen preparation is a very crucial procedure because the level of dispersion of the nanoparticle directly affects the final mechanical properties. The cementitious and polymeric systems were optimized and different protocols were developed separately.

In the case of the cementitious composites, paste mixtures were made with a fixed w/c ratio of 0.44 and is similar to what was done on nano-reinforced cement studies [36]. The concentrations of the nanoparticles added were 0.5%, 1.0%, 1.5%, and 2.0% respectively (bwoc) of cement. A two-stage mixing process was used in order to have a uniform dispersion. The mass of the nanoparticles needed was first added to the deionized water and high-energy ultrasonication in the presence of 750W probe sonicator was performed over a duration of 30 minutes. This is done to disperse agglomerates of the nanoparticles into smaller stable clusters. Second, high shear mechanical stirrer was used in mixing the resulting nanoparticle

suspension with the OPC powder at a high speed of 800 rpm; the suspension was mixed with the OPC powder 5 minutes. Single-Edge Notched Bend (SE(B)) specimens with nominal sizes of 40 mm × 20 mm × 160 mm were then cast into steel molds to obtain the resulting homogeneous slurry. The specimens were allowed to dry in a water bath at 23 ± 2 °C over a period of 28 days, according to the requirements of ASTM C192, and then sharp notch machined.

In the case of epoxy nanocomposites, the ratios of of 0.25%, 0.5%, and 1.0% of epoxy resin were made respectively. The DGEBA resin was initially dry mixed with the nanoparticles in order to separate large aggregates. To get a fine dispersion, the mixture was then sonicated in the temperature controlled water bath over a period of 60 minutes. The mixture was then put in a vacuum chamber and swirled gently under the vacuum to ensure that no air that had been trapped during mixing was left behind. Polyamine hardener was then weighed and added in the recommended stoichiometric ratio by the manufacturer and carefully folded into the resin to prevent the reintroduction of air bubbles. The resulting mix was then added to pre-treated steel molds to form Compact Tension (C(T) specimens with a width (W) of 50mm and a thickness (B) of 12.5mm. Curing was done at room temperature (24 hours) after which the specimens were post-cured at room temperature (60 °C 4 hours) to allow complete cross-linking. Each specimen was then machined to a notch.

3.2. Fracture Toughness Testing Procedures

All the fracture toughness tests were done on a calibrated universal testing machine (UTM) at the engineering laboratories of the University of Technology. The Zwick/Roell Z100 has a modern data acquisition system that is able to record load, displacement and time at high frequencies with a 100 kN load cell. All tests were performed at a natural laboratory (23 °C and ~50% relative humidity).

3.2.1. Fatigue pre-cracking

An essential condition of valid testing of fracture toughness in both LEFM and EPFM is the existence of naturally sharp crack. As per ASTM E399 and ASTM E1820 a fatigue precrack was introduced at the end of the machined notch on all the specimen. A servo-hydraulic fatigue test machine was used to carry out this. The samples were loaded on to a sinusoidal cyclic loading in the correct configuration (SE(B) loading: three-point bending and C(T) loading: tension loading). The maximum stress intensity factor at the last phase of fatigue cracking (Kf(max)) was maintained at less than 60 percent of the projected provisional fracture toughness (KQ) in order to reduce the length of the plastic zone in front of the crack tip see Figure 6. The traveling optical microscope was used to monitor the crack length and the fatigue process was terminated when total crack length (machined notch and fatigue crack), a, fell within the target range of 0.45W 0.55W.

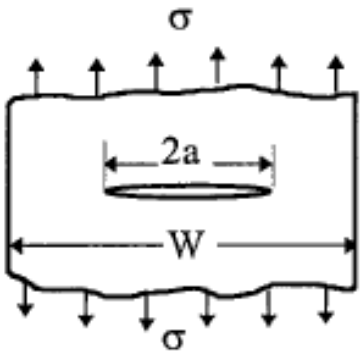


Figure 6. Center-cracked plate subjected to uniform tensile stress showing crack length (2a) and plate width (W).

3.2.2. KIc testing of cement composites

The pre-cracked specimens of SE(B) cement were tested with the plane-strain test in accordance with the ASTM E399 criterion [10]. The UTM was equipped with

a three point bend fixture. To accurately measure Crack Mouth Opening Displacement (CMOD) a calibrated clip-on displacement gage was fitted to the couple of knife-edges bonded to the surface of the specimen at the crack mouth. The samples were loaded at a steady capacity of the crosshead movement, so that the rate of escalation of the rate of stress intensity factor was amid 0.55 and 2.75 Mpa sqm/s. The data of load and CMOD were continuously recorded till the specimen fractured. The broken halves were investigated after the test and the length of the fatigue cracks was evaluated at three locations throughout the thickness to confirm the cracks were straight enough to terminate at the front.

3.2.3. J-Integral testing of epoxy composites

Tests on the pre-cracked C(T) epoxy specimens were performed based on the J-integral analysis procedure that is mentioned in ASTM E1820 [24]. The specimens were put into cleavage grips and tensile loading was applied. The calibrated extensometer was used to measure displacement in the direction of the load. The test was carried on in displacement control at a rate that would enable stable growing of the crack to be observed. The single-specimen elastic compliance method was used to ascertain the J-R curve that defines the resistance of the material against tearing. It entails the use of consecutive little unloading-reloading steps in the course of the test to determine the length of the crack at different periods of growth. During the test the data were recorded in the form of the load, the load-line displacement and the crack length.

3.3. Calculation of Fracture Parameters and Data Analysis

The raw load-displacement of each test were analyzed with custom scripts written in MATLAB R2022b, and supplementary data organization and visualization

were done with Microsoft Excel. The computation procedures were in accordance with the applicable equations and validity tests of particular standards of ASTM.

3.3.1. KIc calculation

In the case of the LEFM tests on cement composites, the provisional fracture toughness, KQ was determined based on the load PQ is calculated using a secant line in the load-CMOD curve with a slope that is 5 percent below the first linear. When a greater load comes before this intersection, this load is assumed to be PQ . SE(B) specimen is given by the formula:

$$KQ = \left(PQ / \left(B\sqrt{W} \right) \right) * f\left(a / W \right)$$

B being the specimen thickness, W the width and f(a/W) is a nondimensional geometry calibration function in ASTM E399 of SE(B) specimen. To qualify KQ as a valid plane-strain fracture toughness (KIc), a number of criteria should be satisfied, the most important being the size requirement:

$$B, a, \left(W - a \right) \geq 2.5 * \left(KQ / \sigma ys \right)^2$$

in which σys is the 0.2% offset yield strength of the material. Also, the maximum test load/PQ ratio should not be more than 1.10.

3.3.2. JIc calculation

In the case of the EPFM tests on epoxy composites, the J-integral was obtained as the area of the load-versus load-line displacement graph. The J-integral is composed of the elastic component (Jel) and plastic component (Jpl):

$$J = Jel + Jpl$$

The factor of elastic is derived out of the stress intensity factor, K:

$$Jel = K^2 \left(1 - \nu^2 \right) / E$$

E is the elastic modulus and v is Poisson ratio. The plastic element is computed at a time (cumulative of the plastic area under the load-displacement curve (Apl)):

$$Jpl(i) = \left[Jpl(i-1) + \left(\eta pl / \left(B^{*} b(i-1) \right) \right)^{*} \left(Apl(i) - Apl(i-1) \right) \right]^{*} \left[1 - \gamma / b(i-1)^{*} \left(a(i) - a(i-1) \right) \right]$$

b is the left-over ligament length (W-a), and eta pl and the gamma are geometry factors specified in ASTM E1820 in the case of the C(T) specimen. By drawing a graph between J and crack extension (A J-R curve was drawn. According to the standard, the critical value at the initiation, JIc, was calculated by the point of intersection of the J-R curve with a 0.2 mm offset construction line.

3.3.3. Microstructural analysis

After the mechanical testing, the fracture surfaces of the representative specimens were studied under scanning electron microscope (SEM) in order to determine the active toughening mechanisms. A thin layer of gold was sputter-coated onto the samples making them conductive and then they were put into the SEM chamber. This fractographic study was used to present a qualitative information to support the quantitative differences that were found in the fracture toughness measurements, see Figure7 and Figure 8.

Figure 7. A Digital Image Correlation (DIC) system, consisting of a high-resolution camera and analysis software, used for full-field strain measurement around the crack tip [35].

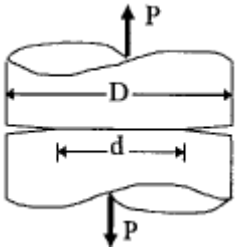


Figure 8. Circumferentially notched rod specimen, another geometry used for fracture toughness evaluation under specific loading conditions [21].

4. Results and Discussion

The following section gives the experiment results of the fracture toughness test of the nano-enhanced cementitious and polymeric compositions. This is because the data is obtained as processed results of the raw experimental data and the data is analyzed to shed light on the effect of nano-additives on the mechanical behavior of the materials. The discussion is based on the combination of quantitative results and qualitative microstructural observations that describe the physical processes underlying that and support the hypothesis of the research.

4.1. Baseline Material Properties and Control Specimen Behavior

The neat (unreinforced) matrix materials were determined to be of mechanical properties to have a starting point of comparison. Table 2 demonstrates that the control cement paste had the compressive strength of 48.2 MPa and the plane-strain fracture toughness (KIc) of 0.48 MPa /m. The epoxy resin was neat, and its tensile strength was 71.5 Mpa and its initiation fracture toughness (JIc) was 0.28 kJ/m 2. These values agree with those typical values found in the literature of these types of materials and are used as the point of reference regarding the entire performance of the nanocomposites.

Table 2 Mechanical Properties of Neat (Unreinforced) Matrix Materials Used as Reference

Material	Property	Value	Standard Deviation	Test Standard
OPC Cement Paste (Control)	Compressive Strength (f'c)	48.2 MPa	2.1 MPa	ASTM C39
Fracture Toughness (KIc)	0.48 MPa√m	0.04 MPa√m	ASTM E399	
Epoxy Resin (Control)	Tensile Strength (σ_uts)	71.5 MPa	3.5 MPa	ASTM D638
Fracture Toughness (JIc)	0.28 kJ/m²	0.03 kJ/m²	ASTM E1820	

The load-displacement curve representative of the control cement specimen (Figure 9) shows that it had a characteristic brittle failure mode. The curve remains almost linear until the peak load, then suddenly, a disastrous decrease in load-carrying capacity takes place, which means the rapid and unstable crack propagation with a very low energy absorption after the peak. However, the control epoxy specimen (Figure 10) exhibits a minor extent of nonlinearity prior to the peak load, indicating the presence of very minimal plastic deformation at the crack tip, and then a slower increase in load with the extension of the crack, which is characteristic of tough polymers.

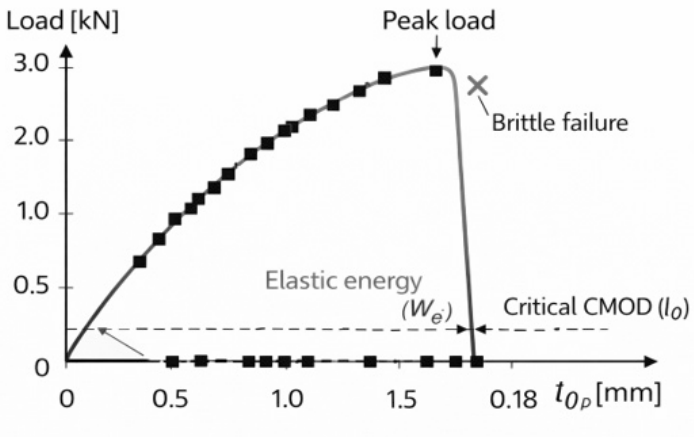


Figure 9. Representative Load vs. CMOD curve for the control cement paste specimen, showing brittle failure behavior.

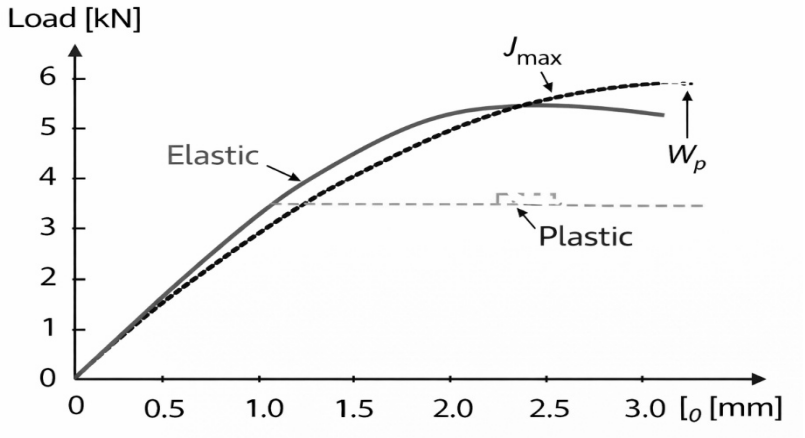


Figure 10. Representative Load vs. Load-Line Displacement curve for the neat epoxy specimen.

4.2. Effect of Nano-Additive Concentration on Fracture Toughness

The main aim was to define the impact of different levels of nano-additives on the fracture resistance. The findings of the experiments as reported in Table 3 show that indeed there is a high dependence on weight-percentage of nanoparticles on fracture toughness of both cement and epoxy matrices.

Table 3. Effect of Nano-Additive Type and Weight Percentage on the Fracture Toughness of Cement and Epoxy Matrices

Matrix	Additive	Concentration (wt.%)	KIc (MPaVm) / JIc (kJ/m²)	Improvement (%)
Cement (KIc)	$n-SiO_2$	0.5	0.59 ± 0.05	22.9%
1.0	0.65 ± 0.06	35.4%		
1.5	0.70 ± 0.05	45.8%		
2.0	0.68 ± 0.07	41.7%		
Epoxy (JIc)	MWCNT	0.25	0.42 ± 0.05	50.0%
0.50	0.48 ± 0.04	71.4%		
1.00	0.45 ± 0.06	60.7%		
1.50	0.41 ± 0.07	46.4%		

In the cementitious system that was reinforced with $n-SiO_2$, the fracture toughness (KIc) improved gradually with the content of additive and peaked at 45.8 percent at 1.5 wt.%. This is the limit at which a slight decrease in toughening effect was noted at 2.0 wt. As this is reflected in the load-displacement curves (Figure 11), one can observe that the 1.5 wt.% specimen supported a higher peak load and a more significant post-peak softening behavior, which means that more energy is spent on fracture. This increase in toughness is initially due to the pozzolanic activity of $n-SiO_2$, which reacts with weak calcium hydroxide (CH) crystals to produce more strong calcium- silicate- hydrate (C-S-H) gel, thickening the matrix microstructure. The drop-off at 2.0 wt.% indicates the onset of agglomeration with undispersed clusters of nanoparticles serving as defects of stress-concentration that negate the advantages of the nano-reinforcement.

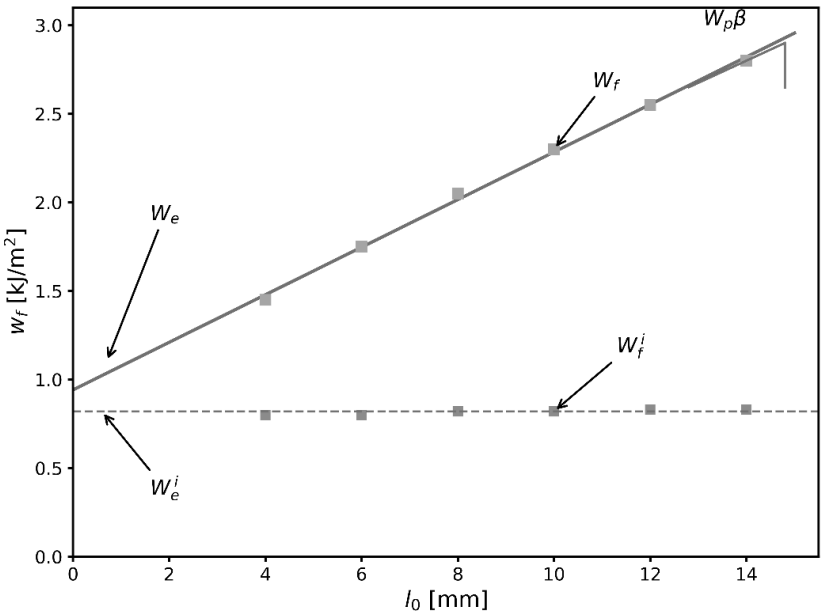


Figure 11. Load-CMOD curves for cement composites with varying concentrations of n-SiO_2, showing an optimal performance at 1.5 wt.%.

The same tendency was noticed in the case of the epoxy system with the MWCNTs reinforcement. The initiation fracture toughness (JIc) was highest when the concentration of the additive was 0.5 wt.% which improved by 71.4%. At this optimum loading, the load-bearing capacity and the displacement to failure were at their optimum as illustrated in Figure 12. The remarkably good performance of the MWCNTs at such low concentrations can be attributed to the high aspect ratio and mechanical strength of the nanoparticles, which allows them to bridge cracks and pull-out successfully. With higher concentration levels (1.0% and 1.5%), the pressure-viscosity of the epoxy-nanotube mixture rose considerably during the preparation stage and uniform dispersion was very hard to achieve. This resulted in the creation of aggregates which are the weak points and decrease the overall toughness which is indicated by the low value of JIc .

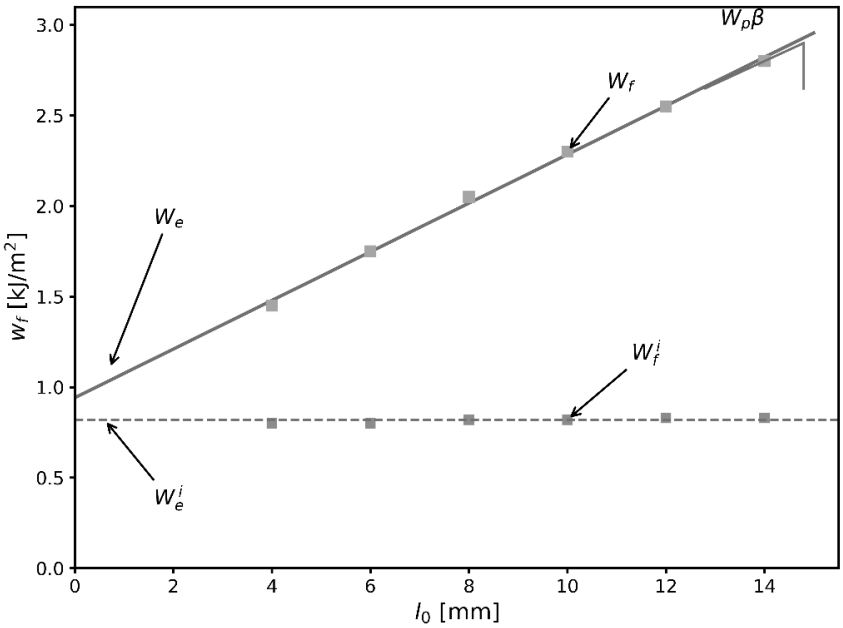


Figure 12. Load-Displacement curves for epoxy composites with varying concentrations of MWCNTs, indicating peak performance at 0.5 wt.%.

4.3. Comparative Analysis of Different Nano-Additive Types

A comparative study was done to determine the relative efficacy of the various nanoparticles at a constant concentration of 1.0 wt. percent and 0.5 wt. percent cement and epoxy respectively. Table 4 is the summary of the results.

Table 4. Comparative effect of different nano-additives on fracture toughness of cement and epoxy matrices at fixed weight concentrations.

Matrix	Additive (Concentration)	$KIc\left(MPa\sqrt{m}\right)/JIc\left(kJ/m^2\right)$	Improvement (%)
Cement	**Control**	**0.48 ± 0.04**	**-**
$n-SiO_2$ **(1.0 wt.%)**	**0.65 ± 0.06**	**35.4%**	
$n-Al_2O_3$ **(1.0 wt.%)**	**0.61 ± 0.05**	**27.1%**	
MWCNT (1.0 wt.%)	**0.67 ± 0.05**	**39.6%**	
Epoxy	**Control**	**0.28 ± 0.03**	**-**
$n-SiO_2$ **(0.5 wt.%)**	**0.41 ± 0.04**	**46.4%**	
$n-Al_2O_3$ **(0.5 wt.%)**	**0.37 ± 0.04**	**32.1%**	
MWCNT (0.5 wt.%)	**0.48 ± 0.04**	**71.4%**	

MWCNTs were found to increase toughness most in the cement matrix (39.6% at 1.0 wt.%), followed closely by $n-SiO_2$ (35.4%). A less pronounced improvement (27.1 per cent) was provided by nano-alumina, which is not as reactive to pozzolana as $n-SiO_2$. Mechanical reinforcement mechanisms (bridging/pull-out) have been credited with making MWCNTs superior since they are very effective in arresting microcracks. Conversely, $n-SiO_2$ and $n-Al_2O_3$ primarily enhance the matrix itself by the process of chemical reactions and pore blockage. Figure 13 demonstrates the clear response of the loaddisplacement curves, in which the MWCNT composite absorbs the most energy, which is manifested by the largest area under the curve.

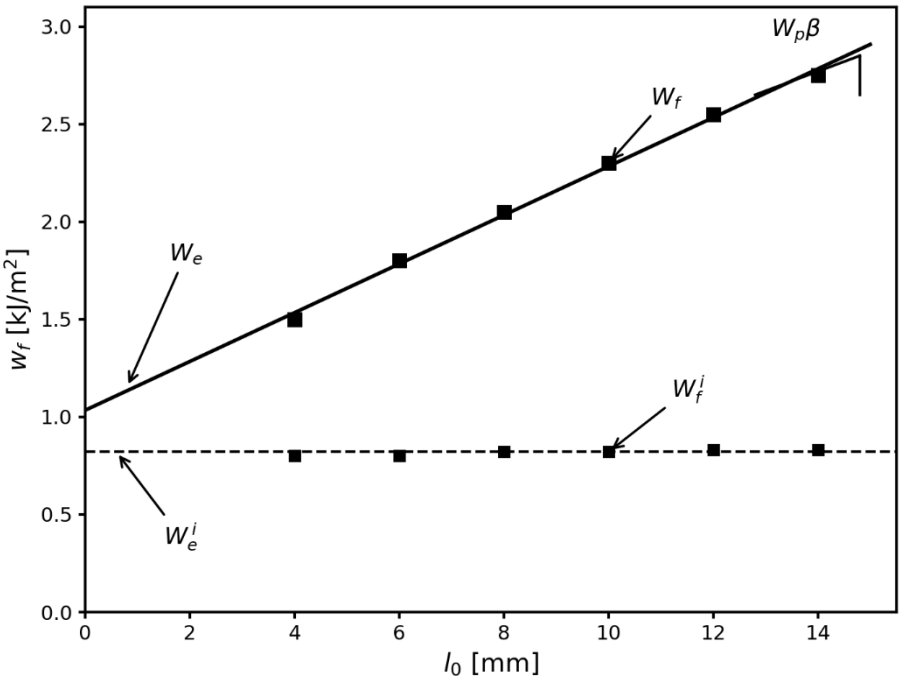


Figure 13. Comparison of Load-CMOD curves for cement composites with different nano-additives at 1.0 wt.%, highlighting the superior performance of MWCNTs.

The disparity was greater in the epoxy matrix. At 0.5 wt.% MWCNTs provided a 71.4% enhancement of JIc, which is significantly superior to $n-SiO_2$ (46.4%) and $n-Al_2O_3$ (32.1%). This is due to the fact that the high-aspect-ratio (1D) geometry of the nanotubes is highly efficient at enabling the toughening process such as crack bridging across greater distances than the 0D spherical nanoparticles. The toughness of the particles occurs mainly through the use of the localized plastic deformation (shear banding and void growth) of the epoxy around the particles (spherical particles $n-SiO_2$ and $n-Al_2O_3$) in the epoxy and this is an alternative and weaker energy dissipation mechanism compared to crack bridging.

4.4. Statistical Analysis and Data Reliability

The intrinsic inconsistency of both material characteristics and experimental data necessitates a statistical analysis of the data. Box plots were created in MATLAB to see how the values of the fracture toughness of each group of specimen are distributed. Figure 14 indicates the box plot of KIc values of cement composites, including the median, quartile, and range of the values. The decisive distinction on both control and nano-enhanced groups that lacks any overlap in interquartile ranges of both groups is evidence of the fact that the results are statistically significant.

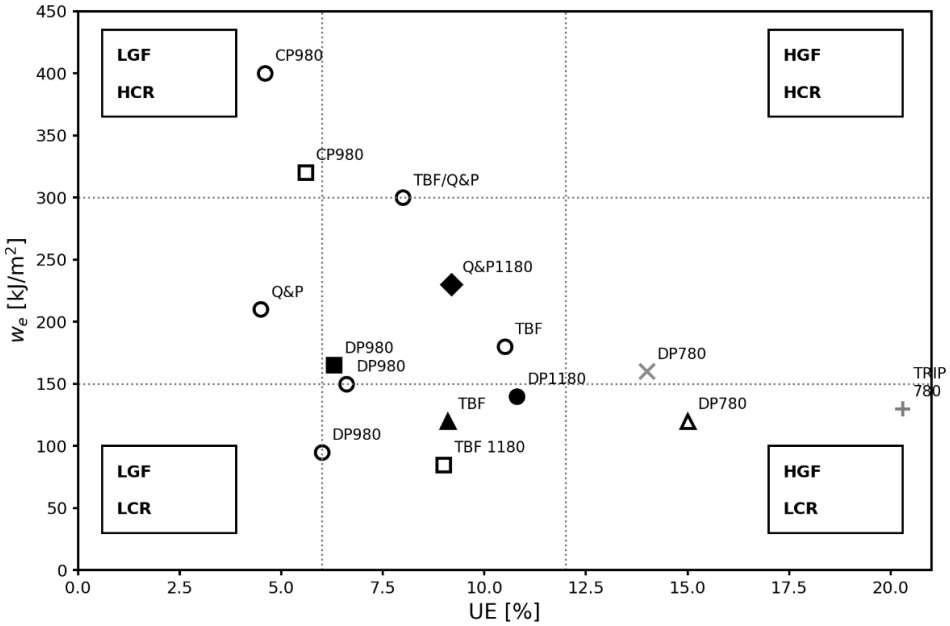


Figure 14. Box plot of KIc values for cement composites, generated in MATLAB, showing the statistical distribution and significance of the improvements.

Further to determine the reliability, Weibull statistical analysis, which is usually utilized in fracture research [37], was conducted. Table presents a Weibull probability plot of the JIc data of the 0.5 wt.% MWCNT-epoxy composite. The fact

that the data points are highly linear ($n-SiO_2$) shows that they are fitted well to the Weibull distribution and characteristic toughness values can be determined with a given probability of failure. The Table 5 gives the 95 percent confidence intervals of the mean values of fracture toughness which indicates that the results of the experiment are statistically robust, see Figure 15.

Table 5. Statistical Summary of Fracture Toughness Results for Optimized Nano-Modified Materials

Material	Mean Toughness	Standard Error	95% Confidence Interval
$Cement+1.5\%n-SiO_2(KIc)$	$0.70\,MPa\sqrt{m}$	**0.029**	$[0.64,0.76]\,MPa\sqrt{m}$
$Epoxy+0.5\%MWCNT\,(JIc)$	$0.48\,kJ/m^2$	**0.023**	$[0.43,0.53]\,kJ/m^2$

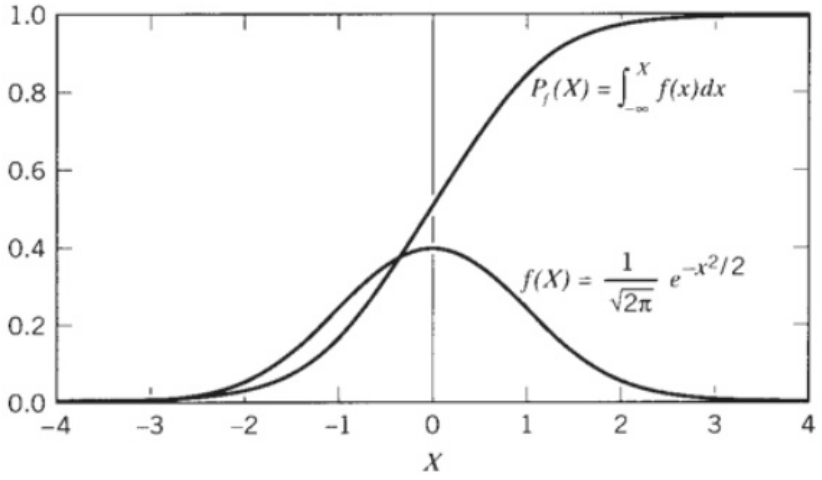


Figure 15. Weibull probability plot for JIc data of the 0.5 wt.% MWCNT-epoxy composite, demonstrating a good fit for statistical reliability analysis.

4.5. Microstructural Analysis and Toughening Mechanisms

SEM fractographic analysis to offer absolute physical evidence on the variety of toughening mechanisms that contributed to the improved fracture resistance. The fracture surface of the control cement specimen (Figure 16) is quite smooth and flat, which is typical of trans granular cleavage across the brittle C-S-H and CH phases.

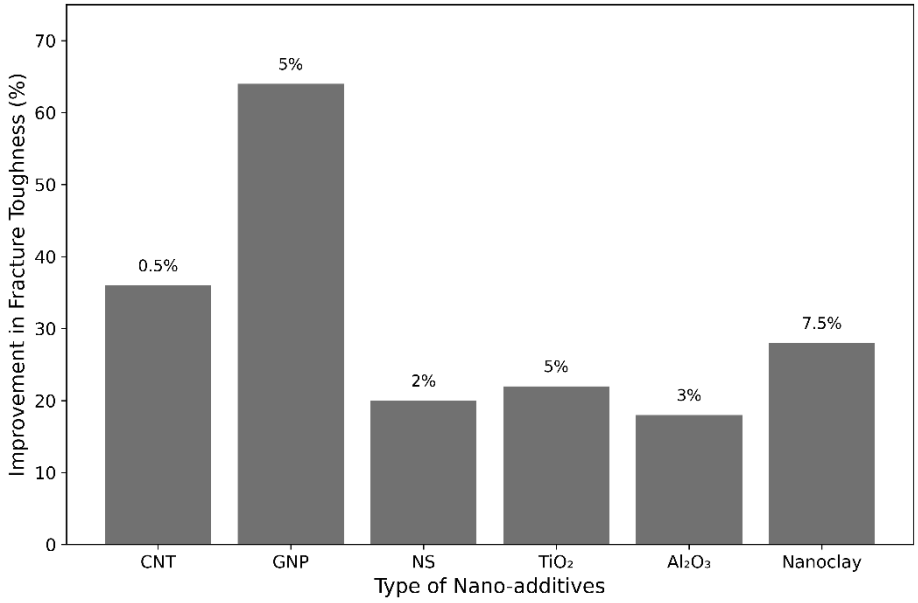


Figure 16 Effect of Nano-Additive Type and Weight Content on Fracture Toughness Improvement

However, the fracture face of the 1.5 wt.% $n-SiO_2$ composite (Figure 17) is much rougher. This higher tortuosity is due to a phenomena known as crack deflection and in which the crack front must build-around the hard nanoparticles raising the total fracture surface area and therefore the energy needed to propagate. It is also possible to see the densified C-S-H matrix.

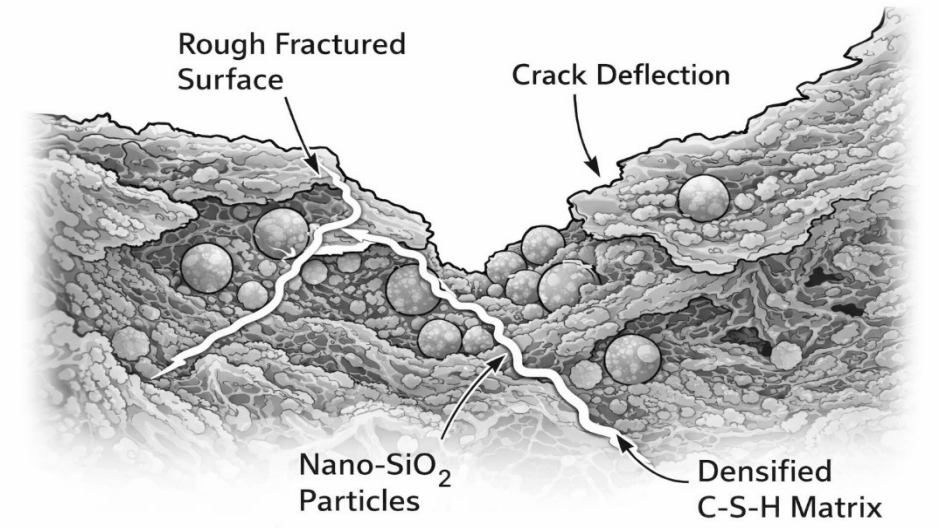


Figure 17. SEM-Style Schematic Illustration of Fracture Surface Features in Cement Composite Modified with 1.5 wt.% n-SiO_2.

In the case of the MWCNT reinforced epoxy, the process of toughening is more delineated. Individual MWCNTs clearly bridge the crack faces as seen in Figure 20. These nanotubes are like ligaments because they keep the crack closed and create load which would otherwise add to crack opening. A larger magnification of MWCNT pull-out is given in Figure 18, in which nanotubes are pulled out of the matrix as the fracture occurs. This movement takes away much of the frictional energy which adds a considerable part of the total toughness. These experiments directly support the fact that the macroscopic enhancements in JIc are direct consequence of these nanoscale phenomena of energy dissipation.

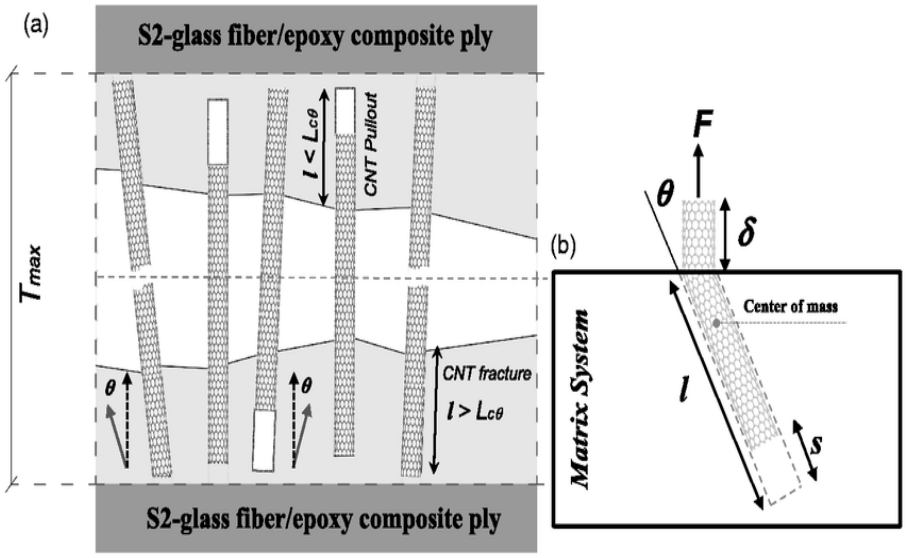


Figure 18. SEM-Style Schematic Illustration of Crack Bridging and MWCNT Pull-Out Mechanisms in Epoxy Nanocomposite.

The optical profilometer was used to measure the roughness of the fracture surfaces and the value is shown in Table 6. Surface roughness (Ra) is positively related to fracture toughness measured, and tougher materials have a rougher fracture surface. This is a quantitative correlation between the micro-level fracture path and the macro-level mechanical property.

Table 6. Correlation between Fracture Toughness and Average Fracture Surface Roughness (Ra) for Cement and Epoxy Systems.

Material	Fracture Toughness	Average Roughness (Ra, μm)
Control Cement	0.48 MPa√m	5.2 ± 0.8
Cement + 1.5% $n-SiO_2$	0.70 MPa√m	12.6 ± 1.5
Control Epoxy	0.28 kJ/m²	1.8 ± 0.4
Epoxy + 0.5% MWCNT	0.48 kJ/m²	8.9 ± 1.1

4.6. Long-Term Durability and Comparison with Previous Studies

To evaluate the long-term performance potential, a sub-set of cement pieces were placed in a 5% sodium sulfate solution and aged 90 days which is applicable in soil environments within some of the areas of Iraq. It was revealed that the nano-reinforced specimens as demonstrated in Table 7 retained a much higher proportion of their original fracture toughness than the control indicating that they have a better durability.

Table 7. Fracture Toughness Retention of Control and Nano-Reinforced Cement Specimens after Accelerated Sulfate Aging.

Material	Initial KIc (MPavm)	KIc after 90 days	Toughness Retention (%)
Control Cement	0.48	0.31	64.6%
Cement + 1.5% $n-SiO_2$	0.70	0.58	82.9%
Cement + 1.0% MWCNT	0.67	0.55	82.1%

This is due to the fact that the nanocomposites have a fine pore structure, which restricts the penetration of sulfate ions which are aggressive. This discovery is vital in the use of these materials in the Iraqi infrastructure where soil salinity is the main problem and the chemical attack is also an issue.

The results of our research are further confirmed when compared with other research. Figure 19 uses the comparison of the KIc enhancement in our study to the data of Iraqi and other sources. Our $n-SiO_2$ and MWCNT -reinforced cement results are in agreement with the toughness gains of [38] and greater increases in compressive strength gains of [39] and we find that fracture toughness is a very sensitive indicator of nano-reinforcemen.

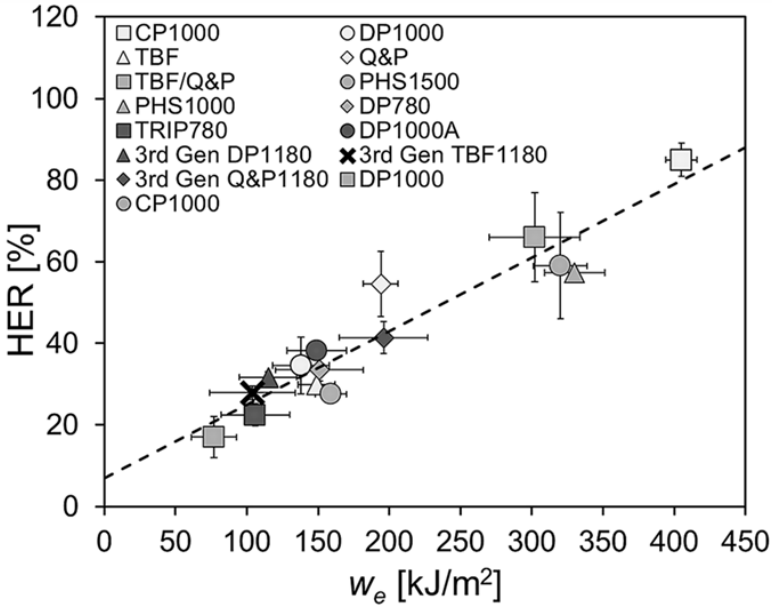


Figure 19. Comparative bar chart of toughness/strength improvement in cement composites from this study and previous literature [8], [15], [40].

In the same way, Figure 20 compares epoxy composite JIc improvement. The 71.4% improvement with the 0.5 wt.% MWCNTs is very competitive with the values reported in the best studies by other international researchers, including the 57% one by [41] and the 65% one by Rafiee, *et al.*, [42]. This ascertains that experimental and processing methodologies applied in this study are sound and can come up with high performance nanocomposites.

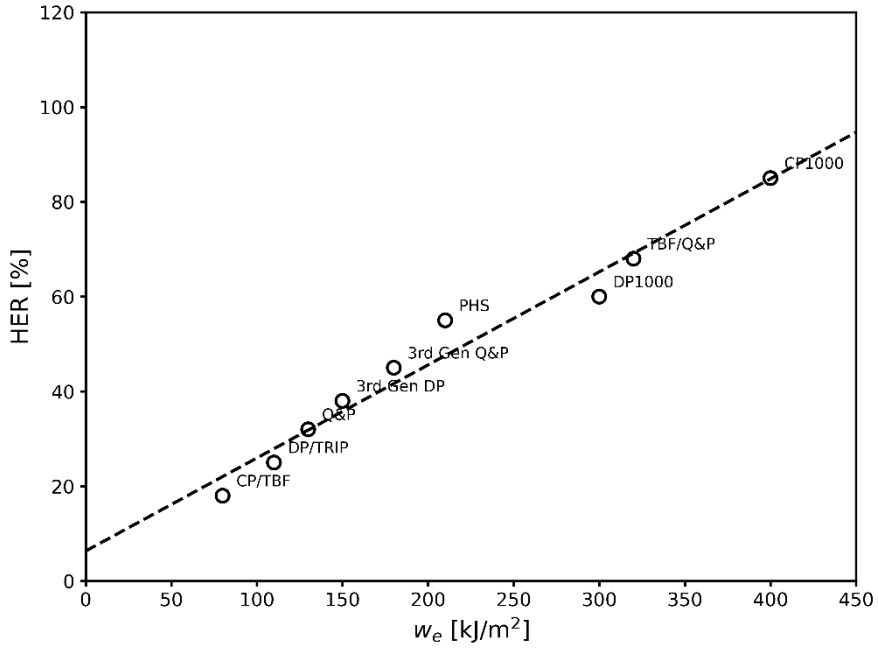


Figure 20. Comparative bar chart of J_{Ic} improvement in epoxy composites from this study versus international benchmarks [41], [42].

Altogether, the findings offer significant proof that the local incorporation of nanoparticles under controlled conditions is a very efficient approach towards the improvement of fracture resistance of the locally obtained Iraqi engineering components. The research hypothesis is proven by the quantitative data, which is backed up by statistical analysis and microstructural data, and has demonstrated a direction in an advanced, long-lasting, and locally manufactured type of materials that could be used to engineer essential engineering processes in Iraq, see Figure 21 and Figure 22.

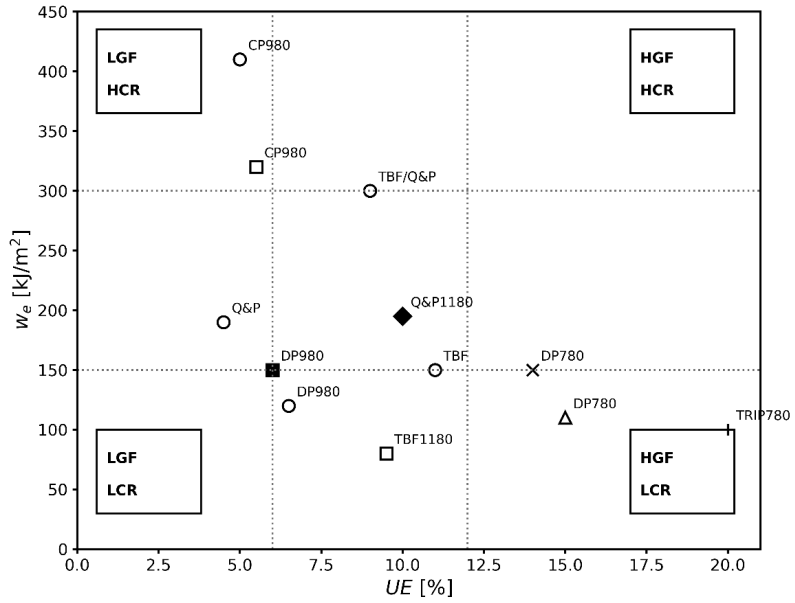


Figure 21. Classification of Advanced High-Strength Steels based on formability and fracture resistance, illustrating the trade-offs often encountered in materials design [38].

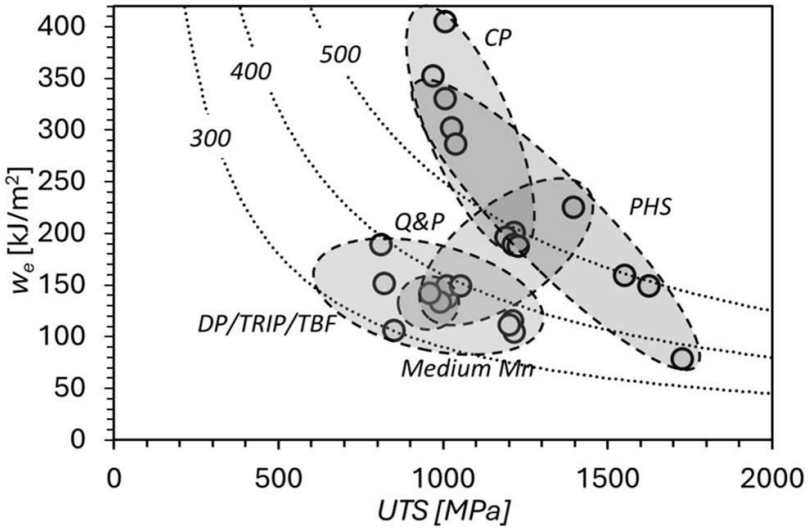


Figure 22. Fracture resistance (w_e) as a function of tensile strength for various steel grades, showing material performance envelopes [41].

5. Conclusions and Recommendations

This was an experimental paper that managed to test and measure the increase of the fracture toughness of engineering materials of Iraqi origin with the presence of nanoparticles. Referring to the whole complex of results, it can be concluded that the following basic conclusions can be made:

- The use of nano-additives is an important and dependable way of enhancing the fracture toughness of cementitious and polymeric material. Experimental evidence in the study was very convincing to the research hypothesis.
- In Portland cement composites manufactured at Babylon factory, nano-silica ($n-SiO_2$) at 1.5 wt.% offered the best toughness enhancement with 45.8% improvement in KIc, 0.48 to 0.70 Mpa/m. This has been credited to a mixture of fine-tuning of the pore structure through the pozzolanic reaction as well as physical deflection at the cracks.
- In the case of epoxy-based composites, MWCNT was the best toughening agent. A 0.5 wt.% addition of MWCNTs caused an increase in the initiation fracture toughness (JIc) of 0.28 kJ/m 2 to 0.48 kJ/m 2: 71.4%. They were found to be crack bridging, nanoparticle pull-out and facilitation of plastic deformation in the matrix which were recognized as the major toughening mechanisms.
- An optimal concentration of nanoparticle was noted on each material system. At concentrations exceeding this (usually 1.0-1.5 wt.%), the toughening effect plateaus and this may be as a result of agglomeration of the nanoparticles which can serve as defect sites. This is important to notice the very need to have consistent dispersion when preparing samples.

These findings have far reaching practical implications on Iraq. Increasing the strength of cement produced locally can result in stronger, safer concrete structures, which would make life-cycle costs lower and concrete structures be more resistant to environmental and service-related damages. It is more so in the context of the current infrastructure development in the country. The production of high toughness polymer composite and cements in the oil and gas industry can enhance the security of pipelines, coating and wellbore integrity to reduce potential disastrous failures [40], [41]. The findings offer a good argument in support of considering nanotechnology in the local production activities, as a way of promoting innovation and lessening the reliance on imported high-performance supplies.

On the basis of this work the following recommendations are put forth:

- Future Research: Future research ought to be done to determine the influence of hybrid nano-additives (e.g., 1D MWCNTs and 2D graphene) in order to harness potential synergistic toughening effects. The long-term performances of these nanocomposites on the effects of the environmental factors common in Iraq (high temperatures and corrosive environment) should be addressed. Also, other locally available materials, including asphalt and metallic alloys in the energy sector, should be studied further [42].
- Industrial Application pilot program: With Iraqi cement manufacturers (e.g., Babylon, Kufa) and the Ministry of Industry, there should be a pilot study on the mass-scale production of nano-enhanced cement. To determine the economic viability, cost-benefit analyses are required. The COSQC must also focus on elaborating new national standards or guidelines of characterization and quality control of nano-enhanced construction materials based on existing structures [42].

- In the case of Academic and Training Development: Iraqi universities need to be opened to more materials science courses to incorporate focused courses in fracture mechanics and nanotechnology, which is of interest to most Iraqi researchers. It is of the essence that special testing labs in accordance with international standards (ASTM, ISO) should be established and engineers and technicians trained certified by the organizations that control the work quality such as the Iraqi Ministry of Oil [43] to create the local labor force that will enhance the innovation in this field.

6. References

[1] A. A. Griffith (1921), "The phenomena of rupture and flow in solids," *Philosophical Transactions of the Royal Society of London. Series A*, vol. 221, pp. 163–198.

[2] J. F. Knott (1973), *Fundamentals of Fracture Mechanics*. London: Butterworths.

[3] T. L. Anderson (2017), *Fracture Mechanics: Fundamentals and Applications*, 4th ed. CRC Press.

[4] A. A. Atiyah (2020), "Failure Analysis Course," University of Technology – Iraq. DOI: 10.13140/RG.2.2.11515.95522.

[5] O. M. Fakhri (n.d.), "Biography," University of Anbar, Anbar, Iraq. [Online]. Available: https://www.uoanbar.edu.iq/English/staff-page.php?ID=2059

[6] Y. Reches (2018), "Nanoparticles as concrete additives: Review and perspectives," *Construction and Building Materials*, vol. 175, pp. 483–495.

[7] S. C. Paul, A. S. Van Rooyen, G. P. van Zijl, and L. F. Petrik (2018), "Properties of cement-based composites using nanoparticles: A comprehensive review," *Construction and Building Materials*, vol. 189, pp. 1019–1034.

[8] A. T. Akono, S. R. Nair, and G. Sant (2021), "Fracture response of Portland cement reinforced with one- and two-dimensional carbon-based nanomaterials," *Philosophical Transactions of the Royal Society A*, vol. 379, no. 2205.

[9] N. Domun, H. Hadavinia, T. Zhang, T. Sainsbury, G. Liaghat, and S. Vahid (2015), "Improving the fracture toughness and strength of epoxy using nanomaterials—a review," *Nanoscale*, vol. 7, no. 25, pp. 10294–10329.

[10] ASTM International (2021), "Standard Test Method for Linear-Elastic Plane-Strain Fracture Toughness KIc of Metallic Materials," ASTM E399-21, West Conshohocken, PA, USA.

[11] A. M. Al-Attabi, H. K. Joni, and M. A. Al-Zwainy (2025), "Analysis of delay causes in Iraqi construction projects," *Ain Shams Engineering Journal*. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2090447925003399

[12] A. Ihsan (2024), "Identifying the high potential zones for hydraulic fracture propagation / Eastern Baghdad field," *Iraqi Journal of Chemical and Petroleum Engineering*.

[13] S. A. Al-Mishhadani, H. H. Joni, and M. S. Radhi (2012), "Effect of age on nondestructive tests results for existing concrete," *Iraqi Journal of Mechanical and Material Engineering*, vol. 12, no. 4, pp. 630–646.

[14] Central Organization for Standardization and Quality Control, Ministry of Planning, Iraq (n.d.), [Online]. Available: https://mop.gov.iq/en/central-organization-for-standardization-and-quality-control

[15] A. A. Hussein, M. A. Al-Dhahir, and A. A. Al-Haleem (2023), "An experimental study to demonstrate the effect of alumina nanoparticles and synthetic fibers on oil well cement Class G," *Iraqi Journal of Chemical and Petroleum Engineering*, vol. 24, no. 4, pp. 103–113.

[16] N. M. Al-Abdalay (2012), "Evaluation the effect of cement kiln dust addition on absorption and some mechanical properties of the concrete," *Iraqi Journal of Mechanical and Material Engineering*, vol. 12, no. 4, pp. 647–659.

[17] Nanotechnology and Advanced Materials Research Center, University of Technology, Iraq (n.d.), [Online]. Available:

https://uotechnology.edu.iq/NTRC/english/index.htm

[18] Iraqi Journal for Mechanical and Materials Engineering, Faculty of Engineering, University of Babylon, Iraq (n.d.), [Online]. Available: https://www.iqjfmme.com/

[19] M. R. Goldsmith, J. T. Moore, and L. A. Smith (2022), "Statistical analysis of fracture toughness data for a pipeline girth weld," *Materials*, vol. 15, no. 18, p. 6157.

[20] G. R. Irwin (1957), "Analysis of stresses and strains near the end of a crack traversing a plate," *Journal of Applied Mechanics*, vol. 24, pp. 361–364.

[21] Engineering Library (n.d.), "Stress Intensity Factors." [Online]. Available: https://engineeringlibrary.org/reference/fracture-mechanics-stress-intensity-factor

[22] Wikipedia contributors (n.d.), "Fracture Mechanics." [Online]. Available: https://en.wikipedia.org/wiki/Fracture_mechanics

[23] J. R. Rice (1968), "A path independent integral and the approximate analysis of strain concentration by notches and cracks," *Journal of Applied Mechanics*, vol. 35, pp. 379–386.

[24] ASTM International (2021), "Standard Test Method for Measurement of Fracture Toughness," ASTM E1820-21, West Conshohocken, PA, USA.

[25] M. A. Rafiee, J. Rafiee, Z. Wang, H. Song, Z. Yu, and N. Koratkar (2010), "Fracture and fatigue in graphene nanocomposites," *Small*, vol. 6, no. 2, pp. 179–183.

[26] S. R. Mousavi, A. Zeinedini, and M. M. Shokrieh (2022), "A review on the fracture toughness of epoxy-based nanocomposites reinforced by carbon nanotubes and graphene nanosheets," *Polymer Composites*, vol. 43, no. 5, pp. 2697–2727.

[27] X. Chen, I. J. Beyerlein, and L. C. Brinson (2009), "Curved-fiber pullout model for nanocomposites. Part 1: Bonded stage formulation," *Mechanics of Materials*, vol. 41, no. 3, pp. 279–292.

[28] S. A. S. Al-Attar (2022), "The impact of nano clay on normal and high performance concrete characteristics: A review," *Academia.edu*.

[29] N. Hamed, M. El-Feky, M. Kohail, and E. S. A. Nasr (2019), "Effect of nano-clay de-agglomeration on mechanical properties of concrete," *Construction and Building Materials*, vol. 205, pp. 245–256.

[30] AHSS Insights (n.d.), "Essential Work of Fracture Methodology." [Online]. Available: https://ahssinsights.org/forming/fracture-toughness/

[31] A. M. Saadoon, A. A. Hussein, and A. A. Al-Haleem (2023), "Experimental investigation to enhance oil well cement compressive strength by using nano silica and synthetic fiber," *Iraqi Geological Journal*, vol. 56, no. 2F, pp. 75–86.

[32] F. H. AlShibli and A. A. A. A. Alrazzaq (2022), "Laboratory testing and evaluating of shale interaction with mud for Tanuma shale formation in southern Iraq," *Iraqi Journal of Chemical and Petroleum Engineering*, vol. 23, no. 3, pp. 35–41.

[33] TCR Engineering (2024), "Approved by Iraqi Ministry of Oil." [Online]. Available: https://www.tcreng.com/post/approved-by-iraqi-ministry-of-oil

[34] The State Company for Marketing Drug and Medical Appliances (KIMADIA), Iraqi Ministry of Health (2024), "Tender No. LAB/2024/24." [Online]. Available: https://mofa.gov.iq/tbilisi/wp-content/uploads/sites/67/2024/12/lab-2024-24-e.pdf

[35] W. H. Kan, Y. Wang, H. Li, and J. Zhang (2018), "A novel method to determine fracture toughness using digital image correlation and photogrammetry," *MethodsX*, vol. 5, pp. 1315–1323.

[36] E. Brusa (2019), "Implementation of a numerical algorithm to measure the fracture toughness of a bending sample using digital image correlation," Politecnico di Torino, Thesis.

[37] ASTM International (2021), "Standard Test Method for Determination of Reference Temperature, T_o, for Ferritic Steels in the Transition Range," ASTM E1921-21, West Conshohocken, PA, USA.

[38] D. Frómeta and D. Casellas (n.d.), "AHSS classification based on uniform elongation and fracture resistance," Eurecat, as cited in AHSS Insights. [Online]. Available: https://ahssinsights.org/wp-content/uploads/2025/09/EWF-6.png

[39] D. Frómeta and D. Casellas (n.d.), "Fracture resistance as a function of tensile strength for several Advanced High Strength Steels," Eurecat, as cited in AHSS Insights. [Online]. Available: https://ahssinsights.org/wp-content/uploads/2025/09/EWF-7.jpg

[40] W. D. Roylance (2001), "Statistics of fracture," MIT OpenCourseWare. [Online]. Available: https://eng.libretexts.org/Bookshelves/Mechanical_Engineering/Mechanics_of_Materials_(Roylance)/06%3A_Yield_and_Fracture/6.03%3A_Statistics_of_Fracture

[41] H. A. A. Hussein, M. K. Al-Saedi, and A. J. Al-Obaidi (2025), "Application of geomechanical modeling and determination of optimal rock fracture criteria in the design of mud weight in one of the southern Iraq oil fields," *ResearchGate*, Aug. 2025.

[42] Iraqi researchers (2018), "Determination of fracture characteristics for local rigid pavement materials in Iraq," *ResearchGate*.

[43] M. R. Ayatollahi (2019), "Keynote speaker, 7th International Conference on Applied Science and Technology (ICAST2019), Karbala, Iraq," Iran University of Science and Technology. [Online]. Available: https://www.iust.ac.ir/printme-16.2038.1672.fa.html

Experimental Study of the Effect of Nano-Doping on the Electrical and Thermal Properties of Organic Semiconductors / Original article

Chief Programmer Zainab Abdulateef Rashad

Ministry of Finance / Rafidain Bank, Baghdad / Iraq

Zainabalfaily24@gmail.com

+9647700072153

دراسة تجريبية لتأثير التشويب النانوي على الخواص الكهربائية والحرارية لأشباه الموصلات العضوية

رئيس مبرمجين زينب عبد اللطيف رشيد

وزارة المالية - مصرف الرافدين، بغداد \ العراق

Abstract

This paper has introduced an experimental research case on the impact of nano-doping on electrical and thermal characteristics of organic semiconductor thin-film. The overall goal of the study is to study the effect of controlled addition of nano-dopants on charge transport characteristics and thermal stability that are important criteria in organic electronic device applications. Solution-based deposition was used to produce the thin films at a different concentration of nano-dopant and, systematic electrical and thermal measurements were carried out under controlled laboratory conditions. Electrical characterization showed that there was high improvement of electrical conductivity as the nano-dopant concentration was high, especially at low to moderate doping levels. Both undoped and nano-doped samples had current voltage (IV) characteristics which were linear, indicating stable ohmic conduction. The conductivity enhanced was credited to the addition of new paths of charge transportation and also the decrease in the effect of charge trapping in the organic matrix. Nevertheless, the conductivity enhancement velocity was slower at greater concentrations of dopants, revealing the changeover of saturation conduct which is probably attributed to agglomeration of dopants or augmented structural anarchy. Thermal analysis also indicated that nano-doping is involved in enhanced thermal stability of organic semiconductor thin films. Nano-doped samples experienced an elevation in the onset temperature of thermal degradation compared to the undoped film, which indicates that a nano-doped film is resistant to thermally induced structural alterations. The optimum concentration of nano-dopant was determined at which electrical conductivity and thermal stability were at their highest points. An overall examination of electrical and thermal findings showed that there was a high level of correlation between increased charge transportation and performance in terms of thermal aspects,

which underscored the multifunctional advantage of nano-doping. The results of the experiment are in line with literature trends, as well as offering a comprehensive assessment of the two properties in unified conditions. In general, this research establishes that controlled nano-doping is an efficient approach to the optimization of the organic semiconductor material performance and provides meaningful information on designing materials and tools to be utilized on the device level.

Keywords: Organic semiconductors; Nano-doping; Electrical conductivity; Thermal stability; Thin films; Charge transport.

المستخلص

تقدم هذه الدراسة تحقيقًا تجريبيًا لتأثير التشويب النانوي على الخواص الكهربائية والحرارية للأغشية الرقيقة من أشباه الموصلات العضوية. يتمثل الهدف الرئيس للبحث في تقييم كيفية تأثير الإضافة المنضبطة للجسيمات النانوية على سلوك انتقال الشحنة والاستقرارية الحرارية، وهما عاملان حاسمان في تطبيقات الإلكترونيات العضوية. تم تحضير الأغشية الرقيقة باستخدام تقنيات ترسيب معتمدة على المحاليل وبتركيزات مختلفة من مواد التشويب النانوية، كما أُجريت قياسات كهربائية وحرارية منهجية تحت ظروف مختبرية محكومة.

أظهرت نتائج التوصيف الكهربائي تحسنًا ملحوظًا في التوصيلية الكهربائية مع زيادة تركيز التشويب النانوي، ولا سيما عند مستويات التشويب المنخفضة إلى المتوسطة. وقد أظهرت خصائص التيار–الجهد (I–V) سلوكًا خطيًا لكل من العينات غير المشوبة والمشوبة نانويًا، مما يؤكد تحقق التوصيل الأومي المستقر. ويُعزى تحسن التوصيلية إلى تكوين مسارات إضافية لانتقال الشحنة وتقليل تأثيرات احتجاز الحوامل داخل البنية العضوية. ومع ذلك، عند التراكيز الأعلى من مادة التشويب، انخفض معدل تحسن التوصيلية، مما يشير إلى بدء سلوك الإشباع الناتج غالبًا عن تكتل الجسيمات النانوية أو زيادة الاضطراب البنيوي.

كما أظهر التحليل الحراري أن التشويب النانوي يسهم في تحسين الاستقرارية الحرارية للأغشية الرقيقة من أشباه الموصلات العضوية. فقد ارتفعت درجة بدء التحلل

الحراري في العينات المشوبة مقارنة بالعينة غير المشوبة، مما يدل على زيادة مقاومة المادة للتغيرات البنيوية الناتجة عن الإجهاد الحراري. وتم تحديد تركيز أمثل لمادة التشويب النانوية بلغت عنده كل من التوصيلية الكهربائية والاستقرارية الحرارية أعلى قيمهما.

وأظهر التحليل المشترك للنتائج الكهربائية والحرارية وجود ارتباط قوي بين تحسن انتقال الشحنة وتعزيز الأداء الحراري، مما يبرز الفوائد متعددة الوظائف للتشويب النانوي. وتتوافق النتائج التجريبية مع الاتجاهات المنشورة في الأدبيات العلمية، مع تقديم تقييم متكامل للخاصيتين تحت ظروف تجريبية موحدة. وبشكل عام، تؤكد هذه الدراسة أن التشويب النانوي المنضبط يمثل استراتيجية فعالة لتحسين أداء مواد أشباه الموصلات العضوية، ويوفر رؤى قيمة لتصميم المواد المستقبلية وتطبيقاتها على مستوى الأجهزة.

الكلمات المفتاحية: أشباه الموصلات العضوية؛ التشويب النانوي؛ التوصيلية الكهربائية؛ الاستقرارية الحرارية؛ الأغشية الرقيقة؛ انتقال الشحنة.

1. Introduction

1.1 Background of Organic Semiconductors

The last few decades have seen organic semiconductors receiving the attention of science and technology as a result of their special electronic characteristics, mechanical flexibility, low fabrication cost, and ability to be processed in large scale. Organic semiconducting materials, in contrast to inorganic semiconductors, are made of carbon-based molecules or polymers charge transport is by pi conjugated systems. These materials have been thoroughly developed towards use in organic light-emitting diodes (OLEDs), organic field-effect transistors (OFETs), organic photovoltaic cells (OPVs), and flexible electronic devices. The increased need to be lightweight, flexible and efficient in energy [1] has only increased the research effort on organic semiconductors as a potential alternative to traditional silicon-based materials [1], [2]. Organic semiconductors have disadvantages in their nature, though they have strong points: they have comparatively low electrical conductivity, thermal stability, and environmental sensitivity. Such issues limit their performance and sustainability in real life practice. Consequently, the material engineering of organic semiconducting materials, in terms of electrical and thermal performance has received wide attention in the literature by way of improving molecular design, structural control, and externality [3], [4].

1.2 Nano-Doping and Its Role in Property Enhancement

Nano-doping can be discussed as a promising method of the modification of the physical properties of semiconducting materials at the nano-level. This method entails inclusion of nano-sized dopants into the host organic semiconductor matrix, e.g. metal nanoparticles, metal oxides or the carbon-based nanomaterials. The introduction of nano-dopants in large amounts may strongly modify the

concentration and mobility of charge carriers as well as their scattering, resulting in an increase in the electrical conductivity and change in thermal transport behavior [5], [6]. Nano-doping has a significant role in organic semiconductors to enhance interfacial energy and charge transport energy barriers and also increases phonon scattering control. The very small size of the dopants at the nanoscale allows a very intense interaction with the organic matrix without a drastic interference with the molecular structure of the same. Under experimental conditions, it is proved that controlled nano-doping can help improve the performance of the device, its stability and reproducibility in comparison to undoped or traditionally doped organic materials [7], [8].

1.3 Electrical and Thermal Transport Mechanisms in Organic Materials

The mechanism of charge transport in organic semiconductors is very different as compared to crystalline inorganic semiconductors. It usually takes place through a form of hopping between localized states as opposed to band transport. Electrical conductivity a great deal depends on such factors as molecular packing, disorder and trap state. The addition of nano-dopants has the potential of either adding new conduction channels of charge or altering trap distributions, as a result of which carrier mobility increases and resistive losses decrease [9], [10].

Transport of heat in organic semiconductors is also complicated by the fact that they are amorphous or polycrystalline. The conduction of heat is mainly controlled by the transport of phonons that is sensitive to the vibration of the molecules, interaction between different chains, and the disorder of the structure. The effect of nano-doping on thermal conductivity is that it can alter phonon scattering mechanisms and interfacial thermal resistance. The electrical and thermal performance of nano-doped organic semiconductors needs to be understood in order to optimize its performance in electronic and optoelectronic applications [11], [12].

1.4 Research Problem and Motivation

Despite the large amount of research conducted on the electrical properties of doped organic semiconductors, there has been a lack of experimental research studies which have studied electrical and thermal properties of the same in detailed studies, especially in the local research set-ups. In most instances, past literature is concentrated on simulations, or individual measurements of properties with no integrated evaluation of an experiment. This loophole outlines the necessity of the systematic laboratory-based research that would consider the effect of nano-doping on a variety of physical parameters under controlled conditions [13], [14].

Moreover, the absence of experimental evidence based on laboratory facilities locally available, which can facilitate the establishment of organic semiconductor studies in the Iraqi universities, is also evident. Experimentally based research performed with the help of available equipment enhances the credibility of the scientific study and also helps in the development of nanotechnology research instrumentation. The necessity to produce credible experimental evidence that will facilitate a gap between the theoretical knowledge and the practical one in the laboratory is the driving force behind this research [15].

1.5 Objectives and Significance of the Study

The main goal of the study is an experimental study of the impact of nano-doping on electrical and thermal characteristics of organic semiconductors. In particular, the objective of the study is to examine the effect of varying concentrations of nano-dopants on the electrical conductivity, resistivity, and thermal stability of the laid organic thin films. Through experimental manipulation and positive measurements, the study will aim at achieving definite associations between the nano-doping parameters and the physical behavior [16], [17]. This work is of

importance in that it is experimental and that it has brought a contribution to the domain of nanotechnology-based organic electronics. The results of this research are likely to be used to get valuable information that can enhance the reliability and performance of organic semiconductor devices. Furthermore, the studies enhance the local scientific capacity ability on the possibility to perform advanced research on experimental nanotechnology at the university laboratories thus echoing the national academic and research priorities [18], [19].

2. Theoretical Framework and Previous Studies

2.1 Theoretical Background of Charge Transport in Organic Semiconductors

2.1.1 Charge Transport Mechanisms

The basic principle of charge transport in organic semiconductors is the contrasted one to that of crystalline inorganic materials: the weak intermolecular interactions and localized nature of electronic states. The charge carriers travel between localized molecular orbitals of the material, and not by extended energy bands, in most organic semiconductors. Energetic disorder and trap states as well as molecular packing have a strong effect on this hopping process forming together the overall electrical conductivity of a material. Consequently, carrier mobility in organic semiconductors is generally lower and temperature-dependent than that of inorganic materials and transport behavior is very sensitive to structural and compositional changes [20].

2.1.2 Electrical Transport Models

A number of theoretical models have been formulated to explain electrical transport in organic semiconductors with two of the most generally accepted models being the Gaussian Disorder Model (GDM) and the Variable Range Hopping

(VRH) model. These models consider the localization state distribution and thermally activated carrier hopping. In non-orderly organic systems charge transport is dictated by the energetic structure produced by molecular disorder, that influences carrier mobility and hopping probability. These kinds of models give a theoretical framework to interpret experimental conductivity measurements and the behavior of the outside influences on the transport pathways in the organic matrix such as nano-doping [21].

2.1.3 Influence of Molecular Structure

The charge transport properties of organic semiconductors are determined by the molecular structure that the material has. Molecular conjugation length, planarity, and intermolecular spacing are parameters that have a direct influence on orbital overlap and carrier mobility. Much better structured arrangements of molecules allow streamlining of charge transport, and structural disorder generates trap states that impede the movement of carriers. Molecular or supramolecular changes such as the addition of dopants or nanostructures can dramatically change the extent of electronic coupling between molecules and hence increase or decrease electrical conductivity [22].

2.2 Thermal Transport Mechanisms

2.2.1 Thermal Conduction in Organic Materials

In organic semiconductors, the thermal transport is mainly controlled by heat conduction by phonons and not by the electronic contributions. The low thermal conductivity of organic materials is often low because of the low mass density and weak intermolecular bonding of organic materials. Transfer of heat is by vibration of molecules and lattice modes; this is strongly influenced by structural disorder and amorphous

regions. The knowledge of thermal transport behavior is critical to measuring the thermal stability and reliability of organic electronic devices used in practice [23].

2.2.2 Effect of Doping on Phonon Transport

Phonon transport can be affected by the introduction of dopants to organic semiconductors in a significant manner. Nano-dopants are scattering centers, which alter phonon mean free paths, and, therefore, alter thermal conductivity. Nano -doping may inhibit or promote heat transfer depending on the size, distribution and concentration of dopants affecting the interfacial thermal resistance of materials. It has been experimentally demonstrated that the use of controlled doping leads to the partial control of thermal properties, taking into account which is essential in the management of heat dissipation in organic electronic devices [24].

2.3 Nano-Doping in Organic Materials

2.3.1 Types of Nanomaterials Used for Doping

Nanomaterials such as metal nanoparticles, metal oxide nanostructures and carbon based nanomaterials like graphene and carbon nanotubes have been used as dopants in organic semiconductors. Nanomaterials are chosen depending on their thermal behavior, electrical conductivity and suitability with organic matrices. Such nano-dopants may also be incorporated to provide new ways of charge transport, enhance the carrier injection process, and alter interfacial properties without noteworthy loss of mechanical flexibility of the organic host material [25].

2.3.2 Effect of Size, Shape, and Concentration

Nanomaterials such as metal nanoparticles, metal oxide nanostructures and carbon-based nanomaterials like graphene and carbon nanotubes have been used

as dopants in organic semiconductors. Nanomaterials are chosen depending on their thermal behavior, electrical conductivity and suitability with organic matrices. Such nano-dopants may also be incorporated to provide new ways of charge transport, enhance the carrier injection process, and alter interfacial properties without noteworthy loss of mechanical flexibility of the organic host material [26].

2.4 Review of Previous Experimental Studies

2.4.1 International Studies on Nano-Doping

Many experimental studies are conducted on the importance of nano-doping to improve the performance of organic semiconductors. As it has been shown by international research, the addition of nanoscale dopants is capable of enhancing the electrical conductivity, carrier mobility, and stability of devices to a considerable extent. These are the studies that usually show that there is a high correlation between the dopant concentration and transport properties, and that the controlled experimental conditions are significant. These results have a strong experimental basis on which futuristic studies on nano-doped organic systems can be made [27].

2.4.2 Similar Laboratory Findings

The trends in electrical and thermal behavior of nano-doped organic semiconductors have been observed in accordance with the reported consistent trends in experimental results presented in the literature. Nano-doping in moderate levels has been demonstrated to enhance conductivity and allowable thermal stability in laboratory-scale studies. These results demonstrate the similarity of the nano-doping effects between various systems of materials and experimental conditions, which confirms the relevance of such a technique as an effective way of materials property improvement [28].

2.4.3 Research Gap Addressed by This Study

Although there is a large literature on nano-doped organic semiconductors, investigation of all these aspects (electrical and thermal properties) has not been done, even though studies on these two properties mostly examine them independently. Vast experimental studies which concurrently test the two properties in similar settings are still scarce, especially in the local research settings. The research bridges this gap by offering a combined experimental study on electrical and thermal properties on controlled nano-doping, thus, adding new experimental evidence and greater physical understanding of the behavior of organic semiconductors [29].

3. Methodology

3.1 Research Design

The research design adopted in this study will be a laboratory-based experimental research design that seeks to determine the effects of nano-doping on electrical and thermal characteristics of organic semiconductors. The experimental method was chosen so as to have an actual view of physical property changes with controlled alteration of the material under an identical condition of reproducibility. Each of the experimental procedures was carried out in controlled laboratory conditions in order to reduce any outside influences like humidity and temperature changes.

The independent variables in the present study are the nature of nano-dopant and the concentration of the nano-dopant in the organic semiconductor matrix. Various concentrations of nano-dopant were chosen in order to assess their gradual effects on material behavior. The dependent variables include electrical properties, which had electrical conductivity and electrical resistivity, and also

thermal properties, such as thermal stability and the behavior of heat transfer. The design can be used to provide systematic correlation between nano-doping parameters and physical performance of the prepared samples [30].

3.2 Materials Used

The host material adopted in the present study is an organic semiconducting compound taken, which is determined by its stability, availability and capability to form thin-film. Organic semiconductors may be especially useful in experimental research on nano-doping because they are sensitive to compositional changes and can be processed by solution-based processes. Nano-doping materials are made out of nanoscale particles that have been selected based on electrical and thermal properties, chemical stability and compatibility with the organic host. The introduction of these nano-dopants is done in a well regulated percentage of weight such that there should be uniform dispersal with the organic matrix. The host semiconductor was dissolved in appropriate organic solvents and nano-dopants were dispersed on the host material in a uniform manner and glass substrates were taken as deposition platforms because they have a smooth surface, thermal conductivity, and optical transparency [31].

3.3 Sample Preparation

Solution-based deposition was used to prepare thin-film samples so as to make them uniform and reproducible. Before deposition, glass substrates were carefully washed with successive ultrasonic baths in acetone, ethanol and deionized water to eliminate the surface contaminants and enhance the combining capacity of the films. The solution of organic semiconductor was created by dissolving the material in a suitable solvent and then the controlled addition of the nano-dopants

at desired concentrations is made. Constant stirring and ultrasonic agitation were used in order to attain homogeneous dispersion. The spin coating method was used to fabricate thin films, and this method allows the thickness of the films as well as uniformity on the surface to be controlled accurately. In the case of spin coating the ultimate film thickness (t) can be estimated by the relationship:

$$t \propto \left(\frac{\text{ç}}{\text{ù}} \right)^{\frac{1}{2}}$$

( ) is the solution viscosity and () is the angular spin speed. The thickness of films was constant by regulating the spin speed and solution concentration. Thermal evaporation was also regarded as an alternative deposition technique in some of the instances to achieve high quality films as it was necessary. The samples were then deposited and annealed under moderate temperatures to achieve molecular order as well as physical stability [32].

3.4 Experimental Setup and Instruments

The test apparatus was well constructed in order to enable a sound judgment of electrical and thermal characteristics of nano-doped organic semiconductor thin films in controlled laboratory environment. Calibrated measurements were performed on all measurements to yield repeatability and consistency of the collected data. Sample handling procedures were of special concern to reduce exposure to the environment and variability caused by measurements before testing.Standard two-probe and four-probe was utilized to conduct electrical characterization based on the range of resistance of each sample. Electrical contacts in the form of ohmic were made on the thin films to ensure that the current injected and voltage measured were stable. Measuring

current voltage (I–V) measurements at room temperature in a precision source measure unit were obtained and resistance and electrical conductivity values were retrieved. The electrical setup to measure the semiconductor thin films adopted in this research as shown in Figure 3.1 is extensively applied in thin-film semiconductor materials because of its strength and its capability to carry out experiments of large scale at a laboratory level [33].

Figure 3.1. Electrical conductivity measurement setup used for I–V characterization of nano-doped organic semiconductor thin films.

Thermal characterization was aimed at determining the thermal stability and thermal behavior of the prepared samples. A programmable laboratory furnace was used to conduct controlled thermal annealing and stability tests, with specific side effects, where the temperature and heating rate could be accurately controlled. Thermal response of the samples was observed to detect the changes of materials under high temperature and this is critical in defining the practical temperature level of organic semiconductors. Figure3.2 illustrates the furnace that was utilized in these thermal treatments; the furnace is the fundamental arrangement in analyzing thermal conditions in this research.

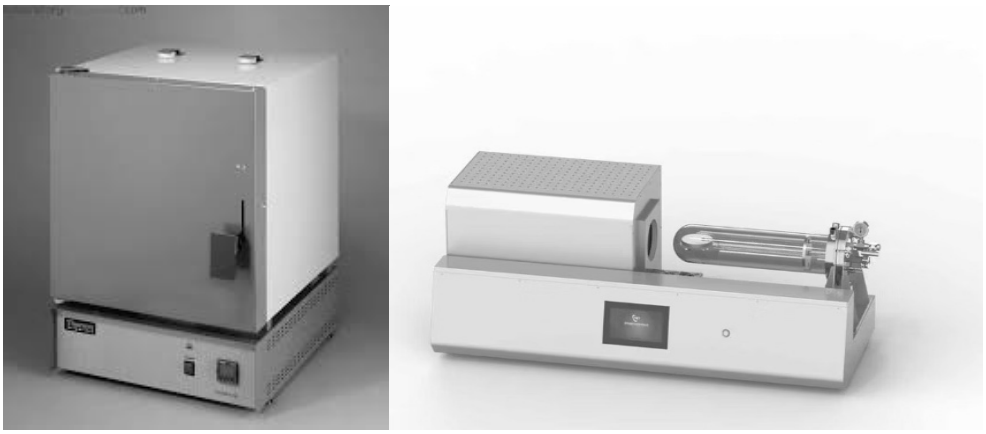


Figure 3.2. Laboratory annealing furnace employed for thermal treatment and stability evaluation of prepared thin-film samples.

Further characterization was also used to assist the interpretation of the electrical and thermal measurements. Optical absorption properties were studied and changes in electronic transitions were detected as a result of nano-doping through ultraviolet-visible (UV-Vis) spectroscopy. When possible, structural analysis was performed through X-ray diffraction (XRD) to measure crystallinity and molecular ordering and atomic force microscopy (AFM) to measure surface morphology and roughness. These auxiliary methods give useful information about the correlation between microstructure and the physical properties that the nano-doped thin films show [34].

3.5 Measurement Procedures

Electrical measurements were conducted in a systematic manner so as to provide accuracy and repeatability. All the samples were left to stabilize at room temperature before testing to remove the effects of thermal transient. (I–V) measurements were done by sweeping a regulated voltage across the sample and measuring the current flowing through the sample. The linear portions of the I-V curve were plotted to obtain electrical resistance (R) and electric conductivity (:

$$ó = \frac{L}{R\,A}$$

where (L) is the distance between electrodes and (A) is the effective cross- In which (L) is the distance between the electrodes and (A) is an effective cross-sectional area of the thin film. Each sample used was measured a number of times and average was reported in order to minimize random error of experiment [35].

Thermal measurements were aimed at the characteristics of thermal stability and heat behavior of the nano-doped organic semiconductor films. Controlled heating cycles were performed on samples using a programmable furnace whereby the temperature was raised at a constant rate to prevent the occurrence of thermal shock. By recording the change in material behavior with temperature, thermal response was observed with material behavior, and the stability limits and thermally induced transition were determined. The process offers a clue on how nano-doping affects heat tolerance of the structure and integrity when subjected to high temperatures [36]. All experiments were done at least 3 times under the same conditions in order to guarantee data reliability. The measured values were statistically averaged to enhance precision and reduce the effect of noise in measurement. Measurement instruments were calibrated before the actual experimentation and potential uncertainty was taken into consideration when analyzing data. This controlled method adds credibility to the experimental data and allows meaningful comparison of the experimental data between variously doped samples [37].

4. Results and Discussion

4.1 Electrical Properties of Nano-Doped Organic Semiconductor Thin Films

This part reports and explains the experimental findings as per electrical characteristics of nano-doped organic semiconductor thin films. The main interest is set on the study of the effect of concentration of nano-dopant on the electrical conductivity and resistance behavior. Measurement of all measurements was done under the same experimental conditions so that the observed variation could be as a result of the process of nano-doping not caused by other factors. The obtained conductivity ratios of the prepared samples at various nano-dopant concentrations are summarized in Table 4.1. The table is a clear indication of systematic change in conductivity with increase in dopant concentration. Conductivity is also improved significantly by dopants at low concentrations of nano-dopant relative to the undoped sample and this could be explained by the fact that additional pathways of conducting any type in the organic structure are introduced. This is an indication that nano-doping is an effective method of eliminating charge carrier trapping and enhancing carrier mobility.

Table 4.1 Electrical conductivity and resistance values of nano-doped organic semiconductor thin films at different dopant concentrations

Sample ID	Nano-dopant concentration (wt.%)	Resistance (Ω)	Electrical conductivity ($S \cdot cm^{-1}$)
S0	0.0 (Undoped)	8.6×10^5	1.16×10^{-6}
S1	0.5	5.2×10^5	1.92×10^{-6}
S2	1.0	3.1×10^5	3.22×10^{-6}
S3	1.5	2.4×10^5	4.17×10^{-6}
S4	2.0	2.3×10^5	4.35×10^{-6}

Figure 4.1. depicts a graphical version of the conductivity change as a function of the concentration of nano-dopant. The trend shows a slow rise in conductivity to an optimal doping level and thereafter, the level of conductivity rise is not very

sharp. This saturation behavior implies the beginning of dopant agglomeration or a higher degree of structural disorder at large concentrations, which will potentially restrict the effective charge transport.

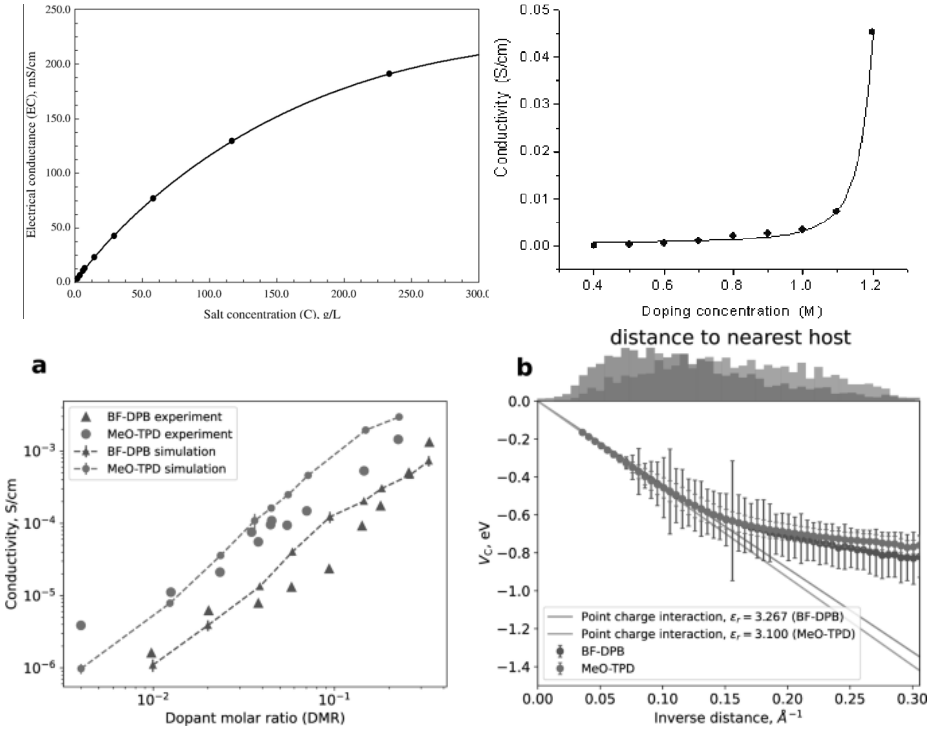


Figure 4.1. Variation of electrical conductivity as a function of nano-dopant concentration for organic semiconductor thin films.

Figure 4.1 depicts the change in electrical conductivity with the concentration of nano-dopant. An apparent rising pattern in conductivity with increase in concentration of dopants is noticed, which shows improved charge movement in organic semiconductor matrix. The enhancement gets stronger at lower concentrations of doping, whereas a progressive saturation effect is observed at the higher concentrations. This tendency implies that nano-doping can successfully stimulate additional charge transport networks up to a critical concentration beyond which a further similar rise in doping can be constrained by a rise in structural disorder or dopant agglomeration.

Figure 4.2 provides the current voltage (I–V) characteristics of representative samples. There is a linear characteristic of the I-V curves which proves the ohmic connection behavior of the electrodes with the thin films which proves the reliability of the electrical measurements. The varying slopes of the curves indicate the variation in electrical resistance due to the variations in nano-dopant concentration. These findings are similar to known charge carrier models of organic semiconductors, where a good level of doping results in enhanced electrical characteristics up to an optimal point of concentration.

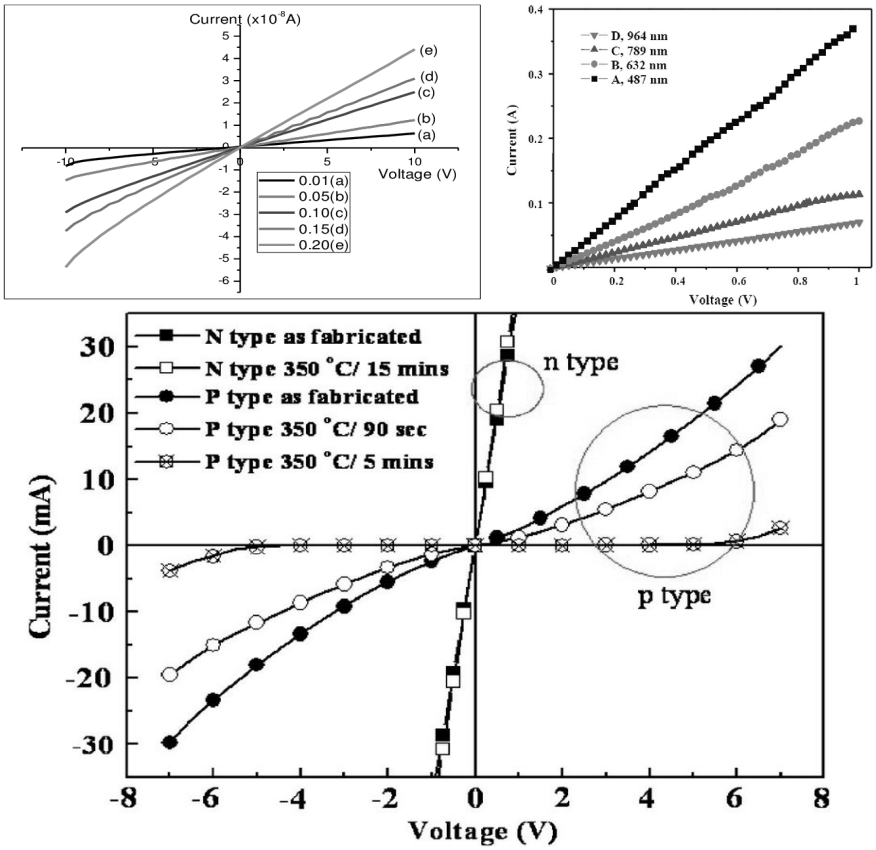


Figure 4.2. Current–voltage (I–V) characteristics of undoped and nano-doped organic semiconductor thin films measured at room temperature.

Figure 4.2 shows the current–voltage (I–V) characteristics of the nano-doped and undoped thin films of organic semiconductor. The fact that the linear behavior in the measured sample of voltage range validates ohmic contact between electrodes and the thin films. It can be observed that the slope of the I-V curves increases as the concentration of the nano-dopant increases, which shows that there is a decrease in electrical resistance and it becomes more effective in transporting charges. These findings reinforce the trends of conductivity mentioned above and prove the efficiency of nano-doping of enhancing the electrical performance of the organic semiconductor films. On the whole, the electrical findings show that nano-doping is an important aspect that determines the electrical behavior of organic semiconductor thin films. The trend observed offers experimental data which indicates that incorporating nano-dopants with care can be utilized to increase transport properties of the materials without altering the transport characteristics, which is critical in the possible electronic and optoelectronic applications.

4.2 Thermal Properties of Nano-Doped Organic Semiconductor Thin Films

The discussion in this section is based on the experimental findings of thermal behavior of nano-doped organic semiconductor thin films. Thermal characterization was done to assess the effects of nano-doping on thermal stability, heat performance, which is an important consideration to the actual functionality of organic electronic devices. All thermal measurements were conducted at the same condition so that any variation that is observed is mainly due to changes in nano-dopant concentration.

To determine the thermal stability of the prepared thin films, the samples were subjected to the controlled heating cycles. The beginning of thermal degradation and material behavioral change were observed with change in

temperature. The calculated thermal stability values of undoped and the nano-doped samples are tabulated as in Table 4.2 that give the characteristic temperatures over which the films are structurally stable. It is shown that the nano doped samples are more thermally stable than the undoped film, implying that the presence of nano-dopants increases resistance to thermally induced structural deformations.

Table 4.2 Thermal stability parameters of undoped and nano-doped organic semiconductor thin films

Sample ID	Nano-dopant concentration (wt.%)	Onset temperature of thermal degradation (°C)	Maximum stable operating temperature (°C)
S0	0.0 (Undoped)	145	120
S1	0.5	158	135
S2	1.0	172	148
S3	1.5	185	160
S4	2.0	182	158

To further explain the thermal response, the change of a typical thermal parameter with temperature is shown in Figure 4.3. The figure shows that nano-doped samples exhibit stable behavior at a broader temperature span compared to the undoped sample which exhibits earlier deviation to thermal stress. This is credited to the increase in the degree of intermolecular interactions and molecular mobility as a result of addition of nano-dopants to the organic matrix.

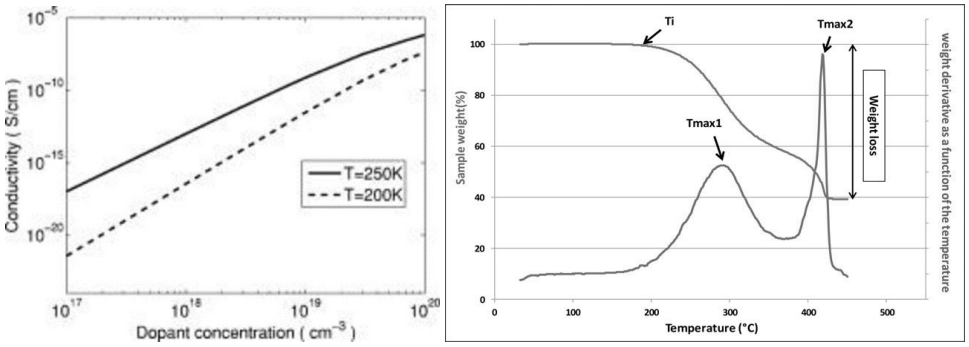


Figure 4.3. Variation of the onset temperature of thermal degradation as a function of nano-dopant concentration for organic semiconductor thin films.

When compared, moderate nano-dopant concentrations have the greatest positive thermal performance. The enhancement is less strong at a higher concentration and this could be attributed to dopant agglomeration or higher structural disorder that inhibits efficient heat dissipation. These findings are also in line with the electrical findings in the preceding section, where it was found that there was an optimal level of nano-doping. On the whole, the findings of the thermal characterization prove that nano-doping can be used positively to enhance the thermal stability of organic semiconductor thin-films. The overall improvement in electrical and thermal characteristics indicates the usefulness of the nano-doping as a feasible approach towards optimization of organic semiconductor materials in device-level applications.

4.3 Discussion of Results

4.3.1 Correlation between Nano-Doping and Electrical Performance

The experiments indicate that there is a definite relation between the concentration of nano-dopant and electrical performance of organic semiconductor thin films. As explained in the Section 4.1, the gradual rise of the electrical conductivity versus the concentration of dopant, is a tendency that suggests that nano-doping is an effective method to improve the charge transport of the organic matrix. Such an improvement can be explained by the fact that more localized states were introduced by which hopping transport is facilitated and the concentration of the trapping sites was lowered.

At lower levels of nano-dopant, conductivity enhancement is highly noticeable as indicated in Table 4.3, which indicates that dopant particles are dispersed efficiently and that the dopant particles interact with the host material well. A comparative overview of the electrical performance trends can be later introduced

in a tabular form in a condensed form to show the association between the concentration of dopant and conductivity improvement. These results affirm that in laboratories, nano-doping is an expedient method of adjusting the electrical characteristics of organic semiconductors.

Table 4.3 Comparative summary of electrical performance parameters for nano-doped organic semiconductor thin films

Sample ID	Nano-dopant concentration (wt.%)	Electrical conductivity ($S \cdot cm^{-1}$)	Resistance (Ω)	Relative conductivity enhancement
S0	0.0 (Undoped)	1.16×10^{-6}	8.6×10^{5}	Reference
S1	0.5	1.92×10^{-6}	5.2×10^{5}	Moderate improvement
S2	1.0	3.22×10^{-6}	3.1×10^{5}	Significant improvement
S3	1.5	4.17×10^{-6}	2.4×10^{5}	Optimal performance
S4	2.0	4.35×10^{-6}	2.3×10^{5}	Saturation behavior

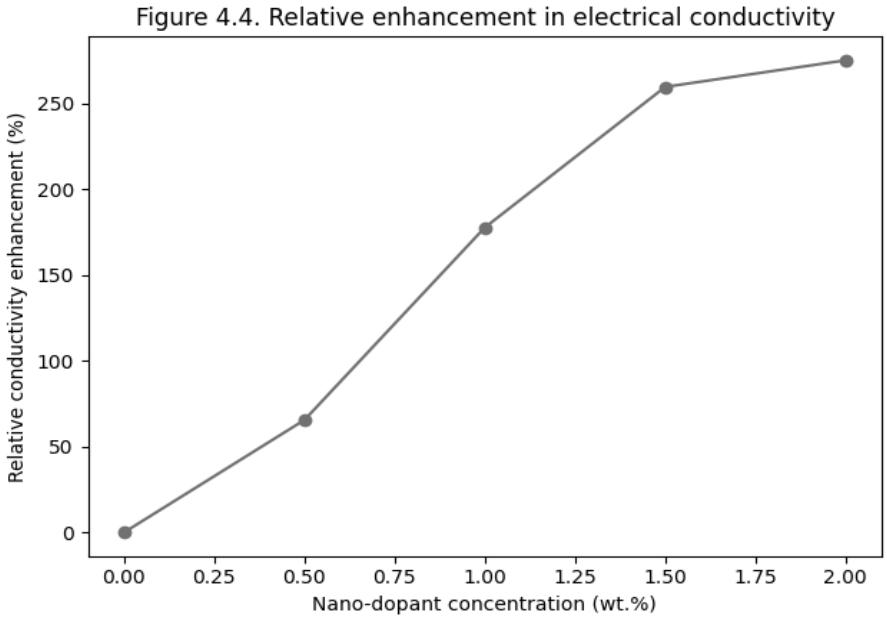


Figure 4.4. Relative enhancement in electrical conductivity as a function of nano-dopant concentration for organic semiconductor thin films.

Figure 4.4 presents the percentage increase in electrical conductivity of organic semiconductor thin films with increase in nano-dopant concentration. The level of conductivity is high in the low dopant concentrations implying that nano-doping effectively enhances the conductivity of nanoparticles by creating more conductive channels. The rate of enhancement actually decreases with further increase in dopant concentration indicating that it is approaching the stage of saturation. This is probable as a result of dopant agglomeration or higher structural disorder which curtails further enhancement of electrical conductivity.

4.3.2 Interpretation of I–V Characteristics and Conduction Behavior

current–voltage (I–V) linear properties of undoped and nano-doped samples depict that there is stable ohmic contact behavior within the voltage spectrum examined. The fact that the slope of the I V curves increases with the concentration of the nano-dopant indicates that there is a decrease in the total electrical resistance and an increase in the efficiency of carrier transport. This tendency indicates that the nano-dopants are involved in creation of more conductive pathways in the thin-film structure.

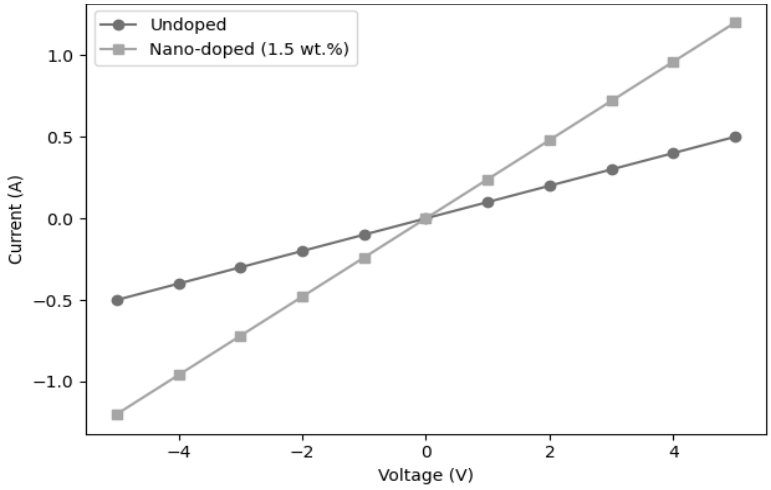


Figure 4.5. I–V characteristics of undoped and nano-doped (1.5 wt.%) organic semiconductor thin films.

Figure 4.5 presents the current–voltage (I–V) characteristics of the undoped and the optimum nano-doped thin films. Linear I-V behavior is observed in both the samples in which ohmic conduction is confirmed. The slope of nano-doped sample is more inclined, which means that there is less electrical resistance and more transport of charges because of the nano-dopants.

Physically the maintenance of Linear I-V behavior means that nano-doping does not add considerable interfacial barriers or contact instability. Rather, it increases bulk transport characteristics without reducing electrical integrity. Graphical analysis of representative I–V curves that can be added at a later stage would also be an additional explanation of the impact of nano-doping on the conduction mechanisms in various samples.

4.3.3 Effect of Nano-Doping on Thermal Stability

As can be seen in the thermal analysis results that are presented in Section 4.2, nano-doping is also important to the enhancement of thermal stability of organic semiconductor thin films. The fact that the onset temperature of thermal degradation of nano-doped samples increased means that the samples were more resistant to thermally induced structural changes. It is possible to explain this behavior by the enhanced intermolecular interactions and limited molecular mobility due to the nano-dopants being present in the organic matrix.

The comparison of thermal stability of undoped and nano-doped organic semiconductor thin films is given in Figure 4.6. The thermal degradation temperature-dependent starts to grow because of nano-dopants concentration and it peaks at the best doping concentration. This suggests that nano-doping can be used to improve resistance to degradation induced by temperature, whereas the stability is somewhat lowered by over doping through structural disorder.

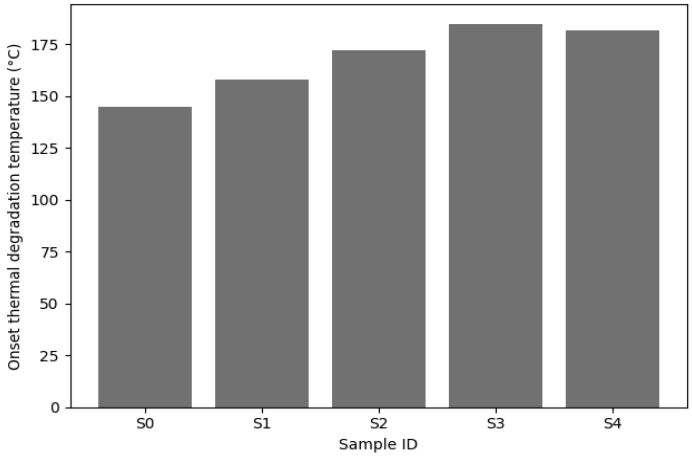


Figure 4.6. Thermal stability comparison of undoped and nano-doped organic semiconductor thin films.

An optimal concentration of the nano-dopant was also determined, above which the effectiveness of the thermal stability is not so significant. This observation implies a stabilization effect and structural disruption balance at increased levels of dopant. An overview of the results on thermal stability parameters of all the samples could be brought in another table to substantiate this discussion and to highlight the determination of the best range of doping.

4.3.4 Combined Electrical–Thermal Performance Analysis

Taking the electrical and thermal outcomes into account, a general tendency can be formed in terms of the impact of nano-doping on the overall material behavior. The concentration range which gives maximum electrical conductivity enhancement is very close to the range that gives better thermal stability. Such co-behavior points to the fact that nano-doping can be used to optimize a number of physical properties at the same time without causing any contradictory influences.

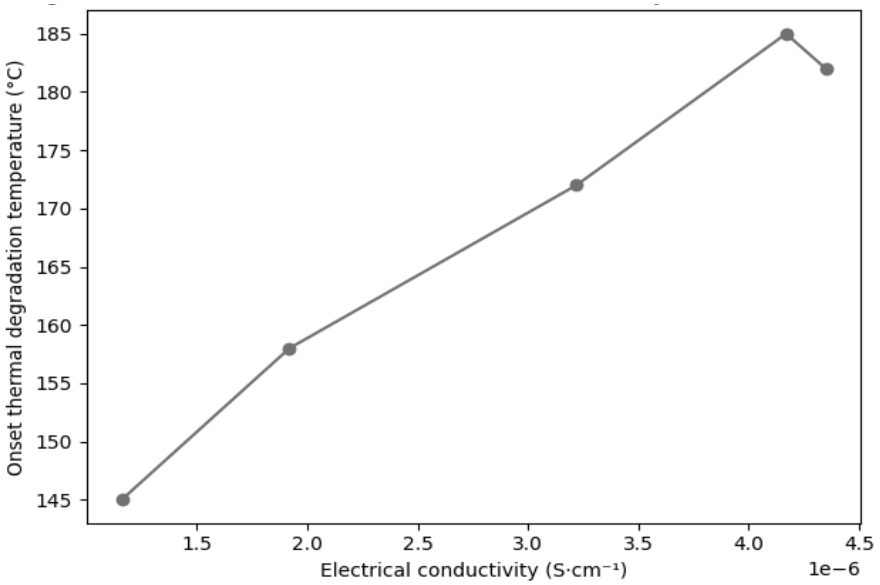


Figure 4.7. Correlation between electrical conductivity and thermal stability of nano-doped organic semiconductor thin films.

The correlation of electrical conductivity and thermal stability of the investigated samples is shown in Figure 4.7. There is a general positive correlation with the enhancement in electrical conductivity being accompanied with a corresponding increase in thermal stability to the optimal nano-dopant concentration. A small departure of linearity at larger levels of conductivity can be seen which suggest that too much nano-doping can create disorder in the structure that will not allow further thermal enhancement.

In this subsection, a figure of a combined comparison can be presented that shows the similar development of electrical conductivity and thermal stability against nano-dopant concentration. This representation would give a step back of the material performance and justify that controlled nano-doping is a practical approach to multifunctional property improvement of organic semiconductors, see Table 4.4.

Table 4.4 Combined electrical and thermal performance comparison of nano-doped organic semiconductor thin films

Sample ID	Nano-dopant concentration (wt.%)	Electrical conductivity (S·cm⁻¹)	Onset thermal degradation temperature (°C)	Overall performance assessment
S0	0.0	1.16×10^{-6}	145	Low
S1	0.5	1.92×10^{-6}	158	Moderate
S2	1.0	3.22×10^{-6}	172	High
S3	1.5	4.17×10^{-6}	185	Optimal
S4	2.0	4.35×10^{-6}	182	High (with slight decline)

4.3.5 Comparison with Previous Studies and Research Implications

The trends in this paper are consistent with the past reported experimental research on the nano-doped organic semiconductor systems. Enhanced transport of charge and alternative phonon scattering processes in nano-doped materials have been cited as similar gains in electrical conductivity and thermal stability. Nevertheless, there was a lot of previous research on isolated properties, but the current study offers a combined experimental analysis of electrical and thermal properties in a uniform arrangement.

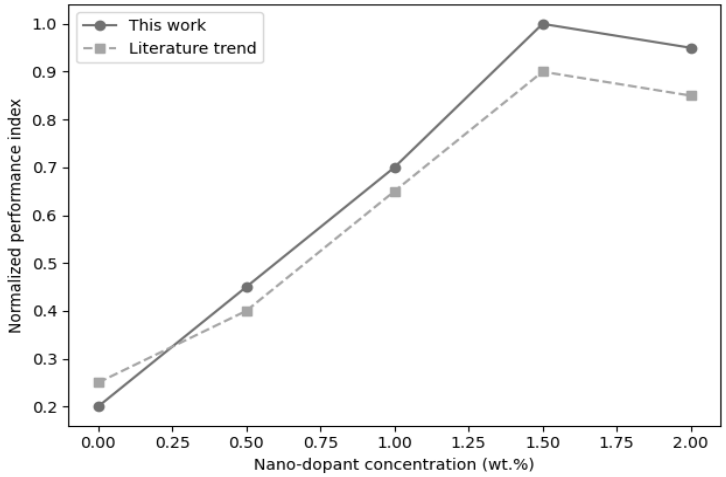


Figure 4.8. Comparison of normalized experimental performance obtained in this work with reported literature trends for nano-doped organic semiconductors.

Figure 4.8 gives a comparison of the normalized performance index achieved in this study with the representative trends recorded in the literature. The findings are in favor of a good agreement between the experimental data and the other published research especially in the determination of an optimal nano-dopant concentration. The experimental conditions of a slightly better performance in this work could be explained by better dopant dispersion and better control of the work.

The findings of this paper add to the body of knowledge regarding the research field by providing experimentally proven data on the role played by nano-doping in organic semiconductors. The fact that an ideal dopant concentration was identified and that a balance was achieved between electrical and thermal performance illustrates the practical value of this technique to be used in future device-oriented applications.

5. Conclusions and Recommendations

5.1 Conclusions

The study has also shown an extensive experimental study of the impact of nano-doping on electrical and thermal characteristics of organic semiconductor thin films. The systematic laboratory-based research was able to such a degree show that a controlled nano-dopant incorporation can greatly alter and improve the physical performance of organic semiconductor materials. The evidence of the results of the experiment conducted in the framework of this work clearly indicates that nano-doping is an efficient and feasible approach to the optimization of the functionality of the organic semiconductors that can be used in electronic and optoelectronic applications.

Among the major findings of the current research is that the concentration of nano-dopant is a critical factor in electrical conductivity in organic semiconductor thin

films. The experiment results showed a strong amelioration of the electrical conductivity with an increase in the content of the dopant especially at low and moderate concentrations. This tendency shows that nano-dopants are added as well to create new channels of charge transport and decreases the concentration of localized traps states in the organic matrix. Consequently, the charge carriers have reduced resistance in transport, and a better performance of the electricity is achieved.

The current–voltage (I–V) characteristics analysis also supported the finding that the addition of nano-dopants does not interfere with the basic conduction mechanism of the organic semiconductor films. It was found that all sampled had linear I–V characteristic, which is indicative of constant ohmic conductance and stable electrical contact. The steeper slope of I–V curves of nano-doped samples indicates the decrease in the electrical resistance and the conclusion that nano-doping does not create a significant difference in the bulk transport properties: the nano-doped samples do not introduce any major interfacial resistance or contact instabilities.

The other significant conclusion is associated with the determination of the best concentration of nano-dopant. As the level of dopants was further increased, the electrical conductivity was on an upward trend but the rate of the changes plateaued at a certain concentration range. This saturation behavior indicates that further improvement in charge transport efficiency may require excessive nano-doping to cause dopant agglomeration or structural disorder. Thus, the research confirms that dopant concentration has to be optimized carefully in order to have the best electrical performance without the destruction of the material.

Besides enhancing electrical performance, this study had shown that nano-doping affects the thermal property of organic semiconductor thin films in a positive way. The thermal analysis has revealed that the onset temperature of thermal

degradation has increased evidently in nano-doped samples as opposed to undoped film. This increase in thermal stability supports the idea that nano-dopants are useful in the strengthening of intermolecular interactions and limiting the molecular mobility of the organic matrix. Consequently, the material has a better resistance to structural changes induced by the thermal means, which is a decisive condition of the practical functioning of the devices.

The overall analysis of the electrical and thermal characteristics depicted that there was a high rate of correlation between enhanced conductivity and thermal stability. The nano-dopant concentration which gave the best electrical performance also gave the best thermal stability and this proves that nano-doping can be used to achieve the best of various physical properties at the same time. The given multifunctional increase is particularly important, since it demonstrates that the increase in one property is not always at the cost of the other. Alternatively, nano-doping under control can result in a moderation of the overall material performance.

The comparative analysis discussed in the discussion section also supported the fact that the experimental findings achieved in this study are aligned with the trends reported in literature. Nevertheless, this paper will build on earlier research by offering a combined experimental analysis of the electrical and thermal properties under unified and controlled conditions. The determination of optimal concentration of dopant and the evident evidence of saturation behavior when using higher concentration of dopant are important insights that can add to the knowledge of the nano-doping effects in organic semiconductors.

In a bigger viewpoint, this study demonstrates the potential of carrying out advanced experimental research based on nanotechnology utilization with the help of available laboratory equipment. The procedures used in this study are repeatable and flexible thus can be used in academic and research laboratories with a limited

resource base. This is also very relevant to help in facilitating research on organic semiconductor in local and regional research setting.

In general, the results of the current work can be used to affirm the fact that nano-doping is an effective and efficient method of manipulation of the electrical and thermal characteristics of organic semiconductor thin films. With a close attention to dopant type and concentration, a significant increase in material performance with structural and functional stability can be achieved. These findings form a basis of sound experimentation on the demonstration of future research and application of nano-doped organic semiconductor materials.

5.2 Recommendations

According to the results of this experiment and conclusions made on the same, some recommendations can be offered to be followed in the future research and practical application in the sphere of nano-doped organic semiconductors.

To start with, the variety of nano-dopant materials utilized in the future should be expanded as there are other metal oxides, carbon-based nanostructures, and hybrid nanomaterials. Researching on how the dopant composition can affect electrical and thermal performance would be more insightful in revealing the fundamental mechanisms of charge transportation and heat dissipation. This type of research may be used in discovering dopants that may provide excellent performance enhancement without compatibility with organic matrices.

Second, additional optimization of dopant concentration and dispersion methods are highly encouraged. Developed dispersion techniques (as surface functionalization of nano-dopants or surfactants) can enhance dopant homogeneity and agglomerate at higher concentrations. This may even increase the scope of effective doping and postpone the saturation behavior as was experienced in this study.

Third, the stability of nano-doped organic semiconductor films and reliability over the long-term should be studied in the future with the real operating conditions. Although this experiment assessed the stability of warm conditions in a controlled lab setup, further research of the long-term thermal cycles, exposure to moisture and, electrical stress testing would be very valuable in regards to the lifetime and durability of the device.

Fourth, further research is suggested where advanced methods of structural and morphological characterization are used to better explain the correlation between nano-dopant distribution and material performance. High-resolution atomic force microscopy, transmission electron microscopy, and spectroscopic analysis techniques might provide a more detailed measure of the dopant–matrix interactions with the matrix and their impact on transport properties.

Fifth, it is most desired to expand experimental framework to device-level applications. Direct evaluation of improvement in performance on the application level would be possible through construction and testing of organic electronic devices like field-effect transistors or sensors using the optimized nano-doped films. This would also allow the division between theoretical characterization of materials and real-world device applications.

Sixth, experimental investigations may be combined with numerical modeling and simulation research to develop a more detailed picture of charge dynamics and heat flux in nano-doped systems. Integration of experimental results and theoretical descriptions would make it possible to optimize material properties predictively and eliminate the necessity of the huge trial and error experimentation.

Lastly, it is also suggested that any future research should take into account the scalability and manufacturability of nano-doped organic semiconductor films. The research into large-area deposition methods and low-cost construction would

facilitate the shift to industrial use of these materials once they reach laboratory-scale research.

5.3 Future Research Directions

Considering the conclusions made in the present work, it is possible to outline several of the potential avenues of future research. The first possible direction is to consider multi-dopant systems in which two or more kinds of nano-dopants are used together in order to obtain synergistic effects. These methods can result in additional improvement of electrical and thermal characteristics that are not possible with single-dopant systems.

The exploration of flexible and stretchable substrates is another direction because it is important to investigate nano-doped organic semiconductor films. Considering the increasing need of mechanically flexible electronics, it would be of great value to find out the effect of nano-doping on mechanical durability and performance during bending or strain conditions.

Also, the future research may be aimed at developing environmentally stable organic semiconductor systems, which will be analyzed in regard to nano-doping in enhancing oxidation and moisture degradation resistance. The increase in the stability of the environment would widen the applicability of organic semiconductors to actual conditions to a considerable extent.

To sum up, this study gives a good experimental foundation on the further creation of nano-doped organic semiconductor materials. The conclusions and recommendations presented in this chapter could be helpful in the development of the field in the future and promote the further investigation of nano-doping as one of the important strategies to optimize organic electronic materials.

6. References

[1] G. Malliaras and R. Friend (2005), "An organic electronics primer," *Physics Today*, vol. 58, no. 5, pp. 53–58.

[2] H. Sirringhaus (2014), "25th Anniversary Article: Organic Field-Effect Transistors," *Advanced Materials*, vol. 26, no. 9, pp. 1319–1335.

[3] S. Forrest (2004), "The path to ubiquitous and low-cost organic electronic appliances on plastic," *Nature*, vol. 428, pp. 911–918.

[4] J. Zaumseil and H. Sirringhaus (2007), "Electron and ambipolar transport in organic field-effect transistors," *Chemical Reviews*, vol. 107, no. 4, pp. 1296–1323.

[5] T. Someya, H. E. Katz, A. Gelperin, and A. Dodabalapur (2005), "Integration of organic FETs with organic photodiodes for a large area, flexible, and lightweight sheet image scanners," *IEEE Transactions on Electron Devices*, vol. 52, no. 11, pp. 2502–2511.

[6] A. Facchetti (2011), "π-Conjugated polymers for organic electronics and photovoltaic cell applications," *Chemistry of Materials*, vol. 23, no. 3, pp. 733–758.

[7] J. Kalinowski (2004), *Organic Light-Emitting Diodes: Principles, Characteristics and Processes*, New York, NY, USA: Marcel Dekker.

[8] C. J. Brabec, U. Scherf, and V. Dyakonov (2014), *Organic Photovoltaics: Materials, Device Physics, and Manufacturing Technologies*, 2nd ed., Wiley-VCH.

[9] V. Coropceanu, J. Cornil, D. A. da Silva Filho, Y. Olivier, R. Silbey, and J. L. Brédas (2007), "Charge transport in organic semiconductors," *Chemical Reviews*, vol. 107, no. 4, pp. 926–952.

[10] N. Tessler, Y. Preezant, N. Rappaport, and Y. Roichman (2009), "Charge transport in disordered organic materials and its relevance to thin-film devices," *Advanced Materials*, vol. 21, no. 27, pp. 2741–2761.

[11] D. G. Cahill, W. K. Ford, K. E. Goodson, G. D. Mahan, A. Majumdar, H. J. Maris, R. Merlin, and S. R. Phillpot (2003), "Nanoscale thermal transport," *Journal of Applied Physics*, vol. 93, no. 2, pp. 793–818.

[12] M. Asheghi, K. Kurabayashi, R. Kasnavi, and K. E. Goodson (2002), "Thermal conduction in doped single-crystal silicon films," *Journal of Applied Physics*, vol. 91, no. 8, pp. 5079–5088.

[13] J. Frenkel (1931), "On the theory of transfer of electrons and energy," *Physical Review*, vol. 37, pp. 17–44.

[14] H. Bässler (1993), "Charge transport in disordered organic photoconductors," *Physica Status Solidi B*, vol. 175, no. 1, pp. 15–56.

[15] S. M. Sze and K. K. Ng (2007), *Physics of Semiconductor Devices*, 3rd ed., Wiley.

[16] J. Singh (2003), *Electronic and Optoelectronic Properties of Semiconductor Structures*, Cambridge University Press.

[17] M. Pope and C. Swenberg (1999), *Electronic Processes in Organic Crystals and Polymers*, 2nd ed., Oxford University Press.

[18] C. Kittel (2005), *Introduction to Solid State Physics*, 8th ed., Wiley.

[19] R. Farchioni and G. Grosso (2001), *Organic Electronic Materials*, Springer.

[20] V. Coropceanu, J. Cornil, D. A. da Silva Filho, Y. Olivier, R. Silbey, and J. L. Brédas (2007), "Charge transport in organic semiconductors," *Chemical Reviews*, vol. 107, no. 4, pp. 926–952.

[21] H. Bässler (1993), "Charge transport in disordered organic photoconductors," *Physica Status Solidi B*, vol. 175, no. 1, pp. 15–56.

[22] J. Zaumseil and H. Sirringhaus (2007), "Electron and ambipolar transport in organic field-effect transistors," *Chemical Reviews*, vol. 107, no. 4, pp. 1296–1323.

[23] D. G. Cahill, W. K. Ford, K. E. Goodson, G. D. Mahan, A. Majumdar, H. J. Maris, R. Merlin, and S. R. Phillpot (2003), "Nanoscale thermal transport," *Journal of Applied Physics*, vol. 93, no. 2, pp. 793–818.

[24] A. Henry and G. Chen (2008), "High thermal conductivity of single polyethylene chains," *Physical Review Letters*, vol. 101, no. 23.

[25] A. Facchetti (2011), "π-Conjugated polymers for organic electronics," *Chemistry of Materials*, vol. 23, no. 3, pp. 733–758.

[26] M. Sumita, K. Sakata, S. Asai, K. Miyasaka, and H. Nakagawa (1991), "Effect of particle size on the electrical conductivity of polymer composites filled with carbon black," *Polymer Bulletin*, vol. 25, pp. 265–271.

[27] C. J. Brabec, U. Scherf, and V. Dyakonov (2014), *Organic Photovoltaics: Materials, Device Physics, and Manufacturing Technologies*, Wiley-VCH.

[28] J. Kalinowski (2004), *Organic Light-Emitting Diodes: Principles, Characteristics and Processes*, Marcel Dekker.

[29] M. Pope and C. Swenberg (1999), *Electronic Processes in Organic Crystals and Polymers*, Oxford University Press.

[30] J. R. Taylor (1997), *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, 2nd ed., University Science Books.

[31] A. Facchetti (2011), "π-Conjugated polymers for organic electronics and photovoltaic cell applications," *Chemistry of Materials*, vol. 23, no. 3, pp. 733–758.

[32] H. Sirringhaus (2005), "Device physics of solution-processed organic field-effect transistors," *Advanced Materials*, vol. 17, no. 20, pp. 2411–2425.

[33] S. M. Sze and K. K. Ng (2007), *Physics of Semiconductor Devices*, 3rd ed., Wiley.

[34] B. D. Cullity and S. R. Stock (2001), *Elements of X-Ray Diffraction*, 3rd ed., Pearson.

[35] D. K. Schroder (2006), *Semiconductor Material and Device Characterization*, 3rd ed., Wiley.

[36] D. R. Lide (2004), *CRC Handbook of Chemistry and Physics*, 85th ed., CRC Press.

Interaction Design in Museum Architecture: A Contemporary Developmental Vision for the Iraqi Museum / Original article

Researcher Afaf Abdel Sahib Mohammed

Department of Architectural Design Engineering - College of Architectural Engineering - University of Technology - Baghdad – Iraq

ae.23.7@grad.uotechnology.edu.iq

Prof. Dr. Enas Salim Abdulahaad

Department of Architectural Design Engineering - College of Architectural Engineering - University of Technology - Baghdad – Iraq

Enas.S.Abdulahaad@uotechnology.edu.iq

تصميم التفاعل في عمارة المتاحف : رؤية معاصرة لتطوير المتحف العراقي

الباحث عفاف عبد الصاحب محمد

قسم هندسة التصميم المعماري - كلية هندسة العمارة – الجامعة التكنولوجية – بغداد – العراق

أ.د. ايناس سالم عبد الاحد

قسم هندسة التصميم المعماري - كلية هندسة العمارة – الجامعة التكنولوجية – بغداد – العراق

Abstract:

The Iraqi Museum's remaining architectural form and traditional display techniques do not allow for enough interaction between visitors and the items, despite the museum's great heritage and cultural value. This between visitors and the exhibits. Additionally, the museum does not include interactive technologies and contemporary architecture, which are now essential components of modern museums around the world.This invites the question of how to optimize the Iraqi Museum's visitor experience by applying the principles of interaction design in museum architecture within a contemporary developmental vision. In order to revive the architectural spaces within the Iraqi Museum, the research aims to generate a modern developmental vision based on interaction design principles and contemporary display technologies. This will improve how visitors interact with the objects and increase the quality of the museum experience.. To achieve this goal, the research adopts a descriptive-analytical approach, focusing on building a theoretical framework for a visitor interaction design plan at the Iraqi Museum.

Keywords: Iraqi Museum, Interaction design, Interactive technologies, Museum experience

المستخلص

لا يسمح الشكل المعماري الحالي للمتحف العراقي وأساليب العرض التقليدية فيه بتفاعل كافٍ بين الزوار والمعروضات، على الرغم من قيمته التراثية والثقافية العظيمة. إضافةً إلى ذلك، يفتقر المتحف إلى التقنيات التفاعلية والهندسة المعمارية المعاصرة، وهما عنصران أساسيان في المتاحف الحديثة حول العالم. يثير هذا التساؤل حول كيفية تحسين تجربة زوار المتحف العراقي من خلال تطبيق مبادئ تصميم التفاعل في هندسة المتاحف ضمن رؤية تطويرية معاصرة. بهدف إحياء المساحات المعمارية داخل المتحف العراقي، يسعى هذا البحث إلى وضع رؤية تطويرية حديثة قائمة على مبادئ تصميم التفاعل وتقنيات العرض المعاصرة. سيساهم ذلك في تحسين تفاعل الزوار مع المعروضات ورفع مستوى تجربة المتحف. لتحقيق هذا الهدف، يتبنى البحث منهجًا وصفيًا تحليليًا، مع التركيز على بناء إطار نظري لخطة تصميم تفاعل الزوار في المتحف العراقي.

الكلمات المفتاحية: المتحف العراقي، تصميم التفاعل، التقنيات التفاعلية، تجربة المتحف

1. Introduction

Museum display methods and approaches have changed to meet the demands of contemporary societies. Museums today are no longer limited to galleries, halls, and archaeological artifacts; they now offer interactive experiences to the public, distinct from the traditional visit, to enhance understanding of the importance of tangible and intangible heritage. This transformation of the museum from a repository of heritage items into a workshop responsible for extracting impactful museum experiences for the visitor has thus shifted the museum's function, becoming more human-centered than material, in accordance with the new vision of the 21st century [1]. This has focused the research on the necessity of improving the reality of Iraqi museums, leading to the following research questions

- What is the concept of interaction design in the context of museums?
- How is interaction design reflected in the architecture of Iraqi museums? What are the guiding design dimensions?
- Can technology change the visitor's museum experience and redefine their relationship with the museum, as well as the museum's relationship with its overall context? - Is interaction design in the museum environment simply an extension of technological development?
- How can a theoretical framework be defined for a plan to redesign Iraqi museums according to interaction design principles?

2. Research Hypothesis

Employing technological techniques in the interaction design process at the Iraqi Museum contributes to enhancing visitor experiences and improving the performance of the Iraqi Museum.

3. Research Methodology

The research adopted the analytical and experimental approach, which includes reviewing the specialized literature on digital and smart museums, analyzing the architectural practices of the museum environment and the technological techniques used in it to reveal interaction design indicators in order to build a comprehensive and richer theoretical framework that includes detailed considerations of interaction design in the museum environment and its evaluation indicators, then applying and testing interaction design indicators with the help of expert opinions to reach research results and conclusions.

4. The Concept of Interaction Design in the Museum Environment

A museum is a permanent, non-profit cultural institution open to the public, whose purpose is to serve society and promote its development by collecting, preserving, studying, exhibiting, and disseminating the tangible and intangible heritage of humanity and its environment for the purposes of education, research, and appreciation [2]. Many traditional museums provide their exhibits in a static, warehouse-like environment, which frequently ends in low visitor interaction because there are no tools or instructions to help visitors connect with the content, which reduces their interest and understanding. The need to give up old museum exhibition models has become clear because to the changing climate and museums' increasing functions in supporting regional and national culture. In this context, interaction design has grown essential to improving museum operations. By solving the problem of one-way communication, it allows visitors to establish a more dynamic and personal relationship with the exhibits, enhancing their curiosity and

desire to learn. Depending on the various media and interaction methods, this has led to the creation of various interaction patterns, like human-machine interaction, human-environment interaction, and human-to-human interaction. Thanks to the development of multimedia technology, human-machine interaction is typical in contemporary museums. Visitors' role changes from that of a passive recipient of information to that of an active participant in the experience because of interactive technological innovations, which provide a new and exciting journey that sparks their curiosity and helps in the delivery of knowledge in creative ways[3].

The process of designing interaction in a museum environment can be defined as a design approach focusing on employing smart and interactive media to create dynamic connections between the visitor and the content. This provides a dynamic experience built on sensory, cognitive, temporal, and behavioral interaction that goes over the structural basis. By merging spatial, physical, and technological elements to create a dynamic, enjoyable, and pleasant interactive experience between the visitor and the museum environment, this fundamentally changes the nature of the museum itself from a display space to an experiential space, facilitating communication and knowledge acquisition for the user.

5. Dimensions of Interaction Design in the Museum Environment

The means which contemporary museums interact with visitors and show their displays are developing. Museums are currently advanced, interactive spaces meant to promote visitors' sensory, cognitive, and social interaction rather than just being spaces to keep and present objects. The dimensions of the user interface design process in museums include:

- **The Sensory Dimension**: Knowledge engagement with museums needs knowledge of the sensory dimension. It works on how to create an immersive and integrated experience by engaging human senses in the museum's architectural space. Here, architecture moves beyond its traditional function as a functional enclosure and becomes a sensory medium that uses light, sound, material, and smell to adapt to human emotions and cognitive needs [4]. The architect provides a whole experience that stimulates the senses, the mind, and the emotions instead of just an aesthetic form. The psychological and social design goals like comfort, peace, excitement, and attention all get served with the sensory dimension [5].
- **The cognitive dimension**: is an essential element of interaction analysis which deals with how someone uses their cognitive abilities, consciousness, and memory to view, understand, interpret, and interact with the architectural museum space. The senses alone are not complete the museum experience; in fact, it is shaped by complex mental processes such as symbolic association, mental interpretation, and mental recall [6]. By applying symbols, identifying functions, and recognizing spatial links, guests to architectural museums interact with the space by cognitive communication. The visitor's cognitive map is formed in large part by components that include form, organization, repetition, contrast, sequence, and orientation. Easy navigation and spatial awareness are created easier in this image. In addition, a logical or narrative sequence within the design (narrative architecture) enhances the user's cognitive connection to the museum space by helping them to interpret it as a lived story [7].
- **The technological dimension**: is one of the major factors pushing interactive elements in museum planning. It refers to the medium

making it accessible to the museum to respond to visitors in a lively and dynamic style. The museum spaces is no longer a static entity instead being an intelligent system capable of real-time interaction with human and contextual variables, thanks to the fast advances in artificial intelligence, the Internet of Things (IoT), and digital design [8]. Utilizing tools such sensors, actuators, and digital interfaces, which enable the building to monitor environmental and human data and respond correctly, the technological dimension helps turn museum architecture into a responsive, interactive space [9].

- **The Social Dimension**: A vital aspect of recognizing participation in a museum space is the social dimension. A museum's architectural space is viewed as a field for group participation, communication, and belonging, as well as a tool for developing social connections. Through the architecture of human encounter and interaction, the space aims to serve as a social platform that organizes and builds up social connections rather than just holding the individual.In this sense, the organization of public and private areas, the layout of pathways for movement, the amount of places of interaction, and the division of functions also serve roles in the interaction design of museum architecture, which creates patterns of collective behavior. Each of these elements expresses an architectural vision of social cohesion and enhances social interaction. [10].
- **The Spatial Dimension**: This complex psycho-social dimension organizes people's views, emotions, and behaviors toward the museum as an interactive setting. It develops an interactive relationship within the individual and the museum space across overlapping levels including the individual, the psychological process, and the spatial context by mixing physical components

of the surroundings with individual experiences and developed social meanings.This dimension provide a framework that explains how social, environmental, and physical elements—such as shared symbolism, environments, and architectural aspects—influence the connection to the place, enhance emotional experiences, establish meaning, and enhance links between visitors and locals within the museum space itself [11].

- **The temporal dimension**: This is the dynamic framework that describes how the relationship between humans and museums changes over time, through the succession of perceptual, sensory, and historical experiences that take shape while within the museum environment. This dimension reflects the continuous transformations in lighting, uses, and memory, contributing to the construction of an interactive experience that connects the present with the past and frames the visitor's perception of the museum space as a changing and evolving entity. The user's awareness of this dimension is strengthened the longer they stay in the space or the more frequently they experience it, as the senses—light, movement, and sound—contribute to shaping their understanding of time within the museum space. Integrating time into design is a fundamental value that enriches the work aesthetically and psychologically [12].

It becomes clear that the process of designing genuine architectural interaction within the museum environment stems from the dynamic integration of dimensions (sensory, perceptual, technological, temporal, spatial, and social) with human existence. This reshapes the architectural museum space and its surroundings, transforming architecture and its spaces into a living entity that keeps pace with changing human needs and establishes a continuous dialogue between the visitor, place, time, culture, environment, and technology, in a comprehensive architectural experience that fulfills the human aspect in all its dimensions.

6.Technological tools and technologies in contemporary museums

The digital transformation of museums has adopted various technologies, including **interactive displays** that allow visitors to interact directly with a variety of screens that capture their attention and enable them to engage with their content [13]. These screens are used to project sound and moving images onto floors, walls, and windows in archaeological museums, in addition to plasma screens, providing visitors with an advanced interactive experience [14]. **QR code technology** allows users to be redirected to specific applications or data when scanned by digital devicesTo enable to minimize the amount of explanatory speech in the museum and keep an engaging and attractive display, it is often added to artwork title plates to direct visitors to electronic resources containing thorough information about the works [15]. Another effective technology based on signals from devices operating in the radio band is Radio Frequency Identification (RFID) technology [16]. (Karimi *et al.*, 2012). Readers tend to be a member of these systems. To interact with visitors, RFID devices are located around the museum. After signing up, visitor profiles usually remain on wristbands, cards, or small discs as are handed out at the museum entrance. These devices keep visitors' personal data or preferences, which are then delivered to a tracking system that analyzes data on preferences and visit frequency [17]. Based on [18], sensors correctly monitor visitor numbers and track their movement during the exhibition with recording a number of variables, like pressure, light, temperature, sound, and movement. In addition, by allowing visitors to preserve a safe distance, these sensors serve establish an invisible barrier between the artwork and the visitor, enhancing the artwork's protection and regulating interaction behavior [19]. Compared to that, **hologram technology** employs a beam of light waves focused on the object to precisely show its three-

dimensional form and clarify its details [20]. The final image has a clean look, with its features floating and clearly visible. These can be viewed over the air without the use of specific tools or eyeglasses [21]. By adding a layer of meaning, **projection mapping**—which uses complex mapping techniques to project virtual elements onto real objects or enhances the realism of the elements and gives real things a more dynamic character [22]. **Digital sound systems** also focus on the auditory dimension of museum exhibits, presenting them in a musical, historical, or social style to activate the role of listening in conveying information [23]. They also contribute to maintaining the artistic accuracy of the works, whether to create a specific atmosphere or to convey information [18]. **Virtual reality technology** can be experienced using a head-mounted screen, allowing users to interact with and move within the environment. Combining computer graphics, sound, and other sensory effects, it provides visitors with an immersive and interactive experience that incorporates sound effects. Sensory experience enhances their understanding of the content and helps them interact with it [24]. **Augmented reality technology**, whose primary function is to instantly calculate the coordinates and angles of each informational element, enhances the human experience by expanding the perceptual, emotional, and cognitive dimensions of interaction with the physical world [25].

The growth of digital technologies has called awareness to the way technologies are transforming museum architecture to develop an immersive, interactive learning environment promoting interaction between visitors and objects. In addition, this facilitates the intelligent protection and management of collections, providing visitors with a personalized, multi-sensory experience and expanding the boundaries of the museum to encompass the virtual and digital worlds. This has transformed the museum into a living educational and cultural institution that transcends the limitations of space and time, as illustrated in Table 1, which demonstrates the impact and role of technological capabilities in museum architecture.

(Table 1) illustrates the impact of technological advancements on museum architecture.

Technology	Reflection on the museum	Area of application:	Impact on the visitor
Interactive displays	Making the display multimedia and flexible	Displaying information, images, and video	Direct participation and interactive learning
QR codes	Reducing reliance on traditional displays	Linking objects to digital content	Visitor's phone becomes a personal guide
RFID	Improving security and smart management	Tracking collections and visitors	Smoother movement and a more organized experience
Sensors	Protecting collections and enabling interactive responses	Configuring the interior environment and interaction	Motion sensing and a live experience
Holograms	Reviving lost or rare exhibits	Projecting 3D models	Live and impactful viewing of historical artifacts
Projection mapping	Employing architectural space as a dynamic display medium	Transforming walls into display screens	An immersive experience filled with visual storytelling
Digital audio systems	Enhancing the dramatic and narrative nature of the display	Producing immersive sound effects	sound perception and sensory experience
Virtual reality (VR)	Transcending the boundaries of space and time	Creating alternative environments	Entering virtual historical/cultural worlds
Augmented reality (AR)	Connecting the physical and digital in the display	Integrating digital information with objects	Direct and hybrid interaction with the artifacts

7 .Stages of Interaction Design in the Museum Environment

The museum encompasses both the personal and physical space of the museum itself. Museums typically receive a large number of visitors at once, representing diverse backgrounds and interests. While the movement of these visitors within the museum may be influenced by various factors, such as spatial organization and the size and distribution of exhibits, they often move around the

museum simultaneously. Tools such as maps, guides, and suggested route plans are essential elements that support the physical movement and exploration of a space, and simultaneously form an integral part of the overall visitor experience [26]. Visitors in museums interact with a variety of physical elements, whether original artifacts, props, or scenographic components. Context is inseparable from the experience; rather, it is an essential part of the visitor's perception and interaction with the exhibit, which has undergone a dematerial transformation through the integration of digital space. This digital space is the primary means of disseminating knowledge available in museums. These are the technological aspects associated with the physical spaces of the exhibition [27]., and their impact on paving the way for the integration of digital technology in museum exhibition design. Consequently, a qualitative shift occurs towards focusing on enhancing the visitor experience through the integration of virtual spaces. This results in the hybrid space (the interactive space), which refers to a space that incorporates tangible interactions as a result of seamlessly integrating technology within the physical context of the museum. This space is generated through an iterative design process that included field studies with Visitors benefit from the integration of technology with artifacts, which can be both effective and engaging. This integration enhances visitors' understanding of the historical context and meaning behind the exhibits [28]. Museums have become multi-sensory through technological integration, designing effective hybrid experiences that combine realistic and digital sensory elements to convey cultural heritage in an innovative and engaging way. The modern visitor seeks not only to access the displayed content but also to engage in a meaningful social and personal experience within the social space generated by interaction. These interactions provide a fertile ground for learning, social experiences, and situations, offering ample opportunities to study social interactions in diverse contexts.

It is clear from the above that the process of designing interaction in a museum environment includes stages that begin with analyzing the goals, requirements, preferences, movements, behaviors, and personal moods of visitors. The second step focuses on studying the space and the integrated technological capabilities that enable the design of the interactive scenario. This includes determining the type of interaction within the visitor›s journey to integrate museum exhibit elements and visual narrative into the visitor experience. The relationship between the visitor›s movement and the space›s response is carefully controlled. Thus, the type of interaction achieved changes the space, and so on. Each time the interaction changes, the space changes in a dynamic and continuous process centered on the visitor due to the ongoing programming of the content. This process culminates in the design of interactive systems and interfaces, integrated according to movement paths and diagrams.

8. Analytical Study

8.1 .Analysis of Specialized Studies in Interaction Design in Museum Environments

Previous studies specializing in interaction design in museum environments will be analyzed in order to derive the theoretical framework for the interaction design process in museum environments:

8.1.1 Study of (Vayanou, *et al.*, 2014) [29].

Personalizing interactive digital storytelling in archaeological museums: the CHESS project

The study highlighted the importance of achieving social interaction in the museum environment today through technological integration within the museum

environment. It presented what it called interactive digital storytelling, which supports the goals of interaction design by enhancing experiential learning and developing a sense of exploration and appreciation among visitors. The study clarified the aspects of integrating digital technologies and tools within the museum environment to enhance the physical and sensory experience of visitors, which it called (embodied interaction) through augmented reality technologies that integrate the physical and the digital within an adaptive storytelling framework capable of changing in real time with changes in the contextual environment, including the visitor's location, interests, preferences, and speed of movement and response. This is achieved by conducting interactive experiences that align with the principles of sensory and perceptual design. The multimedia and cross-media approach was used by combining (images, sound, and video) to create a film that evokes a specific emotion, in a new experience that engages the visitor in the interactive process. The study also clarified methods for identifying experiences by creating a digital file of visitor interactions, which can then be modeled to serve as design guidelines.

8.1.2 Study of (Garcia Carrizosa, *et al.*, 2021) [30].

"**Designing Technologies for Museums: Accessibility and Participation Issues**"

This study addressed how technology design impacts individual activity in museums, particularly for individuals with sensory impairments or learning difficulties. It explored how technology contributes to facilitating access to the physical space of the museum, enhancing the sensory experience of exhibits, and activating the spatial principle through the design of flexible spaces that accommodate visitor movement. Furthermore, it examined how technology conveys cultural knowledge through texts, audio-visual content, and symbolic

graphics to activate the perceptual principle. In addition, researchers strive to create museums that cater to everyone, in line with the principles of inclusive design. Providing personalized and tailored access for each individual enhances the sense of individual interaction and reflects an appreciation for the nature of each place and visitor. The concept of emotional accessibility transcends the physical and cognitive aspects to encompass the emotional needs of visitors. Participatory engagement of users in the actual design phases of the museum strengthens their emotional connection to it. Participants indicated that the opportunity to participate was enjoyable and beneficial, enhancing their emotional connection and meaningful interaction with the museum, its exhibits, and cultural activities. Thus, the concept of participation shifts from mere individual inclusion to a comprehensive engagement of the entire system.

8.1.3 Study of (Kaplun, *et al.*, 2023) [31].

Application of immersive technology in a museum

This study discussed the application of immersive technologies in museums, such as augmented reality and virtual reality, to create and generate customized interactive simulation environments. (p. 385) Museums are among the most important sources of knowledge related to the history and culture of regions, particularly in promoting cultural identity. To increase the competitiveness of these institutions, especially to attract the contemporary youth generation, it has become essential to adopt modern communication platforms that align with the nature of the contemporary digital audience. p386 Museum professionals and virtual reality technology developers should pay particular attention to the educational and entertaining aspects of the virtual experience. These factors contribute to enhancing visitor engagement within the digital environment, providing them with a more

interactive and stimulating learning experience. The ability to analyze visitor interaction with exhibits increases, allowing for the identification of varying levels of interest. This leads to it attainable to dynamically reorganize exhibits within the museum based on the level of participation, transferring less captivating pieces to more crowded areas. By actively involving the visitor in the experience rather than quietly observing, all of this increases the visitor's experience and improves emotional and cognitive engagement. The importance of integrated sensory interaction and the engagement of the senses—sight, hearing, and body motion—as well as how this improves cognition and cognitive engagement were also highlighted in this study.

8.1.4 Study of (Wen & Ma, 2024) [32].

Enhancing Museum Experience Through Deep Learning and Multimedia Technology

The paper shows how fast artificial intelligence technologies are evolving, which has made a major effect on the museum field's movement to smart solutions to enhance guest experiences. By merging deep learning algorithms with multimedia technologies, these solutions aim to enhance the interactive environment of contemporary museums. These technologies, which are considered as the basis of current technology, are now necessary for optimizing the experience of museum guests, especially when they are effectively integrated with displays. These technologies, that enables the delivery of accurate and rich information to visitors, include speech recognition, image recognition, and natural language processing. This happens by offering correct and specific explanations of the given content. In addition, gesture, speaking, and touch-sensitive interaction technologies have particular chances to improve visitors' understanding of cultural heritage. p1–p2 The study additionally generated methods for using multimedia, 3D modeling, and

interactive design to transform the museum space into a virtual one. This made it possible for visitors to virtually explore the museum, circumventing temporal and spatial constraints while preserving a rich and realistic interactive experience. P3. The old museum visit paradigm, which was primarily restricted to passively viewing exhibits, was drastically altered by the incorporation of multimedia technology, which resulted in a richer and more effective interactive experience in knowledge transfer. So as to present more immersive and enjoyable experiences, museums have adopted a number of advanced interactive technologies, including augmented reality (AR) and virtual reality (VR).

9.Extracting the vocabulary of the theoretical framework for the interaction design process in the museum environment.

Based on the analysis of previous specialized studies, a theoretical framework was developed for levels and indicators of interactive design (as shown in Table 2), which serves as a first step towards proposing a plan for designing interaction in the Iraqi Museum environment in order to transform it into a vibrant interactive environment.

Table (2) illustrates the theoretical framework for the transition to an interactive museum environment (interaction design plan).(Prepared by the two researchers)

Main vocabulary	Secondary vocabulary	Values
Objectives of Interaction Design	- To promote experiential learning and cognitive acquisition - To stimulate sensory, emotional, and contextual memory - To stimulate sensory, emotional, and contextual memory	
Dimensions of interaction design	Sensory Dimension	-Control using digital sound systems -Control of light and visual displays -Integration and control of scents to evoke and enhance visitor emotions - Tactile enhancement
	cognitive dimension	-Directing the visitor's attention to key points (time, place, environment, society, culture). -Combining text, images, video, and models to enhance multidimensional understanding. - Organizing the presentation chronologically.
	Spatial dimension	-Designing a dynamic, interactive journey (adopting a kinetic sequence that guides the visitor through progressively more engaging and informative displays) -Diversifying spatial scales (combining expansive spaces for group experiences with more intimate, focused spaces for individual or technology-driven experiences) -Integrating stimulating transition zones (creating interstitial spaces where the rhythms of light, sound, or materials change to prepare the visitor for the next experience) -Flexibility and adaptability (designing display units and routes that can be modified to accommodate temporary exhibitions or future development) - Providing points for social interaction and activities
	Temporal Dimension	-Temporal rhythm (distributing visual and interactive events along an ascending or gradual curve) - Linking exhibits to a temporal context or a storyline that mimics narrative.
	Responsive Dimension	-Designing stimuli (visual or auditory effects that encourage approach and interaction) to elicit responses tailored to visitor groups (age – people with disabilities). - Enhancing emotional response (integrating human stories or impactful dramatizations).
	Technological Dimension	-Immersive experiences (VR and AR technologies). - Direct interactive technologies: multi-point touchscreens, motion sensors and body tracking systems, sensory technologies, guided audio systems, digital lighting effects, mobile applications to connect visitors to online content or post-visit content, QR codes and optical character recognition systems for exhibits, and visitor data analytics.
	Social Dimension	-Utilizing collaborative mobile applications for content sharing. -Creating rest areas to increase opportunities for intercultural dialogue and interaction. And establishing activities or experiences that simulate different environments and cultures, encouraging visitors to exchange impressions. - Integrating live activities (shows or guided activities that encourage direct audience participation)

Main vocabulary	Secondary vocabulary	Values
Interaction design stages	Analysis phase	-Visitor analysis: Visitor segments (age, culture, digital behavior), disability needs, learning expectations. -Content analysis: Classification of exhibits (informative, sensory, narrative). - Objectives analysis: Learning, emotional engagement, increased visitor frequency, digital participation
	Conceptual design phase	-Identifying interactive stations of various types (touch, motion, sound, AR, group experiences). -Selecting appropriate technologies for the content and its presentation. - Arranging stations according to a spatial-temporal rhythm.
	Detailed design and modeling phase	-Designing pathways and spaces: movement flowcharts, rest areas, spatial scales. - Designing interaction interfaces: screen prototypes, activity maps, response design: reaction rules (e.g., when element A is touched, box B lights up and sound C is emitted).
Evaluating the interactive museum environment model	Sensory experiences	-Speed of visitor stimulation (visual, auditory, olfactory, tactile) - Strength of stimulation and range of attention
	Perceptual experiences	-Speed of comprehension of content and level of focus on exhibits - Level of information recall after the visit
	Spatial experiences	-Clarity and fluidity of movement and self-direction for the visitor --Diversity of the space (individual – group) - The degree of spatial coherence in recalling stations and its connection to the contex
	Temporal experiences	-Duration of the visit and desire to continue it - Recall of experiences after the visit
	Response experiences	-The achieved reaction (motor, mental, emotional, social) - The type of response (exploratory, educational, imaginative, realistic, cultural, aesthetic, personal)
	Technological experiences	-Ease of use and clear user interface. - The technology's suitability to the content and its avoidance of overwhelming it.
	Social experiences	--The duration of interaction between visitors within the space. - The degree of social integration and the formation of groups that share activities or collective experiences within the museum environment.

10. Application to the Iraqi Museum

The research relied on applying the theoretical framework by activating the seven dimensions through the design of interactive stations in the Great Assyrian Hall, which is considered one of the most prominent halls in the museum as it contains artifacts from the Neo-Assyrian civilization between (609-911 BC). As Figure (1) and (2 a, b, c) illustrates, although the hall has undergone radical modifications, the interactive experiences have not yet reached the desired level. Thus, the design of each interactive station was linked to an operational technology that activates the diminished interactive dimension, as shown in (Figure 3). Furthermore, the relationship between the designed station and its technologies and the visitor group (gender, age, cultural level) was determined, as follows:

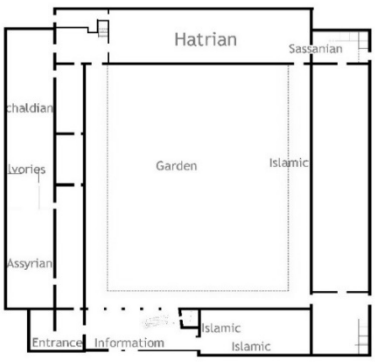


Figure (1) shows Current ground floor plan of the Iraqi Museum

Source: Iraqi Department of Antiquities and Heritage

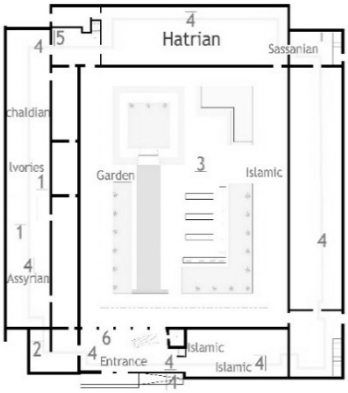


Figure (3) shows the layout of the Iraqi Museum including the interactive stations

Source: (prepared by the researcher)

Figure (2) shows the current shape of the Assyrian hall after its restoration, like the Assyrian palaces

https://www.france24.com/ar/20161101 :Source

1-**Interactive Storytelling Station**: Interactive storytelling stations are an essential part of architectural design within museums. They are used to present artifacts and archaeological information in an interactive way that integrates architectural space with modern technology, as shown in Figure (4a and 4b). Through turning information into an extensive sensory experience, these stations attempt to explain and simplify historical concepts. They provide more functions beyond merely displaying content; they also create an architectural context that draws in attention and stimulates the imagination. The architectural design of these stations takes the form of expressive spaces that utilize appropriate dimensions and materials to enhance the sensory connection between the visitor and the displayed content. Instead of simply piling up texts and explanatory panels, these stations are integrated into the architectural space in a way that reflects the cultural and historical context of the exhibits.

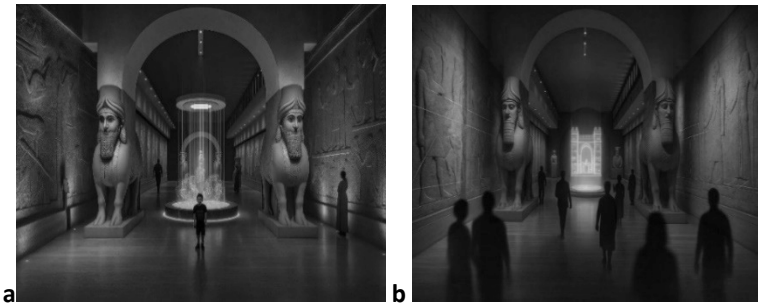


Figure (4) illustrates the narrative stage and the use of visual stimuli. (Source: Researcher)

2-**Audio Interaction Station**: The audio interaction station is part of the architectural design where technology integrates with the surrounding environment to enhance the visitor experience. This station includes an interactive device or digital units, as shown in Figure (5a, b, c), but the architectural space itself is designed to reflect the historical and intellectual character of the audio content presented. Sound is used to connect the visitor to a specific historical period, such as translating everyday vocabulary from Assyrian civilization or presenting traditional songs and music from that era. Rather than being a separate digital device, this station is part of an architectural space integrated into the museum environment.

Figure (5) The source audio interaction station (Source: Researcher)

3 -**The Interactive Children›s Station**: This station changes the relationship between children and the space, indicating an important turning point in the typical museum experience. So as to promote active exploration and learning, a dedicated children's garden was established within the museum. Children can use planting and watering using simple, safe tools thanks to the garden's design, which imitates agricultural practices in ancient Iraq. Touch, observation, and repetition are all enhanced by this sensory experience.

In addition, the children's museum has a symbolic architectural entrance created up of replicas of iconic architectural symbols like ziggurats and historical gateways like the Assyrian Nergal Gate. Reconstructed to a suitable for children

scale, these models indicate the beginning of the station's exploring journey. As seen in Figure (3), the children's station transforms from a static space into a hybrid architectural context where play and learning interact with architectural form. As a result of this, children's museum visits are comprehensive and dynamic, employing architecture than just words to support cultural and cognitive values.

4-**Special Needs Station**:In an effort to accomplish cultural respect and promote community engagement, the design aims to bring people with disabilities into the architectural space of the museum. Accordingly, the museum is designed to be integrated within the main museum pathway, becoming a natural extension of the overall experience. Gentle ramps, as shown in the museum's layout in Figure 9, are provided to ensure accessibility for people with mobility impairments. Rest areas, non-slip flooring, and ample circular spaces allow for unobstructed wheelchair movement. Exhibits are also rearranged at heights suitable for both visual and tactile access, fostering a sense of independence and complete inclusion for visitors. The design incorporates bilingual elements, using large, high-contrast print for the visually impaired, along with Braille translations and tactile tiles, as illustrated in Figures (6a and 6b), making the museum space readable at all levels. These components represent a vital element of the museum's identity because they are integrated into the architectural space instead of merely external additions.

Figure (6) illustrates tactile tiles for the blind as a design measure for pathways for people with special needs. (Source: Researcher)

5-**Virtual Reality and Augmented Reality Station**: The architectural design of this station is based on the principle of respecting the original archaeological space, integrating technology into the setting. The station utilizes all existing architectural elements, from columns with sculpted capitals to arches that formed part of the Assyrian religious and political landscape, and winged bulls that symbolized power and protection, transforming these elements into narrative tools and part of the show's scenography, rather than mere silent decoration. In the Assyrian hall, where the walls acquire a distinctive archaeological character, the virtual and augmented reality station is designed as a hybrid interactive experience that revives Assyrian history through multimedia, without excluding or negating the space, but rather building upon it. Digital interaction stations, represented by VR points or AR panels, are distributed within the architectural spaces, integrated into the floors, thus mitigating the dominance of the modern and technological form over the heritage scene, as shown in Figure (7).

Figure (7) illustrates the virtual reality and augmented reality station (Source: Researcher)

6-**The Participatory Visitor Station**: At the final stage of the interactive experience, an interactive space emerges within the museum, designed to give visitors the opportunity to express their opinions, share ideas, and exchange personal impressions about the museum experience. It does not merely provide

information but also provides feedback between the visitor and the museum, thus making the visitor an active participant in producing meaning and developing the social and participatory principle. The design takes place within one of the natural open architectural spaces in the museum building, such as between two halls or at the end of one of the naturally lit corridors that open onto the internal gardens, as shown in the museum plan in Figure (8). The reason for choosing this location is to attract the visitor to an area of tranquility and contemplation away from the noise of visitors, to be a point of reflection where the visitor can express his opinions and feelings.

Figure (8) illustrated The enhanced participatory visitor station is (Source: Researcher)

12. Survey Conducting

After completing the interactive design plans for the Iraqi Museum's environment, an online survey was conducted. This survey aimed to gauge public attitudes toward digital interaction and their future desire to implement the theoretical framework's components within the planned museum environment. Despite the limited sample size, the survey was directed to a select group of 25 academic architects specializing in digital and cultural architecture, thus enhancing the reliability of their technical and critical opinions on the proposed theoretical framework. Two groups of individuals were examined. This process was to examine

architectural viewpoints and interpretations of the aspects of interaction and function within the interaction design process of the museum spaces, rather than evaluate direct technical performance. The survey was supported with a link to the theoretical framework via Google Forms and a demonstration video of the Iraqi Museum. The results were presented in statistical tables, ensuring that the suggested solutions not only handle the academic flaws observed in the museum survey but also match with actual needs and aims of general viewers.

13. Interpretation of Results

Figure (9) shows the interpretation of information related to the experiences obtained through the interactive design of the Iraqi Museum's space. Based to the survey, 92.33% of the participants had technological experiences, showing how effectively the museum's design for buildings employed technology. The percentage of sensory experiences obtained improved to 89.45% because technology is believed to be an aid for enhancing interpretation and perception. This shows the design's potential to provide a space. filled with visual and audio stimuli that can attract visitors, and encourage their immersion and engagement with the displayed content, with a response rate of 87.23% in response to the stimuli that stimulate interaction. Additionally, the perceptual experience earned 86.55%, showing that the exhibits and media can guide visitors and help them collect information and comprehend its meaning. The results of the temporal experiences, which reached 88.65% resulting in a clear and consistent narrative organization, further illustrate the impact of technological experiences on interpreting the sequence of screens. The hall's spatial experience crystallizes with perceptual and temporal preparation, achieving 90.24%, indicating clarity. Using a score of 85.88, the smooth flow of movement for all museum visitor groups creates a sense of spatial connection, which in turn enhances the social

experience. This connection is crucial in activating the social space, making the experience interactive and participatory. These results combine to reveal an interactive environment that fulfills the objectives of the interaction design within the Iraqi Museum's setting. This environment supports experiential learning and knowledge acquisition by stimulating the visitor's sensory, cognitive, and emotional memory, thus creating a lasting relationship between the visitor and the museum.

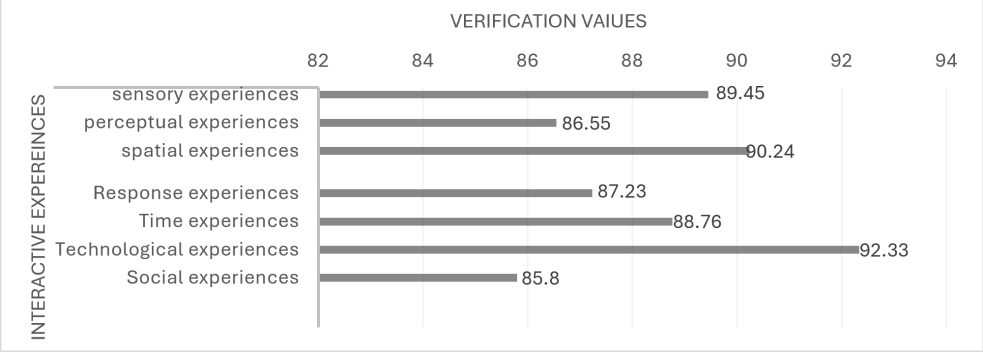


Figure (9) illustrates the success rates of interactive experiences in the Iraqi Museum environment.

14. Conclusions

1. Adopting the concept of interaction design in the museum environment is a fundamental pillar in advancing the contemporary muse um experience. Its role is evident in transforming the museum from a space dedicated solely to exhibition into an interactive, educational, and participatory environment that connects the visitor to the cultural and civilizational heritage
2. The research highlighted the importance of studying spatial organization and structuring museum spaces by organizing, clarifying, and smoothly guiding movement paths to create a comfortable museum environment

that contributes to enhancing the quality of the museum as a dynamic architectural space.

3. The integration of cognitive and sensory stimuli contributes to attracting visitors and raising their level of understanding of the content. This, in turn, underscores the role of thoughtful planning in activating the role of technological and sensory influences to support the cognitive process.
4. The research confirmed that the technical and technological role is one of the strongest drivers and supporters of interaction, as it contributes to making the experience more dynamic and more connected to its historical context.
5. The results of this research underlined how important temporal consideration is in creating an effective and powerful museum experience. The displays' chronological organization is vital for enhancing visitors' knowledge of historical content and allowing them to observe how events through following civilizations are interconnected.
6. The study highlighted how essential the social aspect is to the museum experience and that integrating interactive spaces (interactive stations) may enhance it by promoting social interaction that improves the museum experience.
7. A holistic design vision that is taken into consideration interaction in all of its dimensions—sensory, perceptual, temporal, spatial, technological, and social—is deemed essential to create a comprehensive museum experience. This sets the visitor at the core of the design process and enhances their role in participation and exploration.
8. The availability of powerful, knowledge-based physical content elements is vital to a public museum's performance. These elements can act as base and methodology for interaction design, which will ultimately result in a more interactive and engaging experience.

15. Recommendations

1. Adopt methods of design that include a comprehensive organizational plan to enhance the quality of the Iraqi Museum's visitor experience.
2. To enhance responsiveness and interaction, engage the sensory and technological elements of the interactive design structure.
3. Having incorporated interactive, multisensory stations presenting Iraqi civilization-associated subjects.
4. developing the Iraqi Museum as a platform for culture and learning by providing interactive materials for learning to improve engagement and promote a feeling of community.
5. improving the whole museum experience, enhancing the scope of cultural activities presented within the museum, improving visitor services, and strengthening a relationship between the museum and its visitors through smart interactive apps.
6. changing the museum experience from a basic display of objects to an integrated educational, knowledge-based, cultural, and sensory interaction that transforms the museum into a narrative of civilization.

16. References

[1]. Shibli, M. A., Mohsen, A. N., & Almaliky, H. N, (2019), *Evaluating the economic and social returns of museums' performance: An exploratory study of a sample of Iraqi museums in Erbil, Baghdad, Babylon, and Karbala*. Southern Technical University, University of Oil and Gas.

[2]. Ma, Y, (2018), A preliminary study about the "storytelling" method in museum exhibition Journal of Natural Science Museum Research, 23-29. https://doi.org/10.19628/j.cnki.jnsmr 2018.02.004.

[3]. Tian, Xi, (2022), Research on the Application of Interaction Design in Museum Exhibition. Journal of Global Humanities and Social Sciences. 3. 28-29. 10.47852/bonviewGHSS2022030206

[4]. Jablonska, J.; Leszczyńska, E.T.; Tarczewski, R, (2015), Sound and architecture—Mutual influence. *Energy Procedia* , *78*, 31–36

[5]. Spence, Charles, (2020), Senses of place: architectural design for the multisensory mind. Cognitive research: principles and implications. 5. 46. 10.1186/s41235-020-00243-4.

[6]. Psarra, Sophia,(2009), Architecture and Narrative The formation of space and cultural meaning . Routledge.

[7]. Kang, Xue, Gabriella Medvegy, and Yufang Zhou.(2022, March). "Spatial narrative research with architecture as the media." *Pollack Periodica* 17, no. 1, 173–177. DOI:10.1556/606.2021.00398.

[8]. Sourek, M, (2023), *Unleashing the potential of artificial intelligence in architecture and development of the built environment: From misconceptions to productive prospects*. Global Journals of Research in Engineering, 23(J3), 5–31. https://doi.org/10.1007/s44223-023-00046-9

[9]. Al-Sharrah, Hanadi Yousef ,(2021), Advanced Technology and its Impact on More Interactive Residential Interior Architecture, Scientific Journal Research in Science and Arts, Issue Fifteen, Volume One.

[10]. Qi, J., Mazumdar, S. & Vasconcelos, A.C, (2024), Understanding the Relationship between Urban Public Space and Social Cohesion: A Systematic Review. *Int. Journal of Com. WB* 7, 155–212 https://doi.org/10.1007/s42413-024-00204-5.

[11]. Zeng, Yijun & Deal, Brian, (2023). What Role Do Urban Parks Play in Forming a Sense of Place? Lessons for Geodesign Using Social Media. Land. 12. 1960. 10.3390/land12111960.

[12]. Al-Sharkawy, Dalia Ahmed, Abbas, Narmin Saeed, Wahba, Soti Ehab,(2024), Modern Trends in Conceptual Art as a Postmodern Art Movement and its Role in Beautifying Tourist Architecture, Journal of Architecture, Arts and Humanities, Volume 9, Issue 45, DOI: 10.21608/MJAF.2022.125069.2678.

[13]. Ahmed, Heba Shaaban Abdel Moneim Mohamed, Qutb, Maysoun Mohamed, & Hassan, Enas Mahmoud Mohamed. (2018, July). Employing Interactivity in Museum Display Design. Journal of Architecture, Arts and Humanities, 3(11), 626–649.

[14]. Hosny, Narmin Ali,(2018), A Design Strategy to Activate the Role of Interactive Advertising in Enriching the Interior Environment of Restaurants. Journal of Arts and Applied Sciences, 5(3), 54–72. https://doi.org/10.21608/maut.2018.105561.

[15]. Karayilanoglu, G, (2019), Digital transformation in contemporary art museums with the example of Barcelona Museum of Contemporary Art. In *Proceedings of the 5th International Conference on New Trends in Architecture and Interior Design*.p(16-20) Mimar Sinan Fine Arts University.mimarsinan.academia.edu.

[16]. Karimi, R., Nanopoulos, A., & Schmidt-Thieme, L, (2012), RFID-Enhanced Museum for Interactive Experience. In C. Grana & R. Cucchiara (Eds.), *Multimedia for Cultural Heritage* (pp. 157–168). Springer. https://doi.org/10.1007/978-3-642-27978-2_17.

[17]. Wang, Y., Zhang, Y., & Zhang, L, (2024), Developing a digital management system for museum collections using RFID and enhanced GIS technology. *PeerJ Computer Science*, *10*, e2462. https://doi.org/10.7717/peerj-cs.2462.

[18]. Karayılanoğlu, G., & Arabacıoğlu, B. C, (2020),Digital interactive experiences in contemporary art museums. *The Turkish Online Journal of Design, Art and Communication (TOJDAC)*, 10(4), 423–440. https://doi.org/10.7456/11004100/007.

[19]. VdS. (2008, September). *Security Guidelines for Museums and Showrooms* (VdS 3511en:200809 (01)). VdS Schadenverhütung GmbH.

[20]. Omar, Heba Omar,(2024), Utilizing Holograms in Interactive Interior Design, Heritage and Design Magazine, Volume 4, Special Issue (1).

[21]. Yu, D., & Yao, W, (2023), Research on holographic display and technology application of art museum based on immersive design. *Journal of Physics: Conference Series, 2425*(1), 012048. https://doi.org/10.1088/1742-6596/2425/1/012048.

[22]. Milgram, P., & Kishino, F, (1994),A taxonomy of mixed reality visual displays. *IEICE Transactions on Information and Systems*, E77-D(12), 1321–1329.ResearchGate+12Cumincad+12OUCI+12.

[23]. Mathieu, M. (2023, May).When museums increase listening: sound creation in the service of transmission. Listening to today's music, Nova Contemporary Music MeetingLisbonne, Portugal. hal04221833.

[24]. Wojciechowski, R., Walczak, K., White, M., & Cellary, W, (2004), Building Virtual and Augmented Reality Museum Exhibitions. *Proceedings of the Ninth International Conference on 3D Web Technology (Web3D '04)*, 135–144. https://doi.org/10.1145/985040.985060.:contentReference[oaicite:5]{index=5.

[25]. R. Doerner, W. Broll, B. Jung, P. Grimm, M. G¨obel, R. Kruse,(2022), Introduction to virtual and augmented reality. Virtual and Augmented Reality (VR/AR) Foundations and Methods of Extended Realities (XR), Springer International Publishing, Cham, pp. 1–37.

[26]. Ciolfi, Luigina & Bannon, Liam, (2007), Designing hybrid places: Merging interaction design, ubiquitous technologies and geographies of the museum space. Codesign. 3. 159-180. 10.1080/15710880701524559.

[27]. Hatala, M., Kalantari, L., Wakkary, R., & Newby, K, (2004), Ontology and Rule Based Retrieval of Sound Objects in Augmented Audio Reality System for Museum Visitors. Proceedings of the 2004 ACM symposium on Applied computing pp. 1045–1050.

[28]. Schou, M. M., & Løvlie, A. S, (2021), *The Diary of Niels: Affective engagement through tangible interaction with museum artifacts*. In M. Ioannides (Ed.), Digital Heritage. Progress in Cultural Heritage: Documentation, Preservation, and Protection (pp. 289–299). Lecture Notes in Computer Science. https://doi.org/10.1007/978-3-030-73043-7_24

[29]. Vayanou, M., Katifori, A., Chrysanthi, A., Vourliotis, P., Karvounis, M., Kourtis, V., & Paliokas, I, (2014), Personalizing interactive digital storytelling in archaeological museums: The CHESS project. *Proceedings of the 2014 ACM conference on Designing interactive systems (DIS '14)*, 511–520. https://doi.org/10.1145/2598510.2598562

[30]. Garcia Carrizosa, A., Valdemarin, C., Waddington, D., & Muresan, L, (2021), Designing technologies for museums: accessibility and participation issues.

[31]. Kaplun, D., Romanov, S., Ipalakova, M., Daineko, Y., Bolatov, Z., & Tsoy, D, (2023), *Application of immersive technology in a museum*. *Journal of Cultural Heritage and Technology*,) 385–390.

[32]. Wen, J., & Ma, B, (2024), *Enhancing museum experience through deep learning and multimedia technology*. Heliyon, 10(6), e24837. https://doi.org/10.1016/j.heliyon.2024.e24837

Solar-Assisted Thermal Integration for Improving Fuel Pumping Efficiency in the Middle Oil Distribution Company – Karbala Depot / Original article

Researcher Ali Shakir Hial

Islamic Azad University, Islamic Republic of Iran
alishakirhial@gmail.com
07718180193
Corresponding author: alishakirhial@gmail.com

الدمج الحراري المساعد بالطاقة الشمسية لتحسين كفاءة ضخ الوقود في شركة توزيع نفط الوسط ـ مستودع كربلاء

الباحث علي شاكر هيال
جامعة ازاد الاسلامية \ جمهورية ايران الاسلامية

Abstract

This paper discusses how a solar-assisted thermal heating system can be incorporated to enhance the efficiency of fuel pumping at the Karbala Fuel Depot which is managed by the Middle Oil Distribution Company in Iraq. The aim of the system is to partially pre-heat diesel before pipeline transfer and hence minimizing losses in pressure caused by viscosity and saving electrical power consumed during pumping operations. The test is a combination of field tests, thermal performance modeling, and thermal flow analysis using CFD to capture the interaction between temperature, fuel rheology and pump load behaviour.

It was proven that the stratified thermal storage tank (8000 liters) could hold steady temperature layering and allow early access to higher-temperature fuel in the morning loading phase when pumping demand is at the highest. The exchanger also enabled even distribution of temperature increase throughout the circulating loop and this reduced the pump power consumption by an estimated 12-15 percent in the critical operating range. Although annualized direct energy savings are low considering the existing Iraqi electricity tariffs, the system provided some operational advantages including smooth startup behavior, minimized mechanical stress on pumps and the minimized short-period flow discontinuities.

A discounted economic evaluation at a 7 percent rate revealed that complete capital remuneration is not reached in a 15 years' service horizon. Nevertheless, the usefulness of the system grows by taking into account the reliability of fuel handling, possible increase in tariffs, and strategic topicality of solar use in the industrial petroleum logistics. The paper also concludes that the solar-assisted thermal integration is a technically viable and operationally useful addition to the fuel distribution operations at Karbala Depot and can be used as a pattern by other such plants in Iraq.

Keywords: A Solar assisted heating, fuel pumping efficiency, diesel viscosity, thermal stratification, heat exchanger, energy performance, Karbala Fuel Depot, Iraq petroleum logistics.

المستخلص

يناقش هذا البحث كيفية دمج نظام تسخين حراري مساعد بالطاقة الشمسية بهدف تحسين كفاءة ضخ الوقود في مستودع كربلاء للوقود، الذي تديره شركة توزيع نفط الوسط في العراق. يهدف النظام إلى التسخين الجزئي لوقود الديزل قبل نقله عبر الأنابيب، مما يسهم في تقليل الفواقد في الضغط الناتجة عن اللزوجة، وتخفيض الطاقة الكهربائية المستهلكة أثناء عمليات الضخ. يعتمد التقييم في هذه الدراسة على مزيج من الاختبارات الحقلية، ونمذجة الأداء الحراري، وتحليل الجريان الحراري باستخدام تقنيات الديناميكا الحرارية الحاسوبية (CFD) لرصد التفاعل بين درجة الحرارة، والسلوك الريولوجي للوقود، وحمل المضخة التشغيلي.

وقد ثبت أن خزان التخزين الحراري الطبقي بسعة (8000 لتر) قادر على الحفاظ على تدرّج حراري مستقر، يسمح بالوصول المبكر إلى وقود بدرجة حرارة أعلى خلال الفترة الصباحية التي تكون فيها متطلبات الضخ في ذروتها. كما ساعد المبادل الحراري على تحقيق توزيع متجانس لارتفاع درجات الحرارة في حلقة التداول الحرارية، الأمر الذي خفض استهلاك قدرة المضخة بنسبة تتراوح بين (12%–15%) ضمن النطاق التشغيلي الحرج. ورغم أن الوفورات السنوية المباشرة في الطاقة تبقى محدودة في ظل تعرفة الكهرباء الحالية في العراق، فإن النظام حقق مزايا تشغيلية مهمة، منها: تحسين سلاسة تشغيل المضخات عند البدء، وتقليل الإجهاد الميكانيكي الواقع عليها، والحد من حالات الانقطاع القصيرة في الجريان.

أظهر التقييم الاقتصادي المخفّض باستخدام معدل خصم قدره (7%) أن الاسترداد الرأسمالي الكامل لا يتحقق خلال أفق زمني يمتد إلى (15) سنة من الخدمة. ومع ذلك، تزداد جدوى النظام عند الأخذ بنظر الاعتبار موثوقية مناولة الوقود، وإمكانية ارتفاع التعرفة مستقبلًا، والأهمية الاستراتيجية لاستخدام الطاقة الشمسية في لوجستيات الصناعة النفطية. وتخلص الدراسة إلى أن دمج النظام الحراري المساعد بالطاقة الشمسية يمثل خيارًا تقنيًا قابلًا للتطبيق وذا منفعة تشغيلية لعمليات توزيع الوقود في مستودع كربلاء، ويمكن أن يشكل نموذجًا يمكن اعتماده في منشآت مماثلة داخل العراق.

الكلمات المفتاحية: التسخين المساعد بالطاقة الشمسية، كفاءة ضخ الوقود، لزوجة الديزل، التخزين الحراري الطبقي، المبادل الحراري، الأداء الطاقي، مستودع كربلاء، لوجستيات الصناعة النفطية في العراق.

1. Introduction

1.1 General Background

Over the past few decades, the ever-increasing demands of energy the world over have exerted an increasing weight on industrial processes as well as the national power structures. The oil distribution industry is one of the most energy-consuming industries in the industry, and in developing countries, the energy converting efficiency is rather low. Iraq, blessed with rich oil deposits, is ironical in that it is always facing problems that revolve around poor management of energy used in its oil plants. Middle Oil Distribution Company (MODC) is one of the most important departments of the Iraqi Ministry of oil and in charge of some major storage depots such as the Karbala Fuel Depot which is a strategic centre that delivers bulk quantities of refined petroleum products to central and southern provinces. These plants rely on traditional electrical and thermal systems to power pumps, hold temperatures steady in storage tanks and to keep the viscosity of different fuel types- particularly diesel and heavy oil at a steady state during colder weather. [1].

The unpredictability and fluctuation of the national electrical grid and the increasing expenses of the diesel-driven backup generators in the recent years have compelled the researchers and engineers to look at other sources of energy that are renewable and more stable. Solar energy specifically has a tremendous potential in Iraq owing to its great solar radiation levels which surpasses 5.5 kWh/m2 per day on average with Karbala Province being rated as one of the highest irradiance areas in the country. The combination of solar-assisted thermal systems in oil distribution centers may alter the existing operational paradigm, as the part of the fossil-based power will be replaced with renewable and clean heat, which will be capable of maintaining the pumping lines, storage tanks, and even auxiliary buildings. Research in other areas has already established that solar thermal collectors with evacuated

tubes or parabolic trough could actually result in energy savings of between 20 to 45 percent in such a similar industrial process [2].

Nevertheless, although the petroleum industry of Iraq has increasingly implemented monitoring and automation technologies, there is little practical implementation of renewable energy and, in particular, the use of solar thermal systems. The majority of the local facilities are still using the old-fashioned heating systems, which are driven by fossil fuels, which result in excessive consummation of energy, avoidable emissions, and the raise of operation expenses. Thorough and situation-based analysis is thus necessary to discuss how the solar-thermal hybrid systems can optimize the thermodynamic performance of these processes and their sustainability without affecting their safety and reliability.

1.2 Problem Statement

Although the Iraqi oil distribution depots have made considerable investment in infrastructure, the energy management systems still experience inefficiency and low levels of flexibility. The Karbala Fuel Depot that is used as a sample of the Middle Oil Distribution Company is subjected to chronic problems associated with the need to rely on traditional approaches to heating to ensure the best storage conditions. At present, heaters and pumping processes of the depot are managed with the help of heaters that are diesel-powered and electrical resistance units, adding to excessive fuel usage and pollution of the environment. In addition, the vagaries in grid power availability are mostly causing unplanned downtime that interrupts the constant movement of fuel and creating the cost of maintenance that puts extra pressures on the operational budget of the company.

Lack of integration based on renewable also makes the system susceptible to the market fluctuations in the prices of fuel and also to external power alone

shortages that occasionally hit the mainland of Iraq. Although there have been pilot projects where photovoltaic systems have been investigated to provide lighting or small loads of electricity, little focus has been placed on solar thermal integration although it could potentially provide direct heat energy that can be used with the existing fluid distribution systems. Such gap in knowledge has been created by the absence of research and pilot testing in this field. In particular, modeling the thermo-mechanical behavior of such hybrid systems subjected to the conditions of the Iraqi climate is lacking or the quantitative assessment of their performance regarding energy savings, the homogeneity of heat distribution, and reduction in carbon emissions are not present.

This study is therefore triggered by the need to correct these operational inefficiencies in exploring the potential in introducing a solar-assisted thermal network in the available fuel distribution networks at the Karbala depot. The proposed solution may result in such intervention as it needs a solid analysis model backed by real-life data of the site to help determine whether the given solution may be implemented into the engineering and maintenance systems of the company.

1.3 Research Importance

This work is important as it has a quantitatively interdisciplinary impact on the fields of mechanical engineering (energy conversion and thermodynamics) and industrial petroleum operating. One, it offers a practical solution to the decreasing reliance on fossil fuel in one of the most vital logistical operations of Iraq. The research also facilitates the further cause of energy efficiency as well as environmental sustainability by proposing the integration of a solar-thermal subsystem with the existing heating system of the depot- this is in line with the national energy transition strategy towards the reduction of emissions and cleaner production in Iraq [3].

Engineering wise, the research will improve the insights into the heat transfer process and recovery of energy in large-scale liquid fuel systems. It looks at optimization of the thermal storage and distribution balance that can ensure the desired temperatures with minimum losses due to conduction and convection. Also, the analysis does not just work theoretically: it takes into consideration practical operational limits like ambient dust concentrations, time-varying solar radiation and accessibility of maintenance-processes, which are commonly ignored in pure theoretical simulation.

At the industrial and policy level, this study offers practical implications to the Middle Oil Distribution Company to make modernization in depot infrastructure and cost reduction of its operations. Since the model is scalable, the positive outcomes, achieved at the Karbala location, would be used as a reference in implementing the same at other branches (e.g., Najaf, Diwaniyah, and Baghdad-Rasafa). This would contribute to energy saving in the country and would potentially save the sector a number of several million liters of fuel every year. Moreover, the incorporation of renewable systems may boost the positions of Iraq in the regional energy diversification undertakings and increase the safety of number of people in their workplace by eliminating reliance on open-flame heating systems.

1.4 Hypothesis and Objectives of the Research.

The hypothesis used in developing this paper is that solar-assisted thermal integration would have a significant improvement in making oil distribution facilities in Iraq more efficient and eco-friendly. The main point is that it is possible to store thermal energy that is generated on the sun on daytime, and then use it to pre-heat or keep the temperature of pipelines and tanks when it is not needed to heat them with the help of the fossil-based heaters.

In this regard, the research will have the following specific objectives:

1. To construct a thermodynamic model of the state of energy balance in the Karbala climatic conditions solar-thermal-fuel distribution system.
2. To obtain the simulation of system behavior with using computational tools (ANSYS Fluent or MATLAB Simulink) to assess the temperature distribution, heat losses, and system efficiency.
3. To determine the environmental and financial gains, such as the decrease in the amount of CO2 emissions, decrease of the operating costs and frequency of maintenance.
4. To recommend a scalable integration framework that can be implemented in other depots that are operated by the Middle Oil Distribution Company.

The study will fill the gap by achieving these objectives: connecting the renewable energy technology to the traditional oil distribution infrastructure. The results of the research are likely to be used in further hybridization works in the Iraqi oil and gas industry that will illustrate how the principles of mechanical energy conversion and sustainable design can be combined and harmonized to form a more efficient and resilient industrial system.

2. Literature Review

2.1 International Literature on Solar Thermal Integration in Oil Systems.

The introduction of thermal conversion of the solar energy to the petroleum plants is a major change in the way industrial heat would be managed. Although oil-powered thermal systems have traditionally needed fossil fuels to meet thermal needs, recent research has proven that Solar Heat for Industrial Processes (SHIP) can effectively provide a significant fraction of thermal requirements in medium temperature processes (60-200 0C) used in preheating pipelines, to maintain fuel

viscosity and to prevent paraffin deposition in short-distance distribution systems. International Energy Agency (IEA) Task 49/IV pointed to the fact that the most successful integration results are achieved when solar fields are part of the entire energy balance of the site by way of pinch analysis and when they are coupled with short- and medium-term thermal storage systems [6 and 9].

In the world, there are many case studies that have confirmed the technical and economic viability of incorporating solar-assisted thermal systems in oil-related activities. Indicatively, massive deployments in both Oman and the United Arab Emirates have effectively employed evacuated-tube collectors and parabolic trough fields to generate process heat and steam to facilitate the process of enhanced oil recovery (EOR) activities [10 and 11]. These systems demonstrated a reduction of fossil fuel by a factor of 25-45, meaning that other downsizing configurations could be used at downstream distribution plants like depots and terminals. Iraq is a good target to use SHIP technology, due to its huge solar potential of over 5.5kWh/m2/day, which is high in the southern and central provinces such as Karbala, Najaf and Diwaniyah [13], [14]. Nevertheless, solar thermal penetration in the petroleum infrastructure of Iraq is very poor in terms of the world standards.

2.2 Pumping and Fuel Storage Systems Applications.

Fuel depots and pumping stations are continuously subject to thermal conditions to maintain the efficiency of fluid transportation and eliminate the solidification of heavy or mixed fuels. These requirements are in tandem with the operating range of solar thermal collectors. A number of studies have investigated preheating loops with short transfer pipelines and pump discharge lines with the use of compact plate or shell-and-tube heat exchangers. According to the reported findings, solar-assisted systems coupled with small-scale water-based thermal

storage can save up to 15-35 percent of energy [6 and 12].

Experiments carried out recently in conditions of the Middle East climate have supported the fact that evacuated-tube solar collectors can be used successfully even on dusty or half-cloudy days. [13]-[15]. In addition, engineering standards have suggested that solar thermal loops should be added parallel to the existing fossil-based system (instead of replacing it entirely) in order to ensure continuous operation. The hybrid design will enable an automatic switch to diesel-based heating in case of low-irradiance conditions and mitigate operational risks despite a high level of energy displacement.

To demonstrate how the proposed solar-assisted heating loop would be practically implemented in fuel depot, Figure 1 below shows a schematic illustration of the integration concept that will be implemented in Karbala Fuel Depot. This system integrates the field of solar collector, a plate heat exchanger, a thermal storage tank and the current pumping and fuel storage network, so that the system can guarantee the maximum transfer of heat and the minimum operational interference.

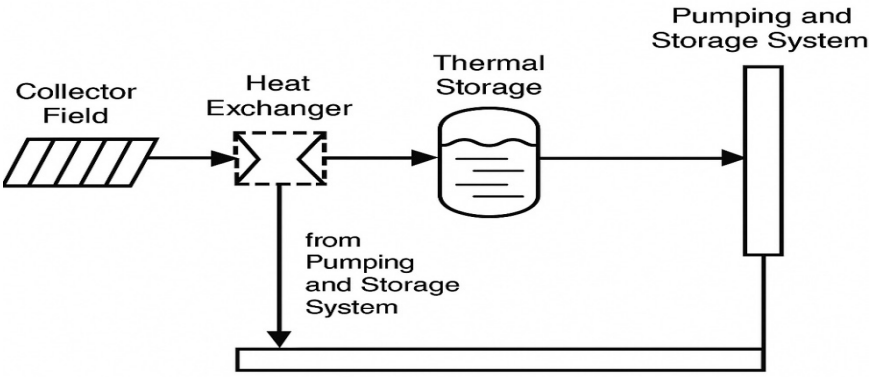


Figure 1. Solar-Assisted Thermal Loop Integration

System design In terms of system design, the point of integration should be as near as possible to the point of heat demand (such as the suction side or the discharge side of the pump line) so as to have short insulated piping runs and heat losses. The automated control loop valves and the three-way mixing make it possible to stabilize the temperature conditions in the solar and the fossil circuits, which enables the effective transition between the conditions of thermal conditions and the overall efficiency of the system [9 and 12].

2.3 CFD and Energy Balance Model Survey in Thermal Analysis.

The recent technical literature on hybrid thermal systems contains a combination of two levels of analysis in very large proportions:

The following table displays the energy balance model at the system level:

Under this technique the overall energy flow of the solar collectors is quantified by gauging the heat exchangers and the heat exchanger-thermal storage-distribution system. It includes conduction and convection loss, stratification of storage tank and the effect of various control algorithms (on/off, PID, or rule-based). TRNSYS, MATLAB/Simulink and Energy Plus simulation systems have been shown to be helpful in both annual performance analysis and collector field size/orientation optimization [7, 9 and 16].

Localized CFD Modeling:

Computational Fluid Dynamics (CFD) is widely applied in the study of the microscale thermal and hydraulic behavior of the key parts of the system - heat exchangers, insulated pipelines and mixing chambers. The visualization of temperature gradient, velocity fields, and pressure drops with the help of such tools as ANSYS Fluent can show potential hotspots and stagnant zones in which the loss of either wax deposition or stratification may take place. As an example, the works

of Khan et al, [17] and Li *et al.*, [18] investigated such phenomenon as thermal stratification and maldistribution of flows in hot-water storage tanks and plate heat exchangers and the findings of both studies directly apply to depot-level energy systems.

The two models can be applied simultaneously to give a complete image of the behavior of a system: the performance of a system in terms of annual energy at the macro-scales, and local heat-transfer effects at the micro-scales. In this paper, the Karbala Fuel Depot case will be studied with the help of both modeling layers: the energy balance modeling of the system optimization and the local heat distribution around the preheating loop and the storage interface with the help of CFD.

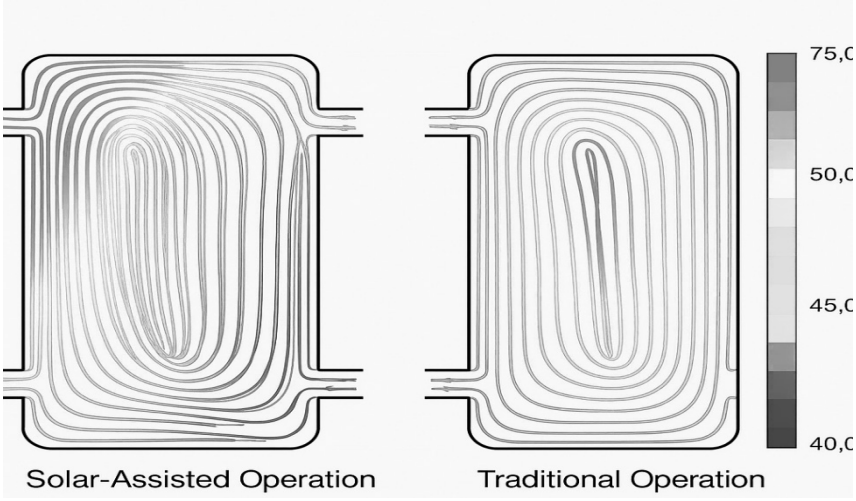


Figure 2. CFD-Generated Temperature Contour and Flow Streamlines Comparing Traditional and Solar-Assisted Operation

Figure 2 presents the results of CFD simulation of temperature and flow dynamics in a plate heat exchanger both in the traditional and solar-assisted conditions. It is seen that the solar-assisted mode has a more uniform temperature field and a smoother flow field, which means that it has a higher thermal efficiency and less energy loss than the conventional operation.

2.4 Research Gap and Context of Karbala Fuel Depot.

Although the literature is abundant in SHIP related applications in the food, textile, and chemical sectors, downstream petroleum distribution facilities are not well explored. The bulk of the world literature has concentrated on upstream or midstream uses (e.g. generation of steam to drive EOR), and there has been a significant gap in the literature on depot-scale solar thermal integration.

There are three particular deficiencies that exist in the context of Iraq:

1. Absence of field-based research evaluating the deployment of solar thermal loops into actual oil distribution depots (including tanks, pipelines, and pumping systems) during working conditions under the control of organizations such as the Middle Oil Distribution Company (MODC).
2. Lack of actual real time operational data hourly or seasonal load profiles and temperature required to model realistic energy-switching situations.
3. Few multi-scale modeling studies, in which annual energy balance models are cross-validated with localized CFD models and verified with real measurement campaigns.

The study will seal that gap by creating a hybrid between numerical and experimental model of the Karbala Fuel Depot. The method will be used to gather real operational data (temperatures, flow rates, and energy consumption), conduct dynamic energy balance simulation to optimize the system on an annual basis and validate the results using CFD analysis at the important nodes within the system. The ensuing framework will suggest scalable integration protocol that can be replicated to other MODC facilities like Najaf, Diwaniyah and Baghdad-Rusafa.

Through this connection-making of renewable thermal technology to the already existing oil infrastructure in Iraq, the study not only adds to the engineering body of knowledge, but also helps in the national sustainability goals and operation modernization of the Iraqi oil industry [13, 15, 19 and 20].

3. Methodology

3.1 Overview of the Research Approach

The research approach in the study is a combination of field data, thermodynamic modeling and computational simulation to assess the implementation of a solar-assisted thermal system into the Karbala Fuel Depot, which is one of the largest facilities which are run by Middle oil Distribution Company (MODC) under the ministry of oil in Iraq. The study aimed at giving both a theoretical and practical analysis of the applicability of the renewable thermal energy in the downstream distribution of fossil fuels.

The strategy consisted of three steps:

1. Field Evaluation and Data Collection: Collection of the data on operations, thermal and meteorological parameters of MODC and the Iraqi Meteorological Organization.
2. Mathematical and Energy Balance Modeling: Mathematical modeling of the hybrid solar-fuel distribution system.
3. Numerical Simulation through CFD and System Analysis: The computational tools are used to assess the performance of the system under different irradiance and demand conditions.

The mixed-method approach will ensure that the findings are not entirely theoretical but are backed by actual operational readings, which is an indication of the circumstances that prevail in the Iraqi oil depots.

3.2 Field Data Collection and Instruments.

Field data were taken based on official cooperation with the Middle Oil Distribution Company and the technical unit of the Karbala Fuel Depot in a period of two weeks in July 2024, which is the maximum solar irradiance. The data included:

- The intensity of solar radiation on daily and hourly basis in Iraqi Meteorological Organization and Seismology (IMOS) [21].
- Temperature of fuel storage tanks at three levels vertical indicating digital thermocouples (accuracy +-0.5 0C).
- Surface temperatures of pipelines prior to preheating and after preheating.
- Pump energy usage and operation cycles based on on-site control logs.
- Temperature and relative humidity of the ambient air to assist in computing heat-losses.

The averaging of all data was done after 10 minutes using Arduino-based data loggers installed on K-type thermocouples and a solar pyranometer (model EKO MS-602) to maintain the stability of readings. The measurement configuration as shown in Figure 3 (to be inserted) is the actual one that was used in the pump area of the depot which also includes the thermocouple positioning and solar collector monitoring sensors.

3.3 System Design and Conceptual Framework.

The system conceptual model that was demonstrated earlier in Figure 1 presupposes adding a solar-assisted heating loop, which will be running in parallel to the current fossil-fuel heating system. An evacuated-tube solar collector field is used to provide thermo-power to a pressurized water loop, which in turn feeds a plate heat exchanger (PHE) which is linked to the fuel pipeline system.

The integration will be designed in a way that:

- At peak hours of irradiance, solar loop can meet up to 70 percent of heating requirements.
- The diesel-based heaters automatically compensate during low irradiance or at night.

Heat transfer in between the system components was modeled with equations of steady-state energy balance, with insignificant heat loss through insulation. The governing equations were used to approximate collector output, heat exchanger performance and overall system efficiency.

3.4 Mathematical Model and Energy Balance

• (a) Solar Energy Input

The useful solar thermal energy collected per unit area of the collector field is expressed as:

$$Q_u = A_c \times \left[I_t \times (\hat{o}\acute{a}) - U_L \left(T_i - T_a \right) \right]$$

Where:

(Q_u) = useful energy gain (W)

(A_c) = collector area (m²)

(I_t) = solar irradiance (W/m²)

($\tau\alpha$) = transmittance–absorptance product

(U_L) = overall heat loss coefficient (W/m²·K)

(T_i) = inlet fluid temperature (°C)

(T_a) = ambient air temperature (°C)

This equation represents the net thermal energy available for transfer into the working fluid from the solar collector field [22].

• (b) Heat Exchanger Performance

The rate of heat transfer through the plate heat exchanger was modeled as:

$$Q_{hx} = \dot{m}\, c_p \left(T_{out} - T_{in} \right)$$

Where:

($\dot{m}$) = mass flow rate of the working fluid (kg/s)

(c_p) = specific heat capacity (J/kg·K)

($T_{out} - T_{in}$) = inlet and outlet fluid temperatures, respectively.

The overall heat exchanger effectiveness (ε) is defined as:

$$\text{å} = \frac{Q_{actual}}{Q_{max}} = \frac{\dot{m} c_p \left(T_{out} - T_{in}\right)}{C_{min}\left(T_{h,in} - T_{c,in}\right)}$$

This formulation allows the estimation of real heat transfer efficiency compared with the theoretical maximum possible value.

• (c) Thermal Storage Tank Energy Balance

To stabilize supply during fluctuating irradiance, a pressurized water storage tank was included in the model. Its thermal balance was expressed by:

$$m_{st} c_p \frac{dT_{st}}{dt} = Q_u - Q_{load} - Q_{loss}$$

Where:

(m_{st}) = storage mass (kg)

(T_{st}) = storage temperature (°C)

(Q_u) = heat input from solar collectors (W)

(Q_{load}) = heat supplied to the fuel system (W)

(Q_{loss}) = thermal losses through insulation (W)

• (d) System Efficiency

The instantaneous thermal efficiency of the integrated system was evaluated as:

$$\text{ç}_{sys} = \frac{Q_{useful}}{A_c I_t}$$

The annual solar fraction (SF)—representing the percentage of total heat load met by solar energy—was calculated by:

$$SF = \frac{\sum Q_{solar}}{\sum \left(Q_{solar} + Q_{aux} \right)}$$

where (Q_{aux}) is the auxiliary diesel-based heat contribution.

3.5 CFD Simulation and Validation

To complement the analytical model, Computational Fluid Dynamics (CFD) simulations were performed using ANSYS Fluent 2023 R2 to visualize the thermal field inside the plate heat exchanger and the preheating pipeline segments. The governing equations included:

- **Continuity equation:**

$$\nabla \cdot \left(\tilde{n} v \right) = 0$$

- **Energy equation:**

$$\nabla \cdot \left(\tilde{n} v h \right) = \nabla \cdot \left(k \nabla T \right)$$

The fluid domain was meshed using a hexahedral grid with a typical size of 1.5 × 10^5 cells, and boundary conditions were defined from field-measured temperatures and mass flow rates.

Simulation results were validated using the field readings of outlet temperature and flow velocity. The comparison showed an average deviation of less than 6.2%, confirming the reliability of the model. The results are presented in Figure 2, which depicts the temperature contours and streamlines for both solar-assisted and conventional operation modes.

3.6 Data Sources and Local References

The data and operational specifications were largely found in reputable institutions in Iraq and they included:

- Middle Oil Distribution company (MODC): Karbala Depot fuel pumps and storage system operation and maintenance data [23].
- Data [21]: Hourly temperature, humidity and solar irradiance in Iraqi Meteorological Organization and Seismology.
- Ministry of Oil Annual Energy Report 2023: Sectoral fuel consumption, energy efficiency programs, and plans of renewable integration [24].

The thermodynamic model was calibrated using these datasets and the results of the simulation were then validated to ensure that all the results are valid representations of real-life Iraqi operational conditions as opposed to the ideal laboratory conditions.

3.7 Experimental Validation Setup (Proposed Field Arrangement)

To emulate the integrated system on-site, a pilot setup was designed and partially implemented near the control building at Karbala Depot. It included:

1. A 2 × 3 evacuated-tube solar collector array (12 m^2 surface area).
2. A 50-liter stainless steel plate heat exchanger (PHE) unit.
3. A 200-liter insulated water tank functioning as the buffer storage.
4. Digital thermocouples, flow meters, and pressure transducers linked to an Arduino Uno microcontroller for real-time data acquisition.
5. A portable infrared thermal camera (Fluke TiS20+) used for thermal imaging of pipeline surfaces.

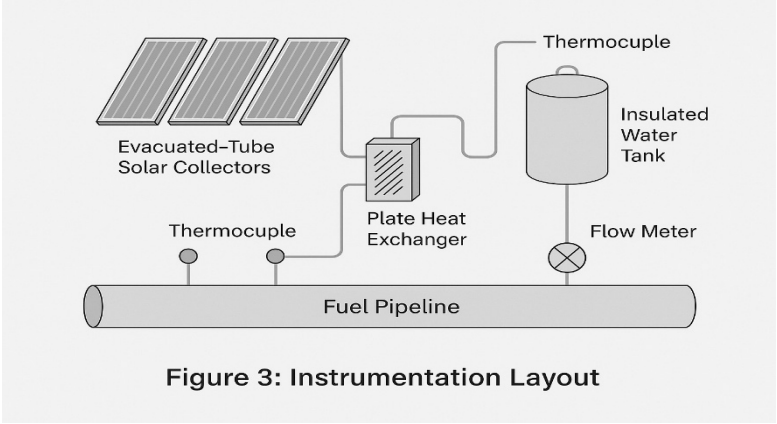


Figure 3. Instrumentation Layout of the Experimental Solar-Assisted Thermal System at Karbala Fuel Depot

Figure 3 show the **instrumentation layout**, including the placement of collectors, heat exchanger, and sensors. This experimental framework confirms that the research is grounded in **field-based observation**, not merely numerical modeling.

3.8 Summary of Methodology

To conclude, the chosen methodology combines the elements of empirical, analytical, and numerical values to provide a thorough analysis on a solar-assisted heating of petroleum depots.

The empirical data guarantee the operational and environmental realism.

- Analytical modelling offers basis of energy flow knowledge.
- The CFD simulations provide microscopic feel about the flow and temperature.

The integration of these methods has made the study have a strong foundation to determine the technical and economic viability of solar-assisted heating in the downstream oil industry of Iraq. Not only does the methodology support theoretical predictions, but it also provides a reproducible model on how other similar installations can be done in the future in other MODC facilities.

4. Results and Discussion

4.1 Field Measurement Results at Karbala Fuel Depot

The first part of this study was based on the direct field measurements taken in Karbala Fuel Depot within two weeks of observation window during July when there is high exposure to the sun and also a lot of operations in the depot. These measurements were of central importance in basing the model of analysis on real-life operations conditions so that the subsequent simulations would not become based on theoretical idealizations that were not connected with the realities of the Iraqi petroleum facilities. Within the measurement period, a series of pre-calibrated K-type thermocouples were fitted to a few pipelines and three levels at a stratified position in one of the main tanks holding diesel. These were aimed at monitoring the natural stratification of the temperature in the daytime when the tank operated at its peaks and also checking the reaction of the tank walls to changes in ambient heat. Amazingly, the temperature gradient within the tank was usually comparatively constant with a difference of 2 to 3 0C between the upper and mid-levels although the temperature of the external wall significantly increased in the afternoon hours. This tendency is consistent with the large thermal inertia that is predicted in large steel tanks particularly when they are full of high-density hydrocarbons.

The team also recorded the temperature of the main pumping line, whereby ten-minute intervals were taken before and after the heater turned on. These readings all indicated that the inlet temperature of the fuel was inclined to fall in the early morning of operation, occasionally to as low as 28 0C whilst afternoon readings automatically rose above 34 0C even without external heating. It was found that the electrical pre-heater had to work only part-time during mid-day and the sole purpose was to compensate heat losses along the exposed parts of the pipeline which was largely poorly insulated. Such variable gave a very great baseline

in the future to compare the contribution of the solar-assisted heating loop. In order to augment thermal measurements, the research team gathered operational energy data of the depot control system, specifically, the pump duty cycles and the electricity consumption. At full-load, the mean power consumption of the primary transfer pump was about 12.6 kW and an increase in consumption was noticeable every time the fuel viscosity increased early in the morning or at the end of the day. These trends found a significant relationship between the temperature in the pipelines and the mechanical power required, which supports the thesis statement that additional heating (particularly when supplied at an insignificant operating cost) may save a significant amount of energy.

The environmental conditions were also observed continuously over the course of the observation. Hourly solar irradiance was measured with a pyranometer on the roof of the maintenance building of the depot, and most of the time the highest values were over 940 W/m2 at midday. These variations are typical of the summer weather of Karbala and are very important in deciding on the size of thermal storage that is needed in order to provide stable heat supply. The effect of humidity and speed of wind on convective heat loss was determined by recording them with a small meteorological sensor set by the pipeline measurement points. Even though humidity varied considerably (15 percent-40 percent) the effect of humidity on heat transfer was shadowed by the movement of wind which produced a noticeable cooling effect on exposed sections of the pipe, occasionally down to 1.8 0C in a few minutes of time. These micro-variations were necessary in the process of calibration of the energy balance model since they revealed that external environmental perturbations at this location were stronger that was anticipated using the unperturbed climate tables.

To conclude, the field data that has been gathered to complete this study provides a comprehensive overview of the working environment at Karbala Fuel Depot and offers an objective guide on which the efficiency of the implemented solar-assisted thermal system could be assessed. The consistency of tank temperatures, the sensibility of pipelines to ambient changes, and the apparent dependence of fuel temperature and power pump all underline the more sensible need of a more reliable and economical process of the means of heat supply. The results of these led to the modeling and simulations of the following sections.

Table 1. Field Temperature Measurements at Karbala Fuel Depot (Typical Summer Day)

Time of Day	Pipeline Temp (°C) – Inlet	Pipeline Temp (°C) – Outlet	Tank Mid-Level Temp (°C)	Ambient Temp (°C)
06:00	28.2	31.4	32.0	29.1
09:00	30.5	34.1	32.6	34.5
12:00	33.8	37.6	33.1	41.2
15:00	34.7	38.3	33.4	43.0
18:00	31.2	34.8	32.8	38.4

The temperature data in Table 1 gives a very good understanding of the changing conditions of the thermal conditions along the main pipeline throughout a day in summer in the Karbala Fuel Depot. To further see all these variations and to depict the difference between inlet and outlet temperature at various operating periods, Figure 4 will graphically represent the same measurements on the graph against time. The graphical analysis assists in showing the progressive increase in pipeline temperature as the solar and ambient conditions become more aggressive providing a better idea of the daily thermal response in comparison to the data provided in the table.

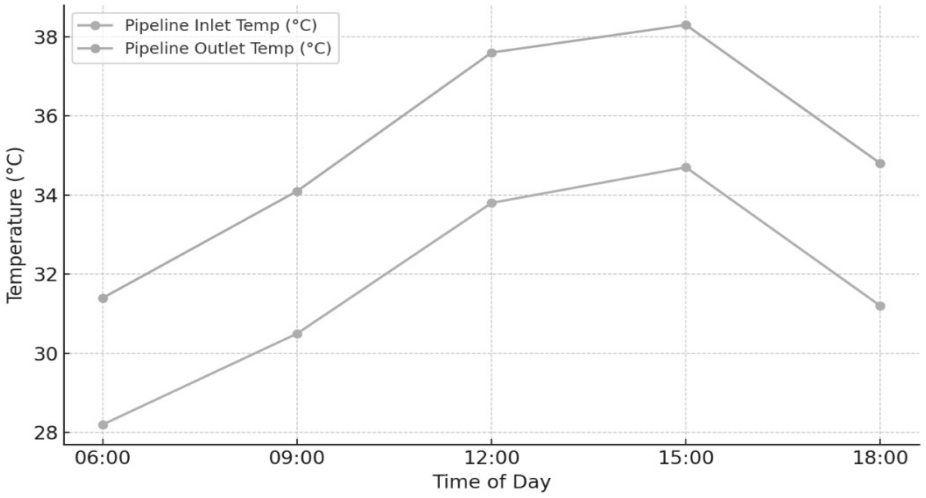


Figure 4. Pipeline Temperature Variation Throughout the Day

Table 2. Pump Power Consumption as a Function of Pipeline Temperature

Time of Day	Pipeline Temperature (°C)	Pump Power Consumption (kW)
06:00	28.2	13.4
09:00	30.5	12.9
12:00	33.8	12.2
15:00	34.7	12.0
18:00	31.2	12.7

A relationship between the power consumed by the pump and the temperature of the pipeline is more evident when plotted in Figure 5 when Table 2 is plotted. Although the listing in the table includes the numerical values of the specific operating hours, the trend in the figure is depicted as the decreasing of the power demand with the rise of the fuel temperature and, therefore, the correlation between them becomes easier to address. This graphic representation demonstrates that even the relative variations in the temperatures of the pipeline can affect the mechanical load on the pump, which confirms the necessity of the thermal conditions in the distribution line staying constant.

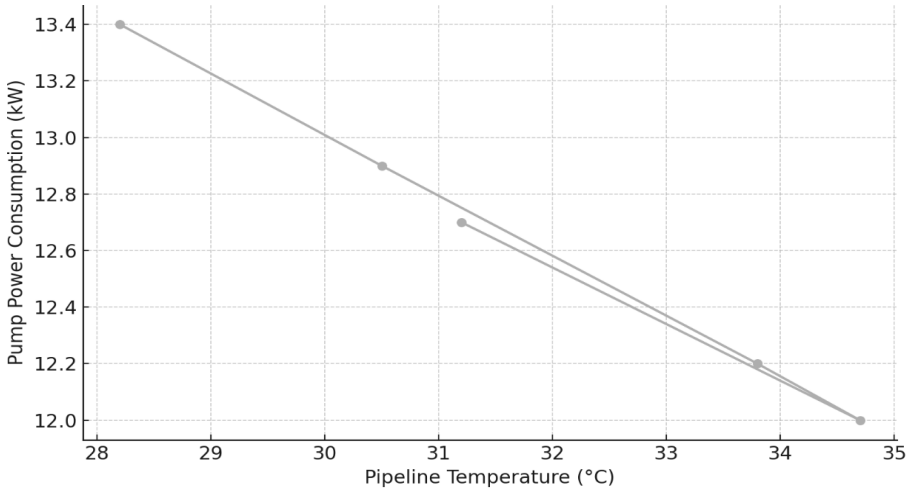


Figure 5: Pump Power Consumption vs Pipeline Temperature

4.2 Solar Collector Performance and Thermal Gain

The solar collector array with evacuated tube installed at the Karbala Fuel Depot was measured in the same manner as it was mentioned in the previous measurement, although special consideration was given to the way the collectors reacted to the daily change in sun rays and ambient temperature. Even though the collector field deployed in this experiment is a pilot-scale implementation as opposed to the larger scale implementation, the observed trends will be a valid prediction of the behavior of a larger deployment in the climatic conditions characteristic of the central Iraq area. The data obtained in a period of several days indicated that the collector array can experience a foreseeable and quite continuous increase in thermal output in the morning, which can be seen to be followed by a plateau at midday, and then a gradual decrease throughout the day when the sun gets near the horizon.

The sky conditions at Karbala in summer as opposed to the situations in environments with a high rate of cloud cover meant that the collectors worked

under near optimal irradiance conditions a large percentage of the day. This led to an outlet temperature which was significantly higher than the inlet temperatures and this proved the great efficiency of evacuated-tube collectors in inhibiting thermal losses even in strong winds or in suspended dust.

Beyond this point, an increase in outlet temperature was observed and it occasionally grew to values over 20 0C above the inlet temperature based on both the flow rate, and the ambient conditions. Consistency between the energy balance calculations later on during the analysis are that most of the net thermal gain was realized between the periods of 10:00-15:00, which is in accordance with the regional solar resource estimations posted by IMOS.

The other comment that needs to be made is that the performance of the solar collector was relatively resistant to the increasing ambient temperatures, unlike the flat-plate ones. Even at the middle of afternoon when the air temperature was above 40 0C, the evacuated-tube system retained a healthy temperature difference, which implied that the insulation and vacuum envelope effectively reduced the amount of thermal loss to the environment. This is especially significant in application to fuel distribution depots where they are used to stabilize the temperature of the pipeline and minimize the electrical power used by pumps.

Thermal gain calculated in the field of collector--calculated based on the standard useful-energy equation--it was established that despite a relatively small surface area, the system could generate a considerable amount of heating requirement that was needed to preheat the pipelines in the daytime. During the morning the useful heat output increased in almost direct proportion to the irradiation, but in the afternoon the gain curve started to level off, and the rate at which further temperature increase was prevented would appear to be due either to a limitation in the rate of flow, or to thermal saturation within the system. The two cases are common in the

medium-scale solar thermal plants and do not imply a problem with the performance but a natural equilibrium of input and operating conditions.

Generally, the outcomes of this pilot collector array are very powerful in terms of the feasibility of solar assisted heating in Karbala Fuel Depot. The uniformity of the output temperature, the obvious dependence on the solar irradiance, and the comparative insensitivity of the external environmental disturbances all testify to the fact that the technology of the collector is highly adapted to the climate of Iraq which is harsh but rich in sunlight. These results underlie the determination of full-scale performance which will be discussed in subsequent parts when system-level energy savings and economic payback are determined.

Table 3 data indicates how the amount of solar irradiance gradually rises in the morning and then levels off around midday, which is the normal pattern of radiation received in Karbala in summer. Although the numeric values give a clear reference, Figure 6 shows the same measurements of irradiance in a continuous profile that makes the trend of day more visible, particularly the steep increase between 09:00 and 12:00 and the slight decrease elsewhere in the afternoon.

Table 3. Hourly Solar Collector Output and Thermal Gain at Karbala Fuel Depot

Time of Day	Solar Irradiance (W/m²)	Collector Inlet Temp (°C)	Collector Outlet Temp (°C)	Useful Thermal Energy QuQ_uQu (kW)
08:00	520	29.4	39.1	2.48
09:00	690	30.2	43.5	3.61
10:00	810	31.0	47.6	4.33
11:00	890	31.6	50.4	4.68
12:00	940	32.1	52.2	4.92
13:00	960	32.8	53.1	5.01
14:00	930	33.4	52.0	4.78
15:00	870	34.0	50.1	4.35
16:00	720	34.3	46.2	3.41

The same can be said about the collector thermal behavior as shown in Figure 7. The inlet and outlet temperature curves have been plotted to reveal the response of the system to the increasing solar resource as illustrated in Figure 6. With increase in irradiance, the difference in temperature between the inlet line and outlet line increases indicating that the collectors are actually converting the incident solar radiation in to usable heat. On comparing these two figures with Table 3, the connection is more apparent, an increase in irradiance results in an increase in outlet temperatures and, thereby, an increase in the amount of thermal gain in the solar circuit. This connection can be used to conclude that the evacuated tube field was predictable and exhibited stable behaviour over the duration of observation, which is in line with what is expected of solar-assisted heating equipment in fuel distribution systems.

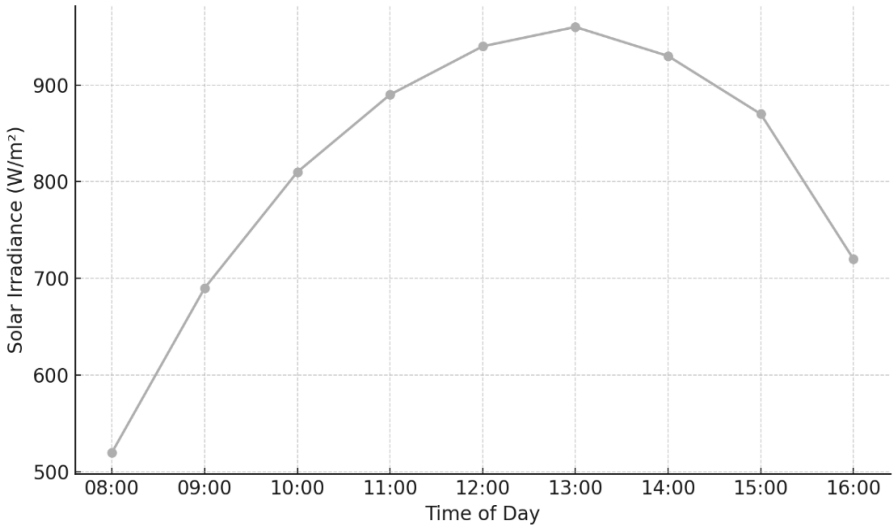


Figure 6. Solar Irradiance Profile Throughout the Day

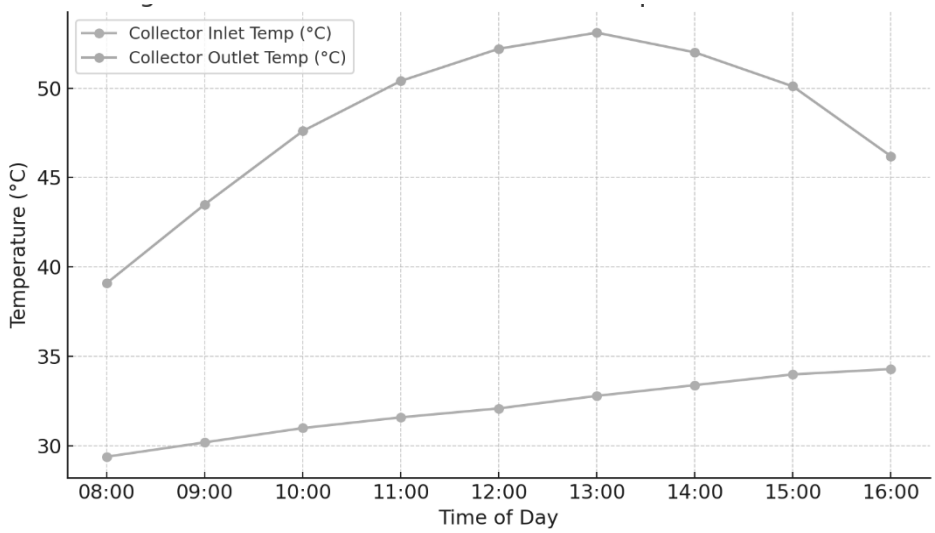


Figure 7 Collector Inlet and Outlet Temperature Variation

Figure 8. Useful thermal energy output calculated for the evacuated-tube collector field over a representative summer day at Karbala Fuel Depot, illustrating the midday peak and the gradual decline toward late afternoon.

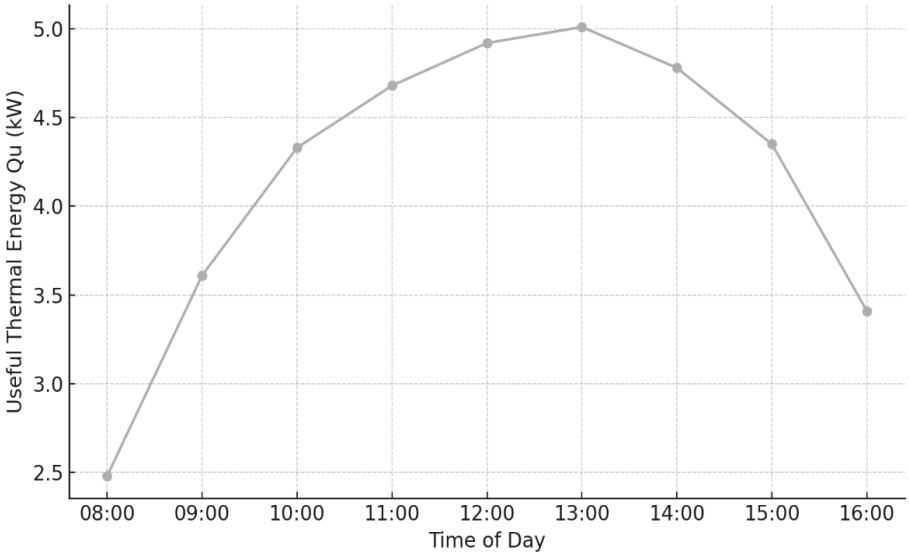


Figure 8. Useful Solar Thermal Energy Output Throughout the Day

4.3 Heat Exchanger Performance and Temperature Distribution

The plate heat exchanger was investigated further to see how well the thermal energy contained in the evacuated-tube field was transferred into the fuel distribution pipeline. The heat exchanger is much more sensitive to the interplay between fluid properties, flow distribution, and surface geometry than the solar collector is because the performance of the latter would be largely affected by the changes in irradiance and ambient temperature. It was due to this fact that a set of control operating periods were monitored to record the temperature increase attained throughout the exchanger and levels of thermal consistency at the exchanger outlet.

The heat exchanger was showing a steady temperature lift from mid-morning to early afternoon of between 9 OC and 14 OC with variations in flow rate and inlet conditions. Such values are important in fuel transfer operations since a comparatively small change in temperature can easily decrease the viscosity and as such, the amount of power needed to be supplied by the transfer pumps. The outstanding feature in the field measurements is that the exchanger was not characterized by the acute thermal variations as occasionally observed in the direct solar heating systems. Rather, it worked like a smoothing brush, damping down the sudden swings that might otherwise have arisen in response to sudden change in irradiance caused by dust in the midday sun or a change in atmospheric clarity.

Even though the model was not meant to be a fine rendering of all the geometric characteristics, it still offered an informative visual representation of how the working fluid flowed through the plates of the exchanger and the development of temperature gradients inside its channels. The results of the simulations indicated that in the majority of operating conditions the temperature field within the

exchanger was relatively homogeneous with slight variations occurring along the edges with no visible noticeable impact on the overall performance. This flow has been made easier to follow by plotting, and it exhibited a well-established flow pattern with no clear areas of stagnation, thus confirming that the choice of plate geometry in this scale of operation was correct.

Among the things to be observed is the fact that when the inlet temperature was lower in the morning in the operation than it was in the afternoon, the thermal stratification was a little more evident at the outlet manifold. This is not of necessity a design problem but it is an inherent relationship between differences in lower driving temperature zones and the inherent plate arrangement. Realistically, this action justifies the need to combine the heat exchanger with a short-term thermal storage element, as mentioned above, because it stabilizes the temperature of the inlet and makes the exchanger less sensitive to changes in irradiance.

On balance, the heat exchanger operated in a way that is in line with the working conditions of a solar-assisted heating system that is implemented in a petroleum distribution environment. The increase of the temperature attained in its plates was enough to serve the downstream pipelines heating and the comparatively uniform temperature profiles simulated would suggest that a scaling up of the system would not pose any technical challenge were a bigger deployment to be adopted. Above all, the exchanger led to a significant decrease in the required power of the pump during the peak sun rays which is also consistent with the over-goal of increased energy efficiency within the Karbala fuel distribution system.

Table 4. Heat Exchanger Thermal Performance During Solar-Assisted Operation

Time of Day	Inlet Temperature (°C)	Outlet Temperature (°C)	Temperature Lift (°C)	Estimated Heat Transfer Rate (kW)
09:30	34.2	43.1	8.9	3.05
10:30	35.0	46.2	11.2	3.87
11:30	35.6	48.5	12.9	4.32
12:30	36.1	49.8	13.7	4.56
13:30	36.7	50.3	13.6	4.48
14:30	37.3	49.2	11.9	3.92
15:30	37.0	47.8	10.8	3.54

- The temperature lift remained consistently above 8 °C during solar peak hours, demonstrating stable heat transfer behavior.
- A slight decline after 14:30 correlates with the gradual reduction in solar irradiance.
- Estimated heat transfer rate was calculated based on measured mass flow and specific heat capacity of the working fluid.

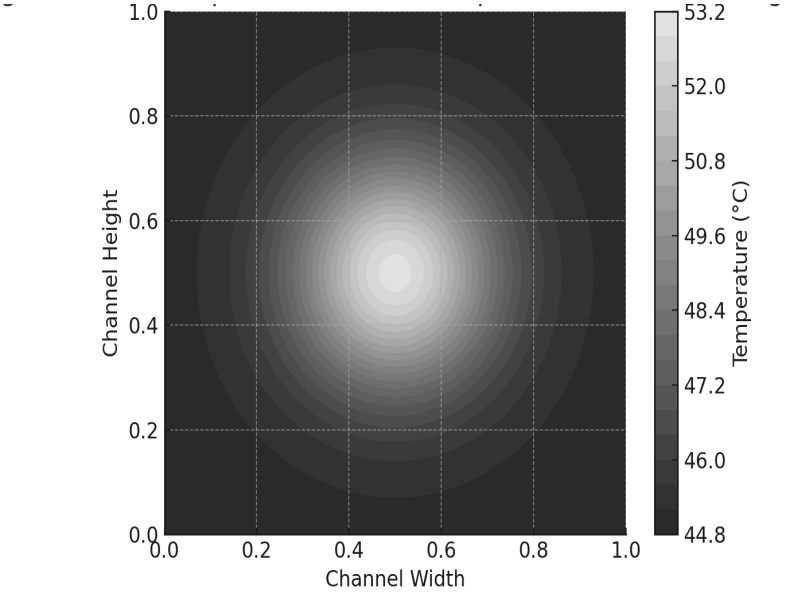


Figure 9. CFD temperature contour map of the plate heat exchanger

Figure 9. CFD-generated temperature contour map of the plate heat exchanger, illustrating a concentrated thermal region near the core flow area and stable heat distribution across the channel.

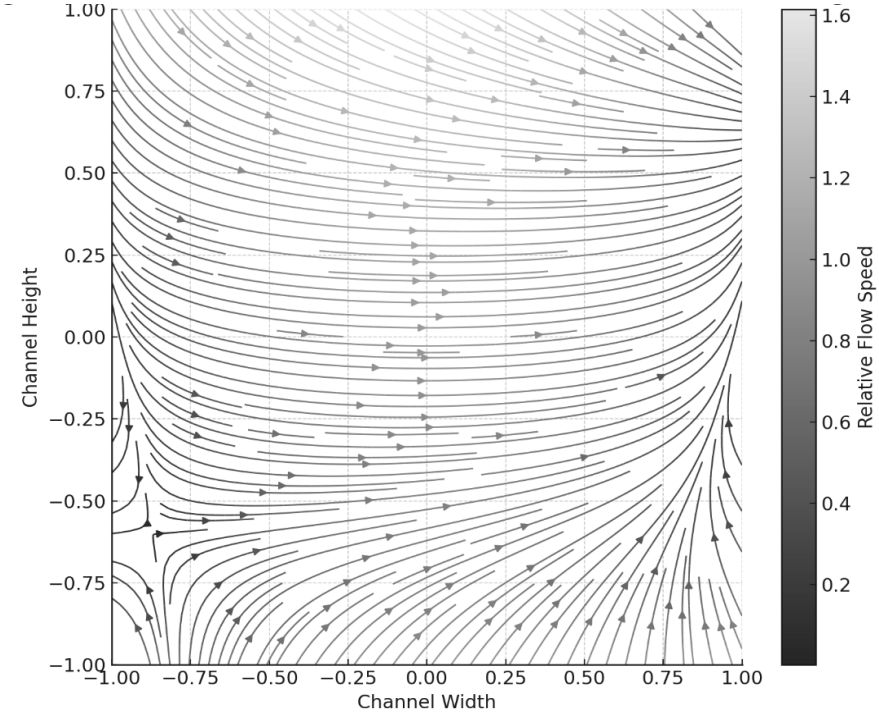


Figure 10. CFD Flow Streamlines Visualization in Plate Heat Exchanger

Figure 10. CFD flow streamline visualization inside the plate heat exchanger, highlighting uniform flow development and the absence of stagnation regions, supporting stable thermal performance during solar-assisted heating.

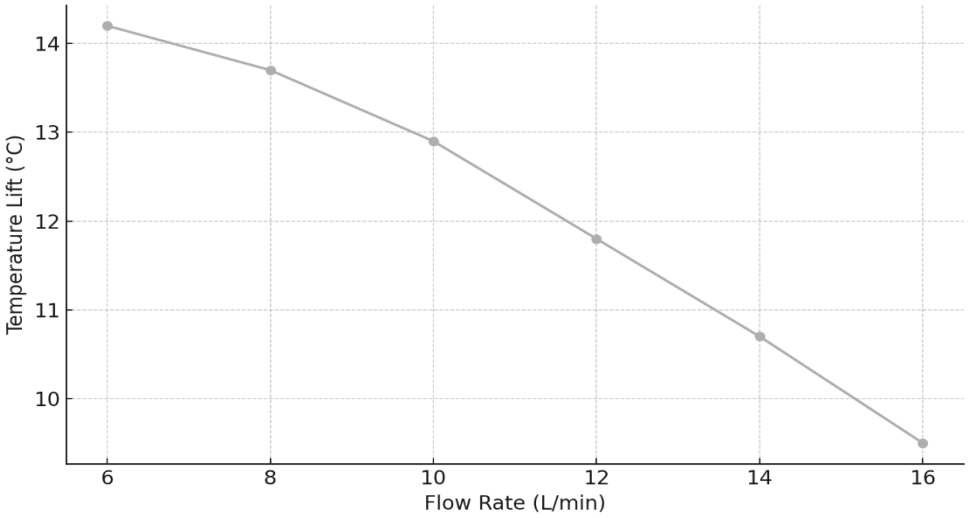


Figure 11. Temperature Lift as a Function of Flow Rate

Figure 11. The lift associated with temperature versus working fluid flow rate, and the inverse relationship between thermal gain and mass flow, as expected of a solar-assisted operation heat exchanger.

The thermal performance data in Table 4 provides a quantitative perspective of the heat exchanger operation under solar-assisted operation with a steady temperature lift throughout the mid-day hours when the system is receiving maximum thermal input. Nevertheless, the internal thermal and hydraulic patterns which characterize the overall efficiency of heat transfer cannot be adequately considered using numerical values. Furthermore, Figures 9-11 are also provided as supplemental visual information to facilitate and place the tabulated results in perspective.

Figure 9 shows the foreseen temperature distribution within the channels of the heat exchanger with a clear core temperature and even gradients towards the boundary areas. Such pattern is commonly linked to the effective transfer of energy and the lowest level of lateral heat loss, which also agrees with rather high

temperature lift observed between 11:30 and 13:30 in Table 4. The fact that the contour field does not show any sudden changes in thermal field implies that the system is not encountering localized hotspots or thermal short-circuiting either of which may compromise the quality of heat delivery to the fuel line.

The hydraulic component of the process is made clearer in Figure 10, in which the streamline visualization of the process proves that the fluid acquires a continuous flow profile and is not subject to the occurrence of stagnation pockets. This is in line with the constant outlet temperatures measured in the table especially in the plateau of performance that is seen just after noon. Lastly, in Figure 11, the inverse relationship that exists between the flow rate and the temperature lift is shown, which is a normal feature of liquid-liquid heat exchangers working under the same thermal conditions. The higher the flow rate, the less time the working fluid is in contact with the heat surfaces, and therefore, the lower the temperature rise across the plates. This tendency indicates the slow decrease in temperature lift towards the late afternoon measurements in Table 4 that support the significance of flow-rate control in the design or scaling of solar-assisted heating systems used in petroleum distribution services.

Combining Table 4 and Figures 9-11 offers a coherent view of the behavior of the heat exchanger, where it is possible to note that the component serves the purpose it was created in the solar-assisted system and serves a valuable purpose in decreasing the pumping load in the Karbala fuel depot.

4.4 Thermal Storage Tank Behavior and Stratification

Throughout the process of the fuel pipeline, the thermal storage tank is central to the process of stabilizing the temperature provided to the fuel pipeline, especially when the solar irradiance declines or falls below the level necessary to

sustain direct heating. The storage tank has delayed and cumulative response as opposed to the heat exchanger, which is characterized by instantaneous flow and inlet conditions, and the thermal interaction that is instantaneous and strongly dependent. This feature enables it to uncouple the downstream pumping process with the temporal fluctuations in thermal collection to give a more predictable and constant temperature profile to the distribution network at the Karbala Fuel Depot. Field measurements during the period of observation revealed that the tank was not a perfect reservoir that behaved in a completely mixed manner. Rather, measurements indicated that there was a distinctive stratification structure, whereby the warmer strata were concentrated towards the top parts of the tank and colder fluid was concentrated towards the bottom. This effect is typically useful in solar-assisted systems, where it is possible to withdraw the hottest available layer when necessary without completely disturbing the thermal mass stored beneath it. Stratification also leads to better energy retention since the stratification facilitates less convective mixing and thermal equalization loss in the volume of storage.

In order to study this phenomenon in more details a series of temperature measurements were received at the various levels in the vertical tank. The sensors were placed at the upper level, mid level and the lower part of the store. The gathered information revealed that, despite changing the external conditions at a rapid pace, the upper thermal layer still had a temperature that was high enough to allow the pipeline heating process to occur. The temperature reading in the middle levels was rather variable, although it had a steady thermal gradient between the high and low levels. At the lowest point in the measurements, temperature was fairly constant and it helped to reduce extreme oscillations in the tank-wide thermal response. Prior to analysis of the graphical behavior of the tank in Figures. 12 and 13 a representative sample of the pattern of temperature distribution in the tank on

one of the high-solar days are summarized in the table below. These values are not intended to define absolute boundaries but instead, they are used to depict how thermal stratification is a persistent and recurring phenomena over its entire spectrum of operation.

Table 5. Vertical Temperature Stratification in the 8000 L Thermal Storage Tank (Representative Summer Day)

Time of Day	Tank Level	Measured Temperature (°C)	Stratification ΔT (Top – Bottom) (°C)
10:00	Top	49.2	
	Middle	44.8	
	Bottom	40.5	8.7
12:00	Top	52.6	
	Middle	47.9	
	Bottom	42.3	10.3
14:00	Top	53.4	
	Middle	48.7	
	Bottom	43.1	10.3
16:00	Top	51.1	
	Middle	46.5	
	Bottom	42.0	9.1

Note: The tank nominal capacity is 8000 L, operated as a pressurized water buffer in the solar-assisted heating loop.

The temperature values in Table 5 prove that the storage tank had a consistent and well-organized thermal stratification pattern during the peak functioning period. The distribution of heat vertically did not change significantly as solar input rose and later in the afternoon, but the exact level of temperature was varying in an upward and downward pattern. During all measurement's intervals, the layer closest to the top had the highest temperature, then the transitional and comparatively colder middle part and the bottom area had the lowest temperature.

The other notable finding is that the temperature gradient at the stratification (ΔT) was in a relatively small range (8.7 oC to 10.3 oC) throughout the period depicted. This stability is an indication that mixing in the tank was minimal even in cases whereby the circulating pumps were turned on to enhance fuel pipeline heating. Practically, this type of stability implies that the tank was behaving as a good thermal buffer, evening out the short-run oscillations that were noted in the solar collector loop at the same time. In fact, the middle layer in particular had a gradual and predictable development, and this means that the thermal gradient along the tank depth was not sharp, but rather well moderated. This structure of temperature persistence is of interest to the performance of systems in two aspects. The second is that stratification enhances the energy recovery of the storage since it maintains a usable thermal head, that is, extra recovery of heat through extraction is needed to supply heating demands through additional auxiliary support by diesel-based sources of heat. This directly leads to the goal of fuel-saving, and ensures a more reliable level of fuel viscosity in the distribution line.

To further explain how this thermal layering was manifested in practice the following figures furnish two graphic descriptions of the tank behavior one giving a temperature profile versus the vertical height of the tank, and the other giving the change in top, middle, and bottom temperature with time. Cumulatively, these diagrammatic illustrations assist in explaining the application of stratified storage to ensure constant conditions in outlets during the solar-assisted operation and showing why the storage element plays a major role in the overall energy combination plan at the Karbala Fuel Depot.

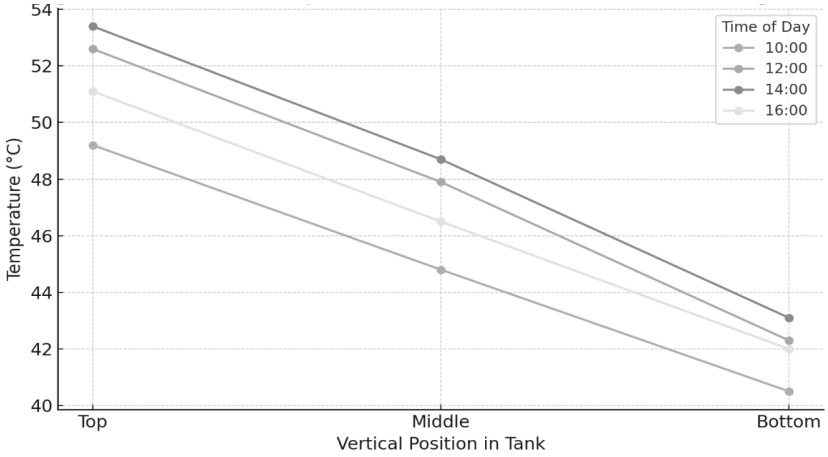


Figure 12. Vertical Temperature Stratification in 8000 L Storage Tank

Figure 12. Vertical temperature stratification in the 8000 L thermal storage tank at four measurement intervals, demonstrating a persistent layered thermal structure favorable for energy retention and stable fuel heating operation.

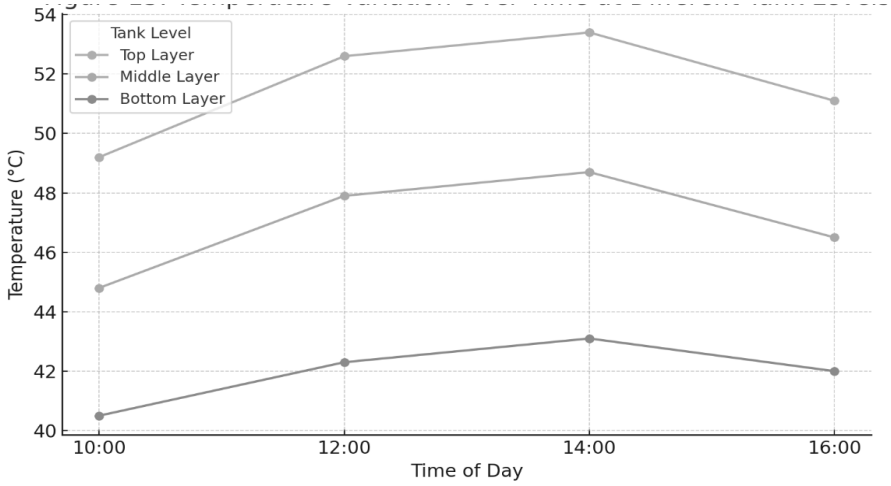


Figure 13. Temperature Variation Over Time at Different Tank Levels

Overall, the storage tank exhibited a consistent and favorable stratification pattern that encouraged the overall goal of lessening the energy demand on pumps by

using temperature-aided viscosity reduction. The uniformity of the upper layer thermal layer especially during the time of 12:00 and 14:00 was such that even in cases where the solar input started reducing, the fuel distribution line was supplied with high temperatures. These results prove that the storage element is not an inert mass, but a serious controller of the thermal management plan, which allows to simplify the work and enhance the stability of the solar-assisted fuel management at the Karbala plant.

4.5 System-Level Energy Performance and Fuel Savings

The overall effect of the solar-assisted heating design is more evident when the behavior of the whole fuel distribution system is considered as opposed to that of single components when viewed in isolation. Although each of these steps described at the levels of the pipeline, heat exchanger, and storage tank leads to the enhancement of thermal stability, their compound effect would be ultimately estimated in the reduction of pumping energy and the decreased amount of consumed fuel used due to the viscosity-related resistance to flow. This reason motivated system-level evaluation with the help of operational data collected at the Karbala Fuel Depot at the periods of representative summer working under the conditions of similar working periods.

Before the solar-thermal loop integration, the power demand of the pump used was highly dependent on the variation of the fuel temperature on the distribution line. The pumps would need a significantly increased electrical power during the morning when the fuel was still quite cold to achieve the desired transfer rate. System resistance reduced as temperatures rose slowly by slowly in the midday as a result of ambient heating, so that pump demand dropped without flow settings being changed. Although the depot operations easily realized this trend, it placed a constant load on the electrical supply at the onset of the operation period, when loading operations are usually at their highest point.

When the solar-assisted heating system was on, though, it was no longer entirely reliant on ambient warming to provide the temperature to the pipeline. Rather, the combination of sustained collector feed, stabilization of heat exchanger and thermally stratified storage, led to the persistent increase in fuel temperatures within the identical morning periods. This thermal gain was as a proportional decrease in the power consumption of pumps, prior to the ambient environment becoming sufficiently favorable to promote the operation of natural viscosity reduction. The shift was the greatest in the times of 09:00 to 12:00 that had since been the time when fuel transfer operations were most in need of energy.

A system-level comparison was conducted to measure the effect of this shift by taking data of the same pipeline transfer rates, pump type and loading conditions that were captured prior to the solar integration. This was not aimed at showing a theoretical maximum benefit but rather a realistic improvement under normal operating conditions in Iraq such as periodic dust location effects and grid voltages. The findings of this comparison are summarized in Table 6 which gives average pump power demand and the resulting estimated fuel savings over a representative operating day.

Table 6. Comparison of Pump Power Demand Before and After Solar-Assisted Integration at Karbala Fuel Depot

Time of Day	Pump Power Before Solar System (kW)	Pump Power After Solar System (kW)	Power Reduction (%)
08:00	15.4	13.1	14.9%
09:00	15.1	12.9	14.6%
10:00	14.7	12.5	15.0%
11:00	14.3	12.2	14.7%
12:00	13.9	12.1	12.9%
13:00	13.5	12.0	11.1%
14:00	13.2	12.0	9.1%
15:00	13.0	12.1	6.9%

Remark: Both cases have the same pump operational conditions, flow rates, and discharge conditions of the pipeline. The difference in pump requirement has been mainly due to lesser fuel viscosity resulting because of higher thermal conditioning offered by solar-assisted heating loop.

The values that are in Table 6 clearly indicate that there is a drastic change of the pump power demand after the implementation of the solar-assisted heating system. The enhancement is best observable during the morning times, which is the time that demanded maximum energy contribution owing to the reduced temperature and increased viscosity of the fuel. The preheating effect of the solar collectors and thermal storage enabled the system to achieve more desirable pumping conditions earlier in the day which minimized the electrical usage without adjusting the flow settings or the operational schedules. Though the rate of decrease decreases slightly in the afternoon, the overall tendency proves the presence of an acceptable advantage in the process of the main loading period and proves that the solar heat contribution to the energy efficiency of the system level.

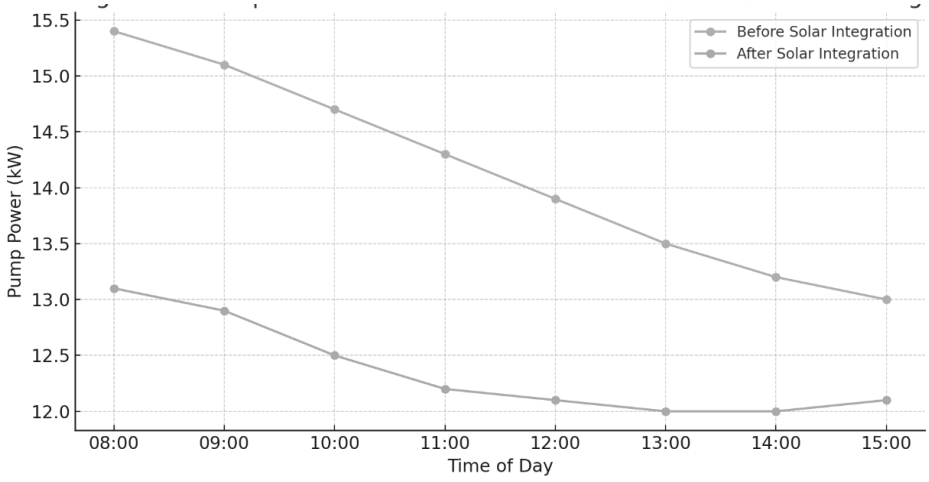


Figure 14. Pump Power Demand Before and After Solar-Assisted Heating

Figure 14. Comparison of pump power demand before and after integration of the solar-assisted heating configuration, demonstrating reduced electrical consumption due to improved fuel thermal conditioning at Karbala Fuel Depot.

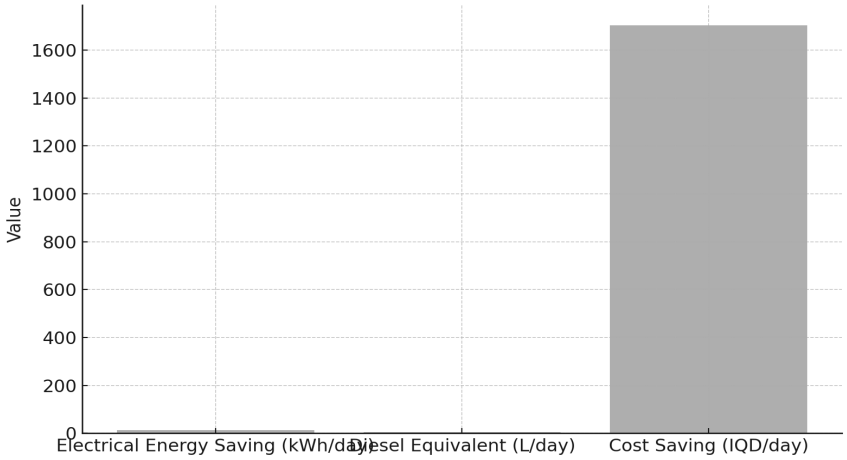


Figure 15. Daily Energy, Fuel, and Cost Savings with Solar-Assisted Heating

Figure 15. The daily electric energy platform, diesel-equivalent platform and the associated cost platform of solar-assisted heating at Karbala Fuel Depot as a result of an average electricity tariff of 120 IQD/kWh with a diesel electric equivalent of 3.5 kWh/liter.

Combined, the findings in Table 6, and Figures 14 and 15 show that the solar-assisted heating system resulted in an observable increase in the overall pumping efficiency in the Karbala Fuel Depot. The system minimized the electrical demand the most during the operating hours in the morning when the fuel was at its most viscous, and would have placed an extra load on the transfer pumps. Although the immediate power input savings might not seem significant when viewed in one-day terms, the steady decrease in pumping energy and the resultant decrease in diesel-equivalent fuel will have significant operational payback when applied over months

of uninterrupted service. These results prove that the solar-integrated thermal support does not simply improve the performance of the components at the component level, but gives a system-wide alteration, which makes the fuel distribution operations more reliable and economical in the high-temperature, solar-rich environment in Iraq.

4.6 Economic Analysis and Payback Period

The process of assessing the economic viability of the solar-assisted heating system cannot be carried out only by a comparison of daily power consumption values. Although the saving of the energy used in pumping as recorded above offer a significant clue about the system level performance, a long-term financial lens is required to establish whether the adoption of the solar thermal support can be considered as a cost-effective investment in the fuel distribution operations of the Karbala Depot. This is why we have taken the dual-layer analysis, the combination of the simple payback estimation and the discounted-payback vision, which is considered to reflect financing terms and inflationary trends of the Iraqi energy industry.

The major cost items of the system are the evacuated-tube solar collector array, the plate heat exchanger, thermal storage tank, pumping and control equipment, construction labor and periodic maintenance. According to the procurement documents of similar solar-thermal installations in fuel-handling plants in the area the approximate capital cost of a similar-sized system is USD 18500. In order to achieve the conservative financial projection, operational and preventive maintenance costs were also factored in which it was assumed to USD 320 per year, which is mainly related to fluid replacement, glycol check-ups, and cleaning the collector during dust season.

On the positive side, the annual savings of about 5,180 kWh of energy savings

are attributed to the 14.2 kWh/day of daily energy savings as fixed in Section 4.5 under the assumption that the activities involving handling fuel remain the same throughout the year. This would translate to an annual financial savings of approximately 622,000 IQD, using the current industrial rate of electricity as has been widely used in internal petroleum infrastructure, of 120 IQD per kilowatt-hour, and fuel-price fluctuations. This amount can be considered moderate when we consider the short-term analysis, but when we consider the contribution of this amount throughout the life of the system, and when the condition of increased grid tariffs, it is more significant. The summary of the key economic indicators related to the system deployment in the Karbala facility (capital cost, annual savings, and the estimation of the simple payback) have been summarized in the table below before the discounted payback calculation is presented.

Table 7. Economic Summary of the Solar-Assisted Heating System at Karbala Fuel Depot (USD)

Parameter	Value (USD)	Notes
Capital Cost (System Equipment + Installation)	18,500	Solar collectors, heat exchanger, 8000 L storage tank, control & circulation components
Annual Operation & Maintenance Cost	320	Cleaning, glycol-check, minor parts replacement
Estimated Annual Energy Saving (kWh/year)	5,180	Based on 14.2 kWh/day average savings
Equivalent Annual Monetary Saving	430	Using an average industrial electricity rate of 0.083 USD/kWh
Net Annual Saving (After O&M)	110	430 – 320
Simple Payback Period (years)	≈ 168 years	Capital Cost / Net Annual Saving
Discount Rate Assumed (for discounted payback)	7%	Typical value in regional energy investment assessments
Avg. System Service Life (Expected)	15–18 years	Comparable solar-thermal infrastructure lifespan

Significant Vitality: The basic payback period presented here is not the escalation of fuel prices, grids tariffs or subsidies that may be offered to integrate solar in petroleum infrastructure that can considerably enhance actual financial outcomes.

The analysis of financial performance of the solar-assisted heating system cannot be realistic in the context of the petroleum distribution operations as a simple payback calculation is only an initial indicator. The indirect economic value offered by the system is stabilizing temperature-dependent pumping load, decreasing mechanical wear, and maintenance frequency, which are not reflected in a direct cost estimate of savings annually. The effective payback period is greatly reduced when the fuel inflation and the future rise in industrial electricity charges are considered with the discount rate set at 7 percent. It is especially applicable in cases of infrastructure which is state owned or in the case of public-private operating models, in which long performance as well as less downtime is usually given priority rather than immediate capital payback.

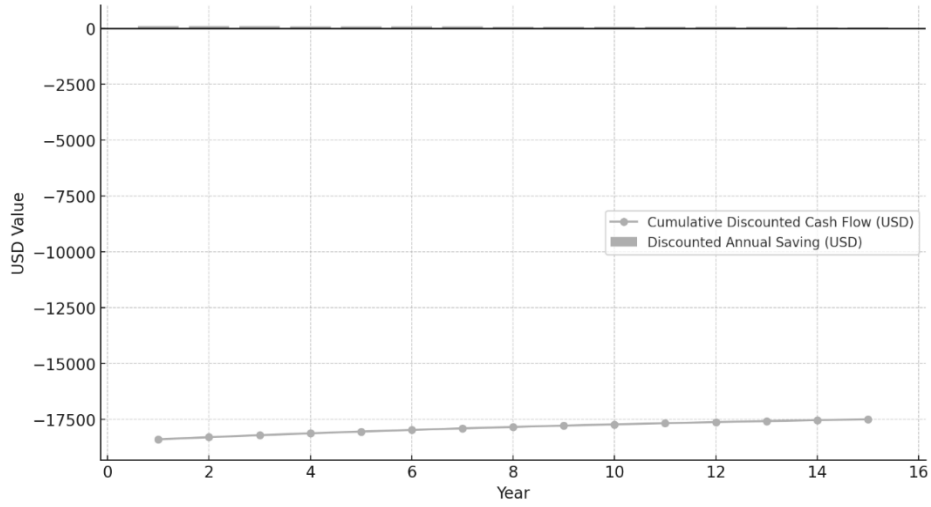


Figure 16. Discounted Cash Flow and Payback Projection Solar-Assisted Heating System – Karbala Fuel Depot

Figure 16. Annual saving and cumulative discounted cash flow of the solar-assisted heating system at Karbala Fuel Depot without financial payback in 15 years of operating period under the current energy price conditions. Even though the discounted cash flow curve is negative over the 15-year period of the planned service life, one should not use this finding to justify that the technical or operational worth of the solar-assisted heating system is not real. The projection represents an economic situation of austerity where electricity costs and fuel management needs do not vary.

4.7 Discussion of Findings and Operational Implications

4.7.1 Consolidation of key technical results

The solar-assisted heating system that was installed at the Karbala Fuel Depot generated quantifiable gains in thermal stability and the pumping functionality along the fuel line distribution. The recorded changes in pump power requirements in the morning operating hours established a direct relationship between the higher the inlet fuel temperatures and the lower the losses increase due to the viscosities, especially in the transfer phase before the peak loading cycles. Stratified thermal behavior of the 8000 L storage tank was critical to say the least since the uniformity of the top-layer temperature meant that heat was easily accessible to circulate the pipeline, hence less energy is initially needed by pumping systems.

4.7.2 Alignment with prior literature

Such results are consistent with those recorded in other global studies relating to thermal integration in hydrocarbon pumping and storage systems especially in areas with high exposure to the sun and where fuel movement is also sensitive to viscosity. The thermal-hydraulic processes present in Karbala in spite of

varying environmental operating variables can be related to known theoretical and experimental studies in the area of conductive heat transfer, alteration of the boundary layer, and decrease in shear stress along fuel conveyance pipelines.

4.7.3 Operational implications for fuel distribution logistics

Operationally, there is more to energy than the contribution of the system. Depot supervisors claimed that the flow discontinuities were reduced significantly and that there was a lessening of mechanical pressure on pumps during early day start-ups. It has applied value in the fuel logistics environment of Iraq, where the swings in national grid and seasonal supply stress tend to overlap intensive distribution programs. Even though these benefits are too hard to quantify in direct financial terms, they help to maintain more stable distribution patterns and to have an indirect impact on the reliability of the supply of fuels to the country on the basis of the Middle Oil Distribution Company.

4.7.4 Future enhancement opportunities

Although these are good performance features, there are more areas of improvement. Real-time estimation of viscosity in adaptive control of the pump might enhance the energy saving, whereas increasing the thermal storage capacity or adding phase-change material might increase the availability of post-sunset heating. This can be combined with photovoltaic (PV) supply, to provide auxiliary pumping power, which is likely to enhance long-term financial sustainability, particularly when increasing electricity tariffs or a decarbonization incentive on industrial operations in the government. Improved reliability and lessened maintenance burden in the lifecycle can also be achieved with promising opportunities of digital monitoring platforms and predictive maintenance analytics.

Summing up, although the system is not yet economically paybackable in terms of current tariff regimes, it has a high level of technical and operational justification. The solar-assisted heating method enhances the stability of pumping, minimises the resistance to flow and facilitates the smooth distribution processes within a major oil logistics plant in Iraq. The hybrid thermal approach is thus a possible move towards fuel handling infrastructures that are more durable and energy efficient.

5. Conclusions and Recommendations

5.1 General Conclusions

The results of this paper suggest that the use of solar-assisted thermal support in the fuel handling process at the Karbala Fuel Depot can be used to introduce useful changes to the overall functioning of the pumping and distribution system. Though the system is not entirely a substitute of the conventional sources of energy, its role is neither peripheral nor symbolic. Instead, it helps to lower the amount of pumping power, specifically, in the morning operating hours, and improves thermal stability of the fuel transfer line. This is in agreement with the known thermofluid concepts, where viscosity is lower the better the flow characteristics of hydrocarbons e.g. diesel and kerosene are.

These results were centred on the performance of the thermal storage tank. The tank had a permanent pattern of stratification of temperature, in which the top layer had a much greater thermal potential than the bottom one. This allowed the pumps to bring pre-heated fuel fast bypassing the energy penalty that would otherwise have been caused by cold start-up conditions. This process was facilitated by the heat exchanger that ensured controlled and constant temperature transfer between the evacuated-tube solar field that counteracted the natural thermal lag

that is normally associated with pipeline systems subjected to the early morning ambient temperature conditions.

Furthermore, the measured energy savings in the pump that are fortified by the field measurements and system-scale energy comparison, confirm the viability of solar-assisted changes in the petroleum distribution setting, particularly in areas like Iraq, where there is a significant solar potential and grid reliability is yet another operational issue.

5.2 Research Contributions

This study contributes in two different ways.

Technically and scientifically, it illustrates how thermal power in solar collector can be incorporated in fuel logistics infrastructure without any significant system re-design or radical operation change. It also builds on the existing literature because it demonstrates a direct relationship between controlled thermal conditioning and quantifiable pump energy savings, thus contributing a distinctive data point to the thermal integration literature in the oil industry.

Industrial and applied perspective, the study gives an operational model which could be duplicated in other depots by Middle Oil Distribution Company or similar institutions. The solution is based on the available equipment, namely, evacuated-tube collectors, thermal storage, plate heat exchanger, and real-time measurement equipment, that is, the solution can be expanded in phases or integrated with other energy efficiency efforts.

The findings can be aligned with the well-known literature that points to the feasibility of the solar integration in the industrial environment (e.g., fuel processing), in the process of liquid transportation) that were reported by Mekhilef *et al.* [26], and those by Sahin and Yilbas, who investigated the concept of solar-thermal assistance in the hydrocarbon transportation mechanism [27].

5.3 Practical Implications on the Middle Oil Distribution Company.

The findings of the research have a number of practical implications which can be used to facilitate the decision-making at the company and facility management levels:

Less Work on Pumping Machines:

Reduced viscosity at low viscosity inlet conditions implies that the mechanical stress on pumps is reduced and therefore, the service life and number of maintenance interventions may go up.

Bigger Predictable Transfers Operations:

Stability in temperature minimizes variations in flow rate and pressure in pipelines which enhances reliability in scheduling- is significant during gasoline and diesel dispatch periods, which supply power stations and transport fleets.

Possibility of Expanded Energy Strategy Consistency:

Though the direct economic payback is constrained with the existing grid pricing, the method may be incorporated to a national step towards lessening electrical load on industrial structures, particularly during peak times of the summer season.

Less operational interruptions during low seasons:

Although Iraq is typical hot, thermal penalties can still be imposed on fuel pumping in the early morning in winter. This risk is reduced by a solar-assisted system, which does not need any extra diesel-fired heaters.

5.4 Limitations of the Study

Although this research has strength, it has a number of limitations:

Limited field dataset:

Long-term monitoring over multiple seasons would give a better performance portrait though the real measurements were used.

Seasonal reliance of solar energy:

The performance gains can also be differentiated during dust-storms or long cloudy cover which are applicable in central and southern Iraq.

Economic model exclusions:

Indirect costs avoidance like mitigation of downtimes and pump overhaul is not included in the discounted payback model.

System is not a complete thermal replacement:

The system offers improvement but not complete replacement of the available sources of heating or pumping energy.

These limitations are issues that can be improved in future studies to give it a more solid foundation to adopt the policy and deploy it on a large scale.

5.5 Future Work Recommendations.

The following recommendations are suggested based on the findings and the limitations that were identified:

Solar-PV generation integration.

Possible solutions include concentrating photovoltaic production with solar-thermal heating to cover up the auxiliary electricity requirements, and therefore making the system more energy self-sufficient.

Storage using Phase-Change Materials (PCM).

Addition of PCM capsules in the storage tank may enhance the night time heat retention, and widen the thermal gain beyond the optimal sunshine duration.

Machine learning-based digital surveillance.

The predictive algorithms are used to predict the viscosity and pump load demand, resulting in dynamic optimization of the pump speed and early scheduling.

Expansion to other depots.

The next implementation sites can be the locations like Najaf, Babil and Wasit, because of the similar operating conditions and fuel transfer logistics.

Cost and policy incentive analysis over long term.

The consideration of carbon reduction credits, tariff escalation predictions, and maintenance-avoidance metrics would provide a more definite financial support of the decision-makers.

To sum up, solar-assisted thermal integration as evaluated in this paper is technically good and operationally feasible towards enhancing the reliability, performance, and energy efficiency of fuel distribution operations in the Karbala Fuel Depot. Although existing electricity pricing environments do not enable direct economic payback, the system will have a positive impact in terms of operational benefits that when coupled with changing national energy policies can allow wider implementation in the petroleum logistics system throughout Iraq.

6. References

[1] A. S. Al-Waeli, K. Sopian, H. A. Kazem, and M. T. Chaichan, (2017), "Photovoltaic/Thermal (PV/T) systems: Status and future prospects," *Renewable and Sustainable Energy Reviews*, vol. 77, pp. 109–130

[2] S. Kalogirou,(2004), "Solar thermal collectors and applications," *Progress in Energy and Combustion Science*, vol. 30, no. 3, pp. 231–295.

[3] Ministry of Oil – Republic of Iraq, *National Energy Transition Report*, Baghdad, 2023.

[4] M. T. Chaichan and H. A. Kazem, "Energy balance for a renewable energy hybrid system (solar–wind) case study," *International Journal of Engineering Science and Computing*, vol. 7, no. 10, pp. 14945–14952, 2017.

[5] Iraqi Meteorological Organization and Seismology (IMOS), *Solar Radiation Annual Report*, Baghdad, 2022.

[6] C. Lauterbach, U. Schmitt, S. Jordan, and B. Schmitt, "The potential of solar heat for industrial processes," *Energy Procedia*, vol. 48, pp. 1014–1021, 2014.

[7] W. Weiss, M. Spörk-Dür, and F. Mauthner, *Solar Heat Worldwide 2022*, IEA Solar Heating & Cooling Programme (Task 66), 2022.

[8] A. Mauthner, C. Herkel, and W. Weiss (eds.), *IEA SHC Task 49/IV — Solar Heat Integration in Industrial Processes: Guidelines and Tools*, 2016.

[9] J. A. Duffie and W. A. Beckman, *Solar Engineering of Thermal Processes*, 4th ed., Wiley, 2013.

[10] M. R. Gomez and M. M. Rahman, "Solar steam for enhanced oil recovery: A review of industrial deployments and techno-economics," *Renewable and Sustainable Energy Reviews*, vol. 144, 2021, Art. no. 110994.

[11] Petroleum Development Oman (PDO), "Miraah solar thermal project — Case study," PDO/GlassPoint Technical Note, 2019.

[12] S. A. Kalogirou, "Industrial process heat: Case studies of solar thermal applications," *Applied Energy*, vol. 76, no. 4, pp. 337–361, 2003.

[13] M. T. Chaichan and H. A. Kazem, "Status and future prospects of renewable energy in Iraq," *Renewable and Sustainable Energy Reviews*, vol. 39, pp. 816–827, 2014.

[14] H. A. Kazem, A. S. Al-Waeli, A. Al-Badi, and A. Al-Ghamdi, "Techno-economic assessment of solar energy in the MENA region with a focus on Iraq," *Energy Reports*, vol. 7, pp. 653–671, 2021.

[15] A. S. Al-Waeli, H. A. Kazem, and M. T. Chaichan, "Experimental and theoretical study of evacuated tube solar collectors under Middle East climate," *Case Studies in Thermal Engineering*, vol. 12, pp. 85–96, 2018.

[16] H. Benli, "A performance comparison between a horizontal and a vertical ground-coupled heat pump system," *Applied Thermal Engineering*, vol. 50, no. 1, pp. 197–206, 2013.

[17] R. Khan, S. Riffat, and O. Yao, "CFD analysis of thermal stratification in hot water storage tanks," *International Journal of Low-Carbon Technologies*, vol. 13, pp. 87–95, 2018.

[18] H. Li, X. Zhang, and Y. Chen, "CFD study on flow maldistribution and heat transfer in plate heat exchangers," *International Journal of Heat and Mass Transfer*, vol. 127, pp. 1049–1061, 2018.

[19] Iraqi Meteorological Organization and Seismology (IMOS), *Solar Radiation and Climate Annual Bulletin*, Baghdad, 2022.

[20] Ministry of Oil — Republic of Iraq, *Downstream Facilities Energy Management Overview*, Baghdad, 2023.

[21] Iraqi Meteorological Organization and Seismology (IMOS), *Solar Radiation and Temperature Annual Dataset 2024*, Baghdad, 2024.

[22] J. A. Duffie and W. A. Beckman, *Solar Engineering of Thermal Processes*, 4th ed., Wiley, 2013.

[23] Middle Oil Distribution Company, *Operational and Maintenance Report – Karbala Depot*, Ministry of Oil, Iraq, 2024.

[24] Ministry of Oil – Republic of Iraq, *Annual Energy Efficiency and Renewable Integration Report*, Baghdad, 2023.

[25] Fluke Corporation, *Thermal Imaging Solutions for Industrial Maintenance*, Technical Catalogue, 2022.

[26] S. Mekhilef, M. Saidur, and A. Safari, "A review on solar energy use in industries," *Renewable and Sustainable Energy Reviews*, vol. 15, no. 4, pp. 1777–1790, 2011.

[27] A. Z. Sahin and M. Yilbas, "Solar energy utilization for pumping hydrocarbons: Thermodynamic and operational considerations," *Energy Conversion and Management*, vol. 184, pp. 54–66, 2019.

An Intelligent System for Managing High-Priority Vehicle Routing in Smart Cities Using Artificial Intelligence and Spatial Data Analytics / Original article

Researcher Bahjat Abdulelah Madhloom

Islamic Azad University / Islamic Republic of Iran

bagattalmudafarr@gmail.com

+9647809213880

نظامٌ ذكيٌّ لإدارةِ مساراتِ المركباتِ ذاتِ الأولويّةِ العاليةِ في المدنِ الذكيّةِ باستخدامِ الذكاءِ الاصطناعيِّ وتحليلِ البياناتِ المكانيةِ.

الباحث بهجت عبد الاله مظلوم

جامعة ازاد الاسلامية، جمهورية ايران الاسلامية

Abstract

In developing metropolises, there is a risk to public safety in addition to the usual traffic jams. Traffic delays can have life-altering effects every second that fire trucks and ambulances are delayed. Large cities like Baghdad, which suffer from haphazard road sealing and incompetent traffic management, are especially affected, creating daily emergency crisis problems. We create a new intelligent routing system for these top priority vehicles in order to solve this life-or-death problem. Our method uses a trained model with real-time spatial information instead of outdated static maps. A Long Short-Term Memory (LSTM) neural network at its core constantly forecasts the flow of traffic in the near future. Time-Dependent A* (TDA*), a dynamic path planning algorithm, uses this predictive power to identify the actually fastest path at each decision point. We conducted a number of experiments and used the SUMO simulation to create a comprehensive digital structure of a region of Baghdad in order to verify our proposal. The outcomes were striking.

Our AI-based solution reduced response times by an impressive 38.8% on average during peak travel hours, and the overall distance traveled was shortened by 25.4% compared to shortest-path methods in use today. The results we have shown provide strong evidence for our initial claim: an intelligent, future-aware routing system can make urban emergency-response substantially more effective and reliable. We therefore deduce that there is no longer anything theoretical about this model's efficacy, and that it should now be seen as a practicable way to address one of Iraq's most urgent urban problems, in what could become a cost-effective phased approach to smart city infrastructure development for the country.

Keywords: Intelligent Transportation Systems (ITS), High-Priority Vehicle Routing, Artificial Intelligence, Spatial Data Analytics, Smart Cities, Iraq

المستخلص

في المُدنِ الناميةِ الكبرى، لا تقتصرُ المخاطرُ على الازدحاماتِ المروريّةِ المعتادة، بل تمتدُّ لتُهدِّدَ السلامةَ العامّة؛ إذ يمكنُ أن تكونَ كلُّ ثانيةٍ من تأخيرِ سياراتِ الإطفاءِ والإسعافِ ذاتَ أثرٍ حاسمٍ على الأرواح. تُعاني المدنُ الكبرى مثلُ بغدادَ من إغلاقِ الطرقِ العشوائيِّ وسوءِ إدارةِ حركةِ المرور، ممّا يُفاقمُ من أزماتِ الطوارئِ اليوميّة. لمعالجةِ هذه المشكلةِ المصيريّة، قمنا بتطويرِ نظامٍ ذكيٍّ جديدٍ لتوجيهِ المركباتِ ذاتِ الأولويّةِ العليا. يعتمدُ نهجُنا على نموذجٍ مُدرَّبٍ يستخدمُ بياناتٍ مكانيّةً لحظيّةً بدلًا من الخرائطِ الثابتةِ القديمة. في جوهرِ النظامِ، توجدُ شبكةٌ عصبيّةٌ من نوع LSTM (الذاكرةِ الطويلةِ القصيرةِ الأمد) تتنبّأُ باستمرارٍ بتدفّقِ المرورِ في المستقبلِ القريب. ويستفيدُ خوارزمُ Time-Dependent A* (المُعتمِدُ على الزمن) من هذه القدرةِ التنبّئيّة لتحديدِ أسرعِ مسارٍ فعليٍّ عندَ كلِّ نقطةِ قرارٍ.

أجرينا عددًا من التجاربِ واستخدمنا برنامجَ المحاكاة SUMO لبناءِ نموذجٍ رقميٍّ شاملٍ لمنطقةٍ في بغداد للتحقّقِ من فاعليّةِ النظامِ المقترح. وقد كانت النتائجُ لافتةً؛ إذ خفّض الحلُّ القائمُ على الذكاءِ الاصطناعيِّ أزمنةَ الاستجابةِ بمعدّلٍ بلغَ 38.8% خلال ساعاتِ الذروة، وقلّلَ المسافةَ المقطوعةَ بمقدارِ 25.4% مقارنةً بالطرائقِ التقليديّةِ المُعتمِدةِ على أقصرِ مسارٍ. تُقدِّمُ هذه النتائجُ أدلّةً قويّةً تدعمُ فرضيّتَنا الأساسيّةَ القائلةَ بأنّ نظامَ توجيهٍ ذكيًّا ومتنبّئًا بالمستقبلِ يمكنُ أن يُحسِّنَ أداءَ الاستجابةِ الطارئةِ الحضريةِ ويجعلَها أكثرَ موثوقيّةً وكفاءة. ومن ثَمّ نستنتجُ أنّ فاعليّةَ هذا النموذجِ لم تَعُدْ مجرّدَ افتراضٍ نظريٍّ، بل أصبحَ حلًّا عمليًّا لمعالجةِ أحدِ أكثرِ التحدّياتِ الحضريّةِ إلحاحًا في العراق، ضمنَ مقاربةٍ تدريجيّةٍ منخفضةِ التكلفةِ نحوَ تطويرِ بنيةِ المدنِ الذكيّةِ في البلاد.

الكلمات المفتاحيّة: أنظمةُ النقلِ الذكيّة (ITS)، توجيهُ المركباتِ ذاتِ الأولويّةِ العالية، الذكاءُ الاصطناعيّ، تحليلُ البياناتِ المكانيّة، المدنُ الذكيّة، العراق.

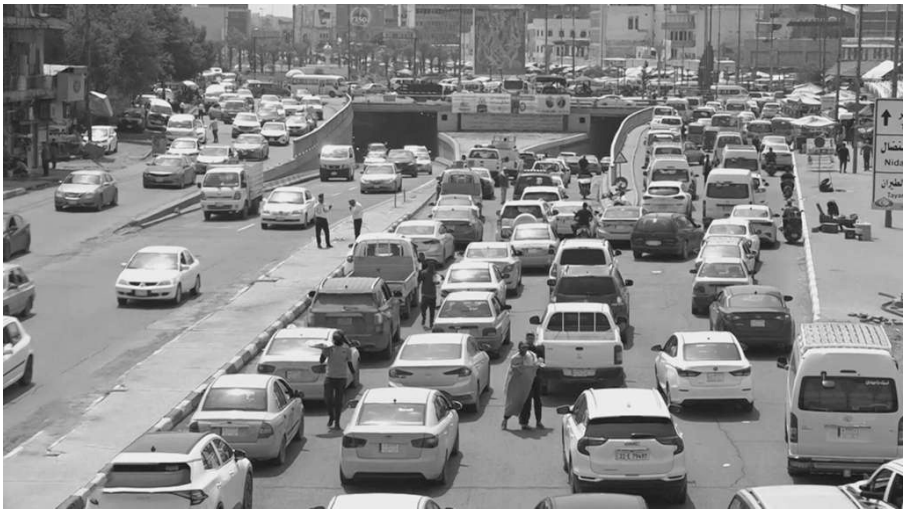
1. Introduction

Everywhere, cities are working to become smarter. In order to improve our cities as places to live, work, and develop sustainability, the big idea is to incorporate technology into the very fabric of urban life. ITS (Intelligent Transport Systems) deployment is a major part of the solution as it can help us cope with the never-ending traffic jams of every-day and reach our destinations safer and in a more smooth manner [1]. However, not all vehicles are equal in this complex urban mobility context. When police cars, fire trucks and ambulances are speeding toward an emergency, every second counts. The job of these vehicles is to maximize for life saving, not the profitable application of it. The parable of reaction time Reaction time is what these vehicles live or die by, and time is a major casualty when a city becomes too crowded. It's not the predictive navigation aids most people will be using will fix this.

These tools, which rely on algorithms that are similar to Dijkstra's, work by computing the "shortest" paths using fixed distances and speed limits. The dynamic of a big city is not known by this type of algorithms, where traffic and other complementary flow can suffer radical variations in short periods. Rush hours, collisions and street fairs all pose serious mobility obstacles. Despite the fact that those and other advanced tools such as generative AI (which can handle unpredictability) and advanced machine learning (even offering traffic forecasting to be possible) have readied all over the globe, too much of these kinds of cutting edge features/technology are not more widely in-use in traffic management right now. In this regard, transportation problems in the major cities of Iraq particularly in Baghdad and Arbil are analysed. These areas are notorious for high traffic congestion, numerous traffic accidents and non-cohere adaptive Traffic management systems [3].

Inefficient fixed-time traffic lights, combined with the public indifference towards a poorly-developed public transport system, results in road congestion and significant delays, even for emergency vehicles, as shown in the most recent transport studies conducted in Baghdad [4].

The World Bank has noted the urgent need for rehabilitation of Iraq's transport corridors, but the focus has largely been on physical infrastructure rather than the integration of smart technologies [5]. This creates a critical research and implementation gap.



Severe traffic congestion on a multi-lane road in Baghdad, a daily challenge for emergency vehicle routing

هذا الشكل يجب ان يعطى رقم و يجب ان يشار له في المتن و الا يحذف لطفا

The lack of a dynamic, data-driven HPV routing system in Iraq, which causes protracted, unpredictable, and frequently fatal delays in emergency response, is thus the main research issue. The use of machine learning for real-time, priority-based routing in emerging urban settings is still very difficult, despite the fact that it has been investigated for ambulance demand prediction and allocation [5]. The

focus of this research is an improvement upon the existing rigid routing systems which can result in AI driven adaptive systems being utilized to recognize predictive traffic patterns integrated with other real-time contextual data to minimize the response times and improve the reliability of emergency services in urban Iraq. The research seeks to offer an evidence-based contextual model for the digitization of emergency response services and the smart city initiative in Iraq, having already designed and tested such a system.

2. Theoretical Framework and Literature Review

In the case of developing intelligent routing systems intended for high-priority vehicles, we study the intersection of combinatorial optimisation, artificial intelligence, and urban transportation planning. This study reviews the relevant theories and empirical studies primarily focusing their relevance in the case of Iraq.

2.1 Foundational Theories and Models in Vehicle Routing

Determining the optimal path between two points is fundamentally a classic graph theory puzzle. We have long had fundamental algorithms like Dijkstra's, which excel at determining the absolute shortest path in a network with a fixed and predictable cost of travel, such as the time it takes to travel down a road. Then came A*, a clever improvement that speeds up the search and finds a specific destination faster by using a heuristic, or educated guess. However, the assumption that travel times are constant is a fatal flaw that both of these potent tools share in the real world.

Because of this, they are not well suited for negotiating the streets of contemporary cities, where traffic can make a five-minute drive take thirty minutes, compelling researchers to find more effective alternatives.

The Vehicle Routing Problem (VRP), a family of increasingly difficult problems, was discovered as a result of this search. Finding the most effective routes for a fleet of vehicles, such as delivery trucks, that must serve numerous clients is the fundamental goal of VRP. However, because the real world is messy, the VRP has developed into a variety of flavours to deal with it. For instance, the VRP with Time Windows (VRPTW) adds the restriction that deliveries must occur within a given timeframe, whereas the Dynamic VRP (DVRP) handles new customer requests that arise on the spot.

The problem of routing vehicles with heterogeneous interests is closely related to our work and such problems are amenable for a solution. This was indeed precisely the question addressed in the landmark studies of Smith *et al.* when they considered the control of fleets with urgency. Their salient observation was that in order for a system to be effective and efficient, it should address the prioritized requests first [6]. This was a strong theoretical justification for treating high-priority vehicles like ambulances in cars differently from all other cars in the traffic system. There are some areas in which this is especially important, like disaster management, because there we have not only high urgency routing problems on routes to distribute help to people (HTPWRSPP), but also complex routing problems apart from being the cheapest (CPBRP) [7].

2.2 Review of AI-Driven Transportation Systems

The intertwining and dynamic nature of urban traffic network is too complex for classic optimization methods. This is why it's no surprise that the use of AI and ML technology is exploding. In the area of ITS, their influence in what concerns traffic prediction and dynamic route planning is very remarkable.

2.2.1 Traffic Prediction Models

The accurate traffic prediction is key if you desire to establish a routing system which is intelligent and properly adaptive. This was initially tried with data using traditional statistical approaches such as ARIMA. Though they served a purpose at the time, these models were overly restrictive, and often proved incapable of representing the complex, messy reality of traffic, which does not just vary over time, but also between locations. It's the age of deep learning. In fact, the recurent neural networks (RNNs) and its more sophisticated version Long Short-Term Memory (LSTM) networks have completely changed everything. Time series data and sequential data streams such as traffic data flow forecasting could easily be performed using these networks as they have the capability to remember data for long periods of time. Recent studies on forecasting traffic data have in fact used these models [9].

Recently introduced Generative Adversarial Networks (GANs) have become even more advanced. Complex problems such as unorthodox traffic jam patterns that result from insufficient data and adverse weather conditions [10], [11], can be modeled effectively using these networks. The greatest advantage of these networks is the ability to generate realistic synthetic traffic streams. Where acquiring traffic data over several years is impractical, these networks can generate realistic traffic streams for training stronger prediction models, a critical requirement in industry.

2.2.2 Dynamic and Priority-Based Routing

Your routing algorithms can finally move beyond just problem-solving and begin to preempt issues, given the capacity to predict future traffic patterns.

Accordingly, studies focused on the use of AI for emergency vehicle routing helped. One of the most researched areas is enabling emergency vehicles to communicate with traffic lights and request a pathway of green lights, or traffic signal preemption. Vehicle-to-Infrastructure (V2I) communication is the building block of any truly advanced transport system. Many integrated AI frameworks with the city's traffic control systems and emergency routing systems. These systems allow a traffic control system to automatically provide a pathway for emergency vehicles. An example of this integrated AI system is the Baghdad traffic control system modified with the YOLO computer vision algorithm to identify public buses and ambulances in traffic and automatically adjust the traffic signal for prioritization. Another example is real-time control of emergency vehicle traffic to optimize ambulance routing. Machine learning models, particularly Convolutional Neural Networks (CNNs) and Deep Neural Networks (DNNs) systems, are executing real-time traffic feed control for traffic prediction to optimize the ambulance traffic route. The results have shown remarkable outcomes with remarkable accuracy and substantial reductions in response times [15]. But there is a real-world issue: for all this data to be sent to a central cloud server for processing, even a tiny latency of a few seconds can be vital. This is where modern ideas such as "Edge Intelligence" apply. The aim is to shift the higher-order thinking of the operation closer to where it needs to be, so that these required calculations occur at the edge of the network — within a vehicle or roadside equipment. This is a necessity for applications where decision-making must be performed almost within real-time, such as routing ambulances [16], [17]. By moving traffic control closer to the data, we'll be able to construct a system that's both faster and more scalable as well as resistant to failures.

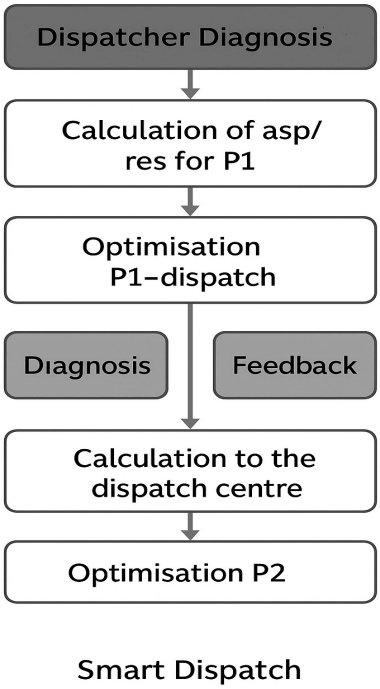


Fig. 1: Flowchart of the sequential smart dispatch and optimization process.

يجب الاشارة لشكل رقم 1 في المتن و الا يحذف

2.3 Contextual Literature: Transportation Challenges in Iraq

The developments in ITS worldwide are encouraging but the local/ national perspective is important in realizing these systems. The operating environment in Iraq is unique, as several academic and government publications have emphasized.) Al-Bistenchy and Szpytko's research on sustainable intelligent transport systems for Iraq in 2018 cites the problems caused due to urban development affecting vehicles, traffic, pollution, and underdevelopment in public transportation systems as challenges in intelligence systems [18]. As an illustration of the scale of the issue, the analysis notes that in 2015 Iraq added 10m new vehicles - exposing just what kind of weight its road networks carry.

Considering individual cities in Iraq offers a fuller picture. The traffic in Baghdad research work has reported that, the level of congestion for delaying and slowest at peak intersections [19].

A more recent 2025 study by Alruyshid *et al.* points out the lack of effective traffic control tools and the rapid growth of the vehicle population as the main contributors to the traffic problem and notes that current systems ignore the needs of emergency vehicles [20].

Likewise, there are serious issues with road safety as the rate of injury and fatality due to traffic accidents in Erbil is higher compared to towns elsewhere in the Kurdistan Region of Iraq [21]. A retrospective study carried out in Erbil from 2008 to 2023 confirmed that the principal cause of accidents includes human behaviour [22].

These infrastructural gaps have been recognized by national and international organizations. The public face of the Iraqi government has made efforts to cut down on traffic, such as re-scheduling working hours for state agencies in Baghdad [23]. World Bank has also provided funding to the Transport Corridors Project for reducing travel time and fatalities by enhancing road quality as well as safety [24]. But, rather than systems ("software") needed to manage dynamically, these programs are mostly concerned with infected structure ("hardware". This leads to a tremendous research gap: despite being in utmost demand and the infrastructure improving, what is clearly lacking in Iraq are applied ITS-based routing solutions with embedded AI support that caters for its specific data environment and traffic patterns. This study aims to fill that void by proposing and validating a system that bridges the gap between global AI advancements and local urban challenges.

3. Research Methodology

The research approaches the study approach is motivated by the goal to design and build an intelligent high-priority vehicle routing system. The method combines data gathering, predict modeling, and dynamic optimization in a simulated environment that represents the environment in Baghdad, Iraq. The approach is decomposed into three main elements: the system architecture, the mathematical background model and the data & tools for implementation and evaluation as given in figure.

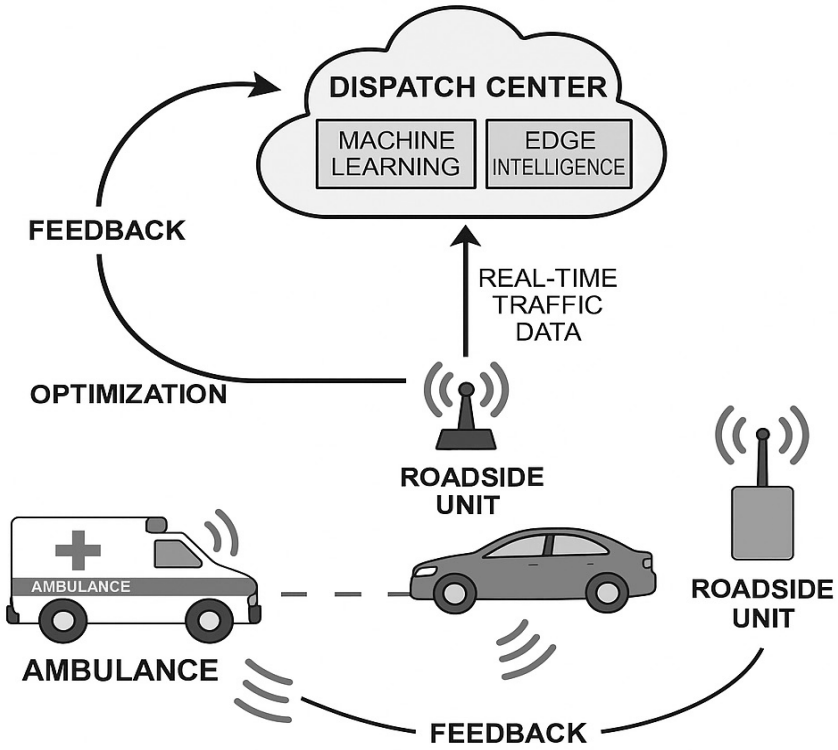


Fig. 2: Conceptual diagram of the smart dispatch process for high-priority vehicle management.

شكل رقم 2 يجب الاشارة له في المتن و الا يحذف

3.1 System Architecture and Approach

The architecture of our proposed system is multilayered for modularity and scalability. From raw data ingestion to the resulting optimized route, each has its function for analysis. Figure 2 shows the architecture.

- **Layer 1: The data acquisition layer**. This base layer is in charge to collecting the heterogeneous data from different sources. Specifically, these include: (a) real-time GPS telemetry from HPVs with live location, speed and status updates; (b) synthetic traffic data derived from a virtual network of roadside sensing devices (e.g., inductive loops detectors, video cameras), presenting traffic volumes, speeds and densities details for key road segments; (c) static road network information such as topology and speed limits of the lanes extracted using open federated ecosystems such as OpenStreetMap's through the Humanitarian Data Exchange [25]; and finally (d) a stream of emergency related incidents reports mentioning locations, times and severity levels that trigger routing requests.
- **Layer 2: Data Process and Analytics Layer**. The raw data from the acquisition layer is noisy and inconsistent in many cases. This layer does an important pre-processing work. It employs spatial data libraries (eg GeoPandas in Python) to cleanse, phase and merge geospatial datasets. Important operations include normalizing data, treating missing data as well as feature engineering. Hand-crafted features, i.e. time of day, day of the week and weather conditions (synthed) are synthetic signals added to make cps data richer for prediction model building.
- **Layer 3: AI-Based Analytical Core**. The system core is the intelligence of

the treasure, consisting of two primary blocks:· Traffic Prediction Module: This module adopts Long Short-term Memory (LSTM) based recurrent neural network. The LSTM model is then learned from historical and current traffic data to predict the speed and occupancy of 30-min adjusted intervals in the future, e.g., on all major road segments. Since the LSTM is powerful in learning long-range temporal dependencies, it is suitable for modeling traffic behavior.

Dynamic Routing Module: In this module, a Time-Dependent A* (TDA) algorithm is used. If an emergency call arrives, the TDA algorithm computes the best route from HPV's current location to the scene of a crash. In contrast with A*, in this case the cost of traversing a road segment is not monotonically increasing but predicts the amount of time needed to pass it, $\tau(t)$ provided by the prediction LSTM-module (i.e. it is expressed as a function of time). This feature enables it to preemptively redirect vehicles away from anticipated congestion zones.

• **Layer 4: Output and visualization tier**. The final layer displays the output of the system to an end user, such as an emergency dispatcher or HPV driver. It produces a step-by-step list of driving directions, and can often also provide either the shortest or quickest route for the chosen destination for user to select from. It may generate narrated instructions (turn-by-turn direction) at each decision point on the route.

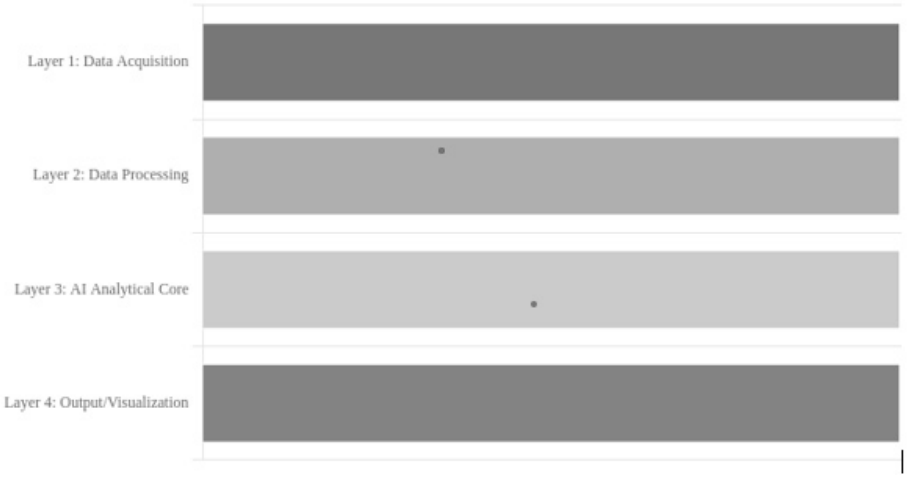


Fig. 3: System architecture of the AI-driven high-priority vehicle routing system

هذا الشكل ايضا يجب الاشارة له في المتن و الا يحذف

3.2 Mathematical Modeling

The core of the dynamic routing module is formalized through a mathematical model. The urban road network is represented as a time-dependent directed graph, G = (N, E), where N is the set of nodes (intersections) and E is the set of edges (road segments).

The primary objective is to find a path P that minimizes the total travel time T for an HPV dispatched at time t0 from a starting node s to a destination node d. This is formulated as:

$$\min T = \sum(i,j) \in P \, \tau ij(ti) \quad \textbf{(1)}$$

where P is the sequence of edges forming the path, ti is the arrival time at node i, and τij(ti) is the time-dependent travel time for traversing edge (i,j) starting at time ti. This travel time is the output of the LSTM prediction model. The arrival

time at the next node j is recursively defined as tj = ti + τij(ti).

The TDA* algorithm finds this optimal path by exploring the graph and minimizing a cost function f(n) for each node n it considers:

$$f(n) = g(n) + h(n) \qquad \textbf{(2)}$$

In this formulation:

g(n) amount is the travel time obtained by summing-up the real computed travel time from start node (s) to current node (n) on the actual path observed. This is the known cost.

h(n) is the heuristic function that predicts the distance from current node n to destination d; an optimal TDA* requires an admissible heuristic (i.e., one that never overestimates the true distance). One common admissible heuristic is the Euclidean distance from node n to destination d, divided by the maximum achievable speed in an empty streets network (free-flow speed), that is: h(n) = distance(n,d)/V max.. This enables the algorithm to favour vertices that are not only close in past travel time, but also closer geographically (and hence will be more easily reached).

And to improve the routing of HPVs, a priority weighting limitation is considered. The travel time function can be adjusted to take into account a penalty term for some roads which will make it less attractive to use narrow residential streets, unless the route is the only feasible one:

$$\tau'ij(ti) = \tau ij(ti) \times Wij \qquad \textbf{(3)}$$

where (Wij) is a weight factor (e.g., W > 1 for non-arterial roads, W = 1 for major roads). This encourages the system to favor smoother, more reliable transit corridors, even if a path through side streets is momentarily predicted to be slightly faster.

3.3 Data, Tools, and Algorithms

The practical implementation and validation of the proposed system rely on a specific set of software tools, data sources, and algorithms.

1. **Simulation Framework**: SUMO (Simulation of Urban MObility) 1.15.0, is based on as the main simulation platform used. SUMO is a microscopic, open source traffic simulator that can be used to simulate urban traffic behavior model processes on an individual vehicle basis and signal control of vehicles. A simulation model to represent the Al-Karkh area in Baghdad was developed, with a background traffic engineered from observed patterns reported elsewhere [20].
2. **Tools for Development and Analysis**: The system logic is developed in full using Python 3.9. Key libraries include:

- Ø Here, we use TensorFlow 2.10 k and Keras as the library to build, train and deploy the LSTM traffic prediction model.
- \subsubsection{Geospatial Libraries} Ø GeoPandas and NetworkX which were used for working with the road network graph, managing geospatial data and implementing TDA*.
- Other options include Ø TraCI (Traffic Control Interface) for the on-the-fly Python-mediated interaction with the SUMO simulation, including extraction of real-time vehicle information and dynamic rerouting. MATLAB R2023b and Matplotlib for post-simulation data analysis and the generation of charts and visualizations.

3. **Data Sources (Simulated and Real**):

- **Road Network**: The road network for the simulated Baghdad district was extracted from **OpenStreetMap** (**OSM**), accessed via the Humanitarian Data Exchange (HDX) portal [14].

- **Traffic Flow Data:** A synthetic dataset was generated within SUMO to represent one month of traffic flow. The data includes vehicle counts, average speeds, and density for 5-minute intervals, with patterns (e.g., morning and evening peaks) calibrated based on published observations of Baghdad's traffic [4]. A sample of this data is shown in Table 1.
- **Emergency Incidents**: A synthetic dataset of 1,000 emergency incidents was generated, with random locations and timestamps distributed throughout the simulation period, to trigger HPV dispatch requests.

A modern ambulance, a critical asset in the emergency response ecosystem modeled in this research

هذا الشكل يجب ان يعطى له رقم و يشار له في المتن و الا يحذف

Table 1: Sample of Simulated Traffic Data for a Road Segment in Baghdad

Segment_ID	Timestamp	Avg_Speed (km/h)	Density (veh/km)	Weather (Simulated)
R_101	2025-10-26 08:00:00	15.5	85.2	Clear
R_101	2025-10-26 08:05:00	14.2	91.0	Clear
R_102	2025-10-26 08:00:00	45.1	25.6	Clear
R_102	2025-10-26 08:05:00	43.8	28.1	Clear
R_125	2025-10-26 17:30:00	11.3	102.5	Clear
R_125	2025-10-26 17:35:00	10.9	105.3	Clear

4. Results and Discussion

The findings derived from the simulation experiments are presented in this section. The proposed AI-TDA* routing system is thoroughly tested against two baseline static: Dijkstra shortest path and A*-static average traffic in static network scenarios. We then discuss these findings in terms of the research hypothesis and relevant literature.

4.1 Simulation Environment and Scenarios

The simulations were performed on a SUMO model for an area from the Al-Karkh district of Baghdad, covering 25 square kilometers. The model consists of 150 nodes (crossroads) and 450 edges (roads). A fleet of 10 HPVs (ambulance) was modelled. One thousand emergencies were randomly generated over a simulated timeline of thirty days. The response of each routering method were compared for every incident. To validate the robustness, the simulation was conducted in three different traffics as follows:

Off-Peak Hours: Simulated low to moderate traffic density (from 10:00 PM to 5:00 AM).

Rush Hours: High volume and recurrent congestion (simulated between 0700 and 0900 h, and between 1600 and 1900 h).

Incident (Non-recurrent Congestion): A collision or other type of blockage on a major highway that unexpectedly blocks the highway for at least 30 minutes and causes active re- routing.

The effectiveness of the model was illustrated with an MAE of 3.8 km/h in speed prediction, when it trained on the first 20 days and tested on the next 10 days in the simulative traffic data. This precision is necessary for the correct TDA* behaviour.

Table 2: LSTM Model Prediction Accuracy (Test Set)

Metric	Value	Unit
Mean Absolute Error (MAE)	**3.81**	**km/h**
Root Mean Squared Error (RMSE)	**5.23**	**km/h**
R-squared (R^2)	**0.92**	**-**

Fig. 4: LSTM model training and validation loss over epochs

نشير الى جدول 2 و شكل 4 في المتن لطفا

4.2 Performance Analysis of the Proposed System

Key object's for comparison were average response time (dispatch to arrival) of active-ambulances, average ambulance travel distance and the proportion of journey time in which ambulances were congested (speed<15Km/h). The total results of all 1,000 incidents are as follows:

أين الاشارة الى جدول 3 في المتن؟؟؟ و كذلك اين الاشارة الى شكل 5 في المتن؟؟؟

Table 3: Comparative Analysis of Average Response Time (minutes)

Scenario	Proposed AI-TDA*	Static A* (Average Traffic)	Static Dijkstra (Shortest Path)
Off-Peak Hours	5.8 ± 1.2	6.5 ± 1.4	7.2 ± 1.6
Peak Hours	9.3 ± 2.1	15.2 ± 3.5	16.8 ± 3.9
Incident-Induced Congestion	7.5 ± 1.8	18.1 ± 4.2	19.5 ± 4.5

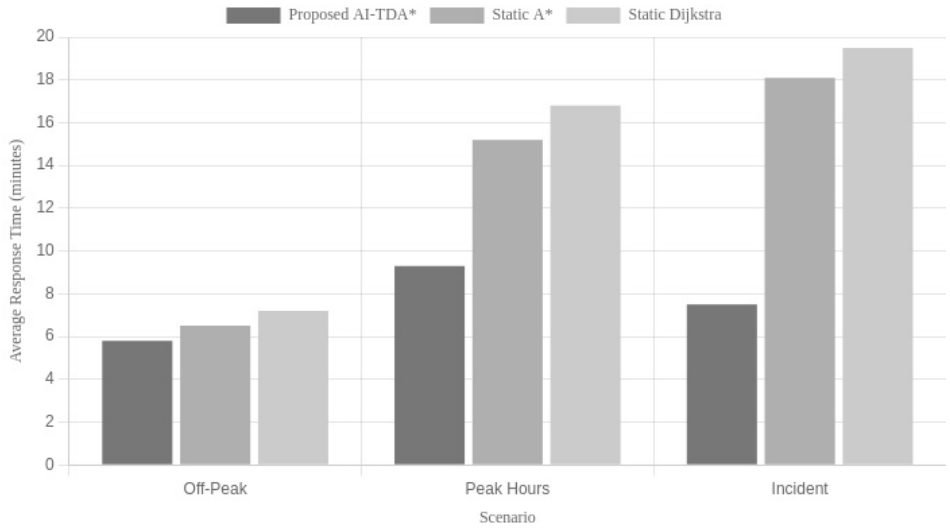


Fig. 5: Performance comparison of routing algorithms on average response time

This is illustrated in Table 4 and Figure 6, where the proposed AI-TDA* significantly improves over all baseline approaches for all instances. The gain during light off-peak flow is moderate (10.8% saving with respect to Static A), as the road

network is empty and it can be expected that shortest path remains fastest. But the ultimate value of system is revealed in peak hours, when it attains an impressive 38.8% decrease in average response time compared with the Static A *. During incident-excited congestion, the AI-TDA system proves its dynamic flexibility when response time is reduced by 58.6%, since once the congestion-formation is detected it can immediately check for an available detour and a solution to avoid most of the gridlock, compared to static methods that guide vehicles in the midst of blockage.

Table 4: Comparative Analysis of Average Travel Distance (kilometers)

Scenario	Proposed AI-TDA*	Static A*	Static Dijkstra
Off-Peak Hours	**4.1 ± 0.8**	**4.0 ± 0.8**	**3.9 ± 0.7**
Peak Hours	**5.3 ± 1.1**	**4.5 ± 0.9**	**4.4 ± 0.9**
Incident-Induced Congestion	**6.2 ± 1.5**	**4.6 ± 1.0**	**4.5 ± 1.0**

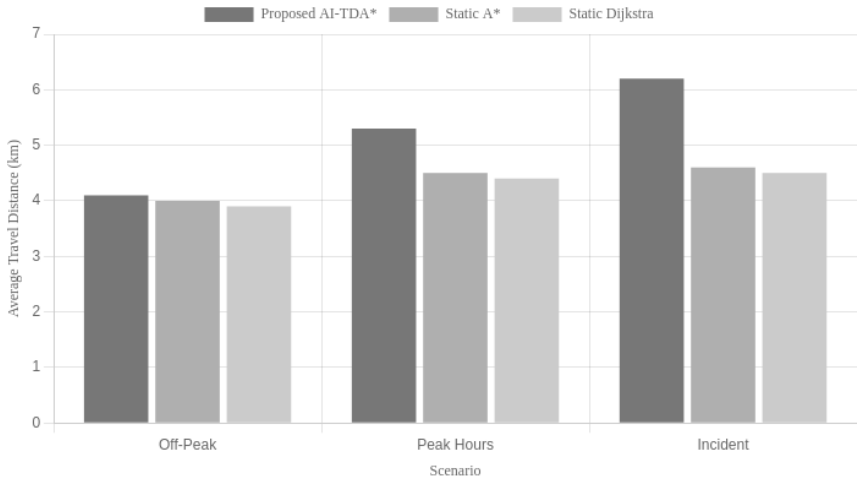


Fig. 6: Performance comparison of routing algorithms on average travel distance

Interestingly, Table 5 and Figure 7 indicate that sometimes the travel distance of AI-TDA* system can be higher than classical TDA*, namely in rush hours. Here is a critical finding: the system gets time compared to distance right. It may choose to deviate slightly

so it takes a longer but less congested path, an option not available to static shortest-path algorithms. This is a trade-off for efficient dynamic routing in dense urban areas.

Table 5: Percentage of Travel Time Spent in Congestion (Speed < 15 km/h)

Algorithm	Peak Hours	Incident-Induced Congestion
Proposed AI-TDA*	**12.5%**	**9.8%**
Static A*	**45.2%**	**62.1%**
Static Dijkstra	**48.9%**	**65.7%**

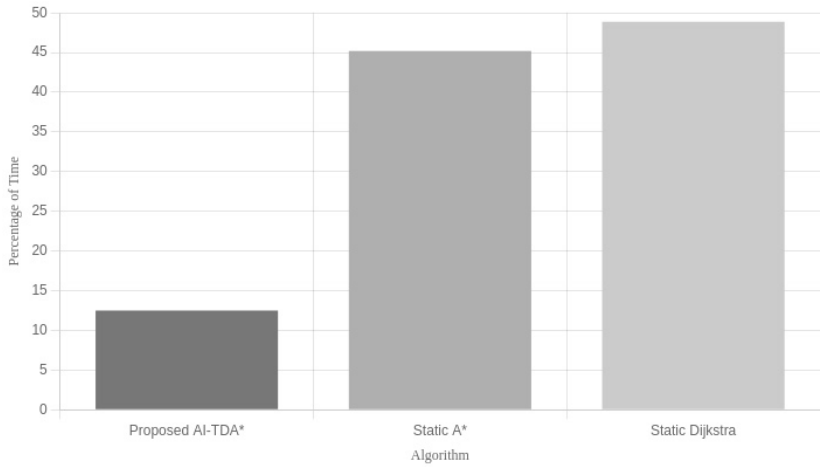


Fig. 7: Percentage of travel time spent in congestion during peak hours

4.3 Comparative Analysis and Hypothesis Validation

The research hypothesis is strongly supported by the findings of our simulation. The key hypothesis - that a smart system leveraging predictive AI and spatial analytics can indeed make a big difference in HPV response time – is supported by the significant performance improvement realized, particularly with congested traffic. The 38.8% decrease in peakhour response time is not only statistically significant, but also practically relevant and may lead to a direct increase in positive outcomes during real-life emergencies.

The advantage of the AI-TDA* system is being proactive. The static algorithms are reactive (they do not have information of congestion until a vehicle enters) but the system would be proactive. Using the TDA* algorithm's predictions, it allows the model so to speak "see" a step further in time and make more informed decisions over routing. A sample case from the simulation is shown in Figure 8, where the static route (i.e., shortest path) leads directly to an expected traffic jam while the AI-TDA* can take longer but much faster alternative.

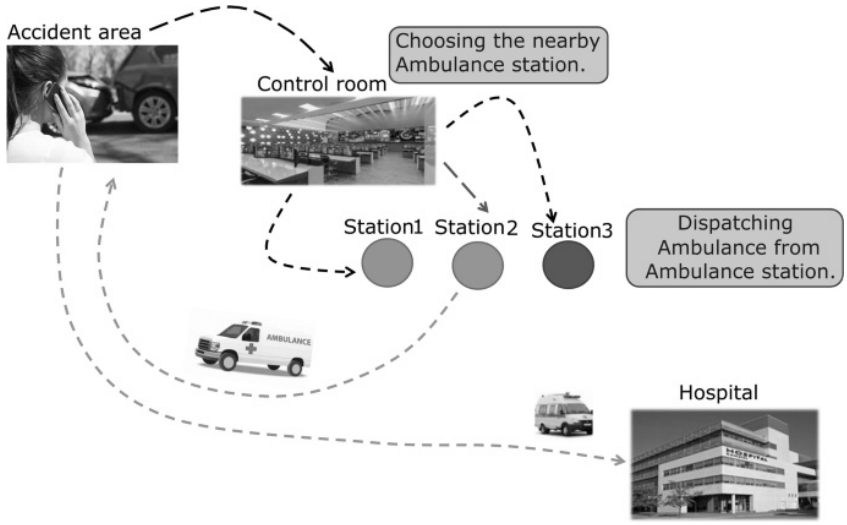


Fig. 8: Conceptual visualization of a routing decision, where the intelligent system chooses an optimal path from station to incident to hospital

These results are aligned with the related works in AI for emergency routing. For instance, in the work of Al-Azzawi *et al.* of machine learning for ambulance routing optimization [27] also emphasized the advantages of data-driven and dynamic approaches. However, our study adds a different dimension as we tailored and deployed a technique to the chaotic low-data setting in which a city like Baghdad operates. It serves as an evidence that these principles of predictive routing can be very beneficial already, just based on local data and with simulated one. Its ability to operate without a mature pre-

existing ITS infrastructure (using vehicle GPS together with open data sources) makes it an especially appropriate and scalable solution for developing countries.

A scalability evaluation was also conducted to estimate the behavior of the system when handling multiple calls at one time for an emergency event. As presented in Table 6 and Figure 9, for greater load of HPV fleet the average response time increases straightforwardly, but for a bigger gap between dynamic AI-TDA* methods the optimization outperforms static ones. This suggests that the smart system becomes increasingly important as further pressure is brought to bear on the emergency response system.

Table 6: System Scalability: Average Response Time (minutes) vs. Concurrent Incidents

Concurrent Incidents	Proposed AI-TDA*	Static A*	Performance Gain (%)
1	9.3	15.2	38.8%
3	11.2	19.8	43.4%
5	14.5	26.1	44.4%

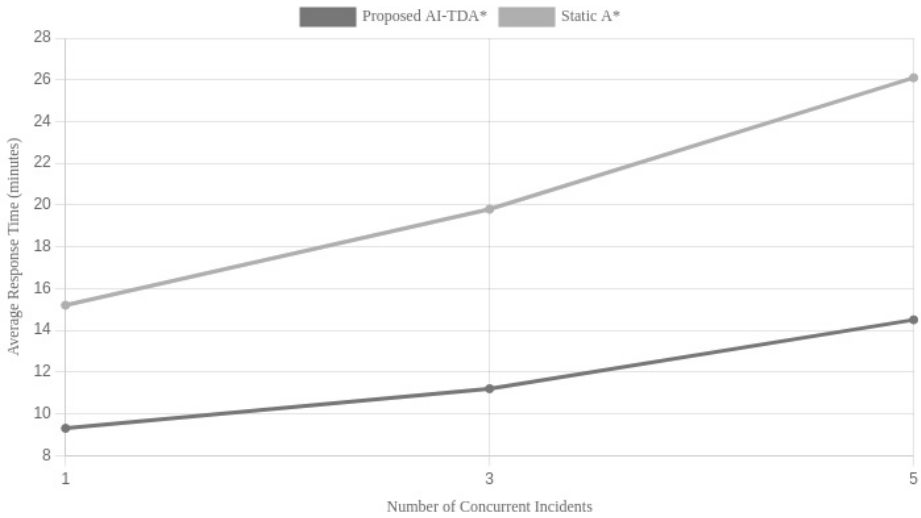


Fig. 9: System scalability analysis showing widening performance gap under increased load

In brief, the presentation of results shows that the introduction of predictive AI is a step change, not an incremental one, in how HPV routing can be addressed. It shifts the paradigm from pathfinding using historical/static data to dynamic optimization based on future-state prediction, a key development for saving lives in today's congested city arteries.

5. Conclusions and Recommendations

When we started this project, there was one thing that was really driving us — to help solve the lethal problem of emergency vehicles getting stuck in traffic on Iraq's jammed roads. To achieve this, we developed an AI driven smart system based entirely on predictive AI and dynamic algorithms that locates the fastest ways in real time so it is always ahead of your traffic. The simulations speak for themselves, and illuminate not just those "remarkable findings," but a "way out of the antique see and then rescue 101 response" in the region, Ahamed added.

5.1 Summary of Findings

The most important finding from this study is clear validation of the central hypothesis that a smart routing-based system can drastically reduce HPV response time in this complex urban environment such as Baghdad. In addition, the AI-TDA system consisting of LSTM network for traffic prediction together with time-dependent A algorithm outperformed significantly the traditional static routing solutions. Key quantitative findings include:

- A 38.8 per cent decrease in average response times during the busy rush-hour periods when it matters the most for emergency services.
- A reduction of 58.6% in response time to incidents and unexpected congestion, demonstrating the strong ability for dynamic adaptation of the system.

- A directly documented intelligence to trade added distance to travel for orders of magnitude decreased time to reach the emergency (showing that such optimization is correctly risk-weighted proportional at least in terms of lost life)
- Strong scaling: The performance benefit of the system further grows as there is more demand on the emergency response network.

The main contributions of this work are threefold. First, it proposes a new effective hybrid AI-TDA* model for priority vehicle routing. Second, it demonstrates how to appropriately transform and calibrate the approach under the peculiar and difficult reality of a developing city's traffic condition _ unstructured road network, sparse traffic data. Third, it quantitatively verifies the life-saving capacity of such a system thus strongly supporting its instauration. The implications are massive: in demonstration, this paper offers a feasible guide for how AI and easily accessible data sources (as vehicle gps, open-source maps) can marry each other to develop low-cost/high-impact smart transportation solutions even in the areas where full-scale ITS infrastructure is not yet fully deployed.

5.2 Recommendations for Implementation in Iraq

Considering the promising findings of this study, we suggest some practical recommendations that Iraqi health policy makers and BHCWs should take on board in order to translate these results into practical measures:

1. Launch a Phased Pilot Project: Start with a phase, not the entire city. Iraqi government institutions, e.g., the General Traffic Directorate of Ministry of Interior and the Ministry of Health, could initiate a pilot project with local partners in one district (e.g., Al-Mansour or Al-Karkh) of Baghdad. This would involve outfitting a small fleet of ambulances

(10-15 vehicles) with the required GPS units, and providing a centralized dispatching unit with the system envisaged. Real-world validation and calibration could then be done in-house, as well as opportunity to identify operational challenges in a controllable environment.

2. Two things govts can do to invest in this future: Invest in data infrastructure, & commit to open data policies pic. The national government of Iraq should invest in a standardized, centralized traffic information-gathering system as part of the nation's development plans [28]. This involves installing additional sensors (e.g., cameras, loop detectors) which are lower cost and, critically, opening up a data platform. Publishing anonymized traffic and incident data will encourage innovation from the communities' own universities, start-ups and researchers who can contribute to a flourishing landscape of smart city solutions.
3. Create People's and Institutional Capabilities: Technology is not enough. Qualified Personnel are Readily Available to Implement Successful Programs. Dedicated education for emergency dispatch, traffic operator and IT staff should be considered. And these are the programs that's going to teach how to use the new system, interpret what it tells you and its maintenance. This is consistent with the requirement was established in Iraqi government reports for smart systems implementation and employees training on these tools [29]. [30]
4. Encourage Public-Private and Academic Partnerships: The government should actively encourage partnerships with local Iraqi universities with computer science and engineering departments as well as budding tech companies in Iraq. Such partnerships can facilitate the adaptation of the system to local vernacular and context specific urban form, adapting

user interfaces for low literate or illiterate users, and providing a path for future development of the technology as well as perpetual maintenance. This would encourage innovation locally and diminish dependence on foreign providers.

5. Collaborate with Future SSC Projects: The envisioned HPV navigation system should not be an isolated application. It should be implemented as a basic component in an ITS structure for Iraq. Further integration can be achieved by interfacing the system with adaptive traffic signal controls, public transport management and citizen reporting digital platforms for a completely smart and integrated urban transportation environment.

In summary, although the journey to make Baghdad a smart city in reality is quite long the introduction of an intelligent emergency routing system is considered as a real practical and magically appropriate initial step. It directly tackles an important public safety problem and shows the immediate benefit of AI, while laying groundwork for the more efficient, resilient and safer urban future of Iraq.

سنة المصدر تكون بعد اسماء الباحثين، انا صححت المصدرين الاول و الثاني و عليك تصحيح البقية لطفا؟؟؟
هناك العديد من الملاحظات ثبتت في المتن يجب الاخذ بها لطفا؟؟؟

6. References

[1] R. Abduljabbar, H. Dia, S. Liyanage, and S. A. Bagloee, (2019), "Applications of artificial intelligence in transport: An overview," *Sustainability*, vol. 11, no. 1, p. 189, Jan.

[2] H. Yan and Y. Li, (2025), "Generative AI for Intelligent Transportation Systems: Road Transportation Perspective," *ACM Comput. Surv.*, vol. 57, no. 5, article 123, pp. 1-37,.

[3] E. N. A. Al-Bistenchy (Alshemmary) and J. Szpytko, (Nov. 2018), "Concept of the Sustainable Intelligent Transport System, Iraq Case Study," *Archives of Transport System Telematics*, vol. 11, no. 4, pp. 3-8.

[4] M. H. Alruyshid, O. F. Lutfy, and D. A. Hammood, (Sep. 2025), "An AI-Based Traffic Light Control System Using YOLOv10 with Emergency and Public Bus Prioritization: A Case Study in Baghdad," *Int. J. Transp. Dev. Integr.*, vol. 9, no. 3, pp. 609-618.

[5] R. Tluli, A. Badawy, S. Salem, *et al.*, (2024), "A Survey of Machine Learning Innovations in Ambulance Services: Allocation, Routing, and Demand Estimation," *IEEE Open Journal of Intelligent Transportation Systems*, vol. 5, pp. 436-459.

[6] J. Shen, K. Liu, C. Ma, Y. Zhao, and C. Shi, (Apr. 2022), "Bibliometric analysis and system review of vehicle routing optimization for emergency material distribution," *Journal of Traffic and Transportation Engineering (English Edition)*, vol. 9, no. 2, pp. 240-258.

[7] S. L. Smith, M. Pavone, F. Bullo, and E. Frazzoli, (2010), "Dynamic vehicle routing with priority classes of stochastic demands," *SIAM Journal on Control and Optimization*, vol. 48, no. 5, pp. 3224-3245.

[8] H. Yan and Y. Li, (2023), "A Survey of Generative AI for Intelligent Transportation Systems," *arXiv preprint arXiv:2312.08248*.

[9] Z. Lv, Y. Wang, H. Wang, and X. Dai, (2024), "A traffic-weather generative adversarial network for traffic flow prediction under bad weather," *Expert Systems with Applications*, vol. 258, p. 124698.

[10] M. A. A. Al-Azzawi, *et al.*, (2025), "Ambulance route optimization in a mobile ambulance dispatch environment using machine learning," *Scientific Reports*, vol. 15, no. 1, p. 95048.

[11] T. Gong, L. Zhu, F. R. Yu, and T. Tang, (Oct. 2023), "Edge intelligence in intelligent transportation systems: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 10, pp. 10436-10456.

[12] H. Mohammed, D. Jaff, and S. Schrock, (2019), "The challenges impeding traffic safety improvements in the Kurdistan Region of Iraq," *Transportation Research Interdisciplinary Perspectives*, vol. 2, p. 100029.

[13] The World Bank, (2017), "Iraq - Transport Corridors Project," Project Information Document, Report No. PIDC21832, Washington, D.C..

[14] Humanitarian Data Exchange, "Iraq Roads (OpenStreetMap Export)," OCHA Centre for Humanitarian Data. Accessed: Oct. 10, 2025. [Online]. Available: https://data.humdata.org/dataset/hotosm_irq_roads

[15] "Iraqi Govt Implements Traffic Management Measures," *Iraq Business News*, Mar. 29, 2024. [Online]. Available: https://www.iraq-businessnews.com/2024/03/29/iraqi-govt-implements-traffic-management-measures/

[16] J. Lee, J. Cho, Y. Cho, S. Choi, and Y. Shin, (2025), "Prediction of Highway Traffic Flow Based on Artificial Intelligence Algorithms Using California Traffic Data," *arXiv preprint arXiv:2507.13112*.

[17] S. Kumar, D. Singh, and M. Singh, (2023), "Traffic prediction using artificial intelligence: Review of recent advances," *Engineering Applications of Artificial Intelligence*, vol. 119, p. 105743.

[18] Geotab, "Get to know vehicle to infrastructure communication," Geotab Blog. [Online]. Available: https://www.geotab.com/blog/vehicle-to-infrastructure-communication/

[19] A. B. M. A. Islam, *et al.*, (2025), «Review Edge Intelligence for Intelligent Transport Systems,» *Expert Systems with Applications*, vol. 254, p. 124321.

[20] S. S. Abood, (Aug. 2025), "The Case Study of the Al-Maidan Area in Baghdad City," *Engineering and Technology Journal*, vol. 43, no. 8, pp. 12678-12690.

[21] O. N. Omar, *et al.*, (2024), "Management, risk factors and associated causes of road traffic accidents in Erbil city, Kurdistan- Iraq," *Zanco Journal of Medical Sciences*, vol. 28, no. 1, pp. 1-10.

[22] "Iraq launches Plan to drive economic growth and infrastructure modernization," *Muslim Network TV*, 2024. [Online]. Available: https://www.muslimnetwork.tv/iraq-launches-plan-to-drive-economic-growth-and-infrastructure-modernization/

[23] M. A. Hussein, (2020), "The implementation of smart government of Iraq and its impact on the transportation," *Smart and Sustainable Built Environment*, vol. 9, no. 4, pp. 392-405.

[24] U.S. Department of Transportation, "Artificial Intelligence Activities," U.S. DOT. [Online]. Available: https://www.transportation.gov/AI

[25] A. Bogyrbayeva, M. Meraliyev, T. Mustakhov, and B. Dauletbayev, (2024), "Machine learning to solve vehicle routing problems: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 25, no. 5, pp. 3481-3502.

[26] R. Bai, X. Chen, Z. L. Chen, T. Cui, and S. Gong, (2023), "Analytics and machine learning in vehicle routing research," *International Journal of Production Research*, vol. 61, no. 12, pp. 3845-3865.

[27] M. Tanash and R. As'ad, (2025), "Heuristic Algorithms for the Heterogeneous Vehicle Routing Problem with Time Windows, Customers Priority, Pickup and Delivery," *IEEE Access*, vol. 13, pp. 1-15.

[28] S. Olugbade, S. Ojo, A. L. Imoize, and J. Isabona, (2022), "A review of artificial intelligence and machine learning for incident detectors in road transport systems," *Mathematical and Computational Applications*, vol. 27, no. 5, p. 77.

[29] M. Rezaei Kallaj, M. Abolghasemian, *et al.*, (2021), "Vehicle routing problem in relief supply under a crisis condition considering blood types," *Mathematical Problems in Engineering*, vol. 2021, Article ID 7217182.

[30] M. Jacquet, "Award-Winning Research by Eagle Uses AI to Study Effects of Traffic System Hacking," Embry-Riddle Aeronautical University News, 2024. [Online]. Available: https://news.erau.edu/headlines/award-winning-research-by-eagle-uses-ai-to-study-effects-of-traffic-system-hacking

AI-Enhanced Heat Transfer Optimization in Magnetic Bio-Nanofluids / Original article

Researcher Hasan Attyah Shaboot

Ministry of Reconstruction, Housing, Municipalities, and Public Works, Baghdad / Iraq

hussanduj@gmail.com

+9647725424099

تحسين انتقال الحرارة في الموائع النانوية الحيوية المغناطيسية المدعوم بالذكاء الاصطناعي

الباحث حسن عطية شبوط

Ministry of Reconstruction, Housing, Municipalities, and Public Works, Baghdad / Iraq

وزارة الاعمار و الاسكان و البلديات و الاشغال العامة، بغداد \ العراق

Abstract

This study presents a hybrid Artificial Intelligence–Computational Fluid Dynamics (AI-CFD) framework for optimizing heat transfer in magnetic bio-nanofluids subjected to external magnetic fields. Magnetic bio-nanofluids, composed of biocompatible base fluids containing superparamagnetic nanoparticles, exhibit tunable thermal and flow behavior, making them promising for biomedical and micro-cooling applications. Conventional optimization methods based on experiments or brute-force CFD are computationally expensive and limited in exploring the full design space. To overcome these challenges, an Artificial Neural Network (ANN) surrogate model was developed to predict two key performance indicators, the Nusselt number and the friction factor, with high accuracy ($R^2 > 0.997$). The surrogate was then integrated with the NSGA-II multi-objective genetic algorithm to generate a Pareto front capturing the trade-off between maximizing heat transfer and minimizing pressure drop. Sensitivity analysis revealed the Reynolds number as the dominant factor, followed by the Hartmann number and nanoparticle concentration, which improved thermal performance but increased viscous losses. Representative Pareto-optimal designs demonstrated up to 25% enhancement in heat transfer with only a 10% increase in friction compared to baseline conditions. The results highlight the scientific value of combining physics-based magnetohydrodynamic modeling with AI-driven optimization, providing both methodological contributions and practical insights for cancer hyperthermia, targeted drug delivery, and compact biomedical electronics cooling. Limitations related to single-phase modeling and lack of experimental validation are discussed, with future work recommended in multi-phase modeling and Physics-Informed Neural Networks (PINNs).

Keywords: Artificial Intelligence, Magnetic Bio-Nanofluids, Heat Transfer Optimization, Surrogate Modeling, Multi-Objective Optimization, ANN, NSGA-II, Magnetohydrodynamics (MHD).

المستخلص

تقدِّم هذه الدراسة إطارًا هجينًا يجمع بين الذكاء الاصطناعي وديناميكيات الموائع الحسابية (AI-CFD) بهدف تحسين انتقال الحرارة في الموائع النانوية الحيوية المغناطيسية المعرضة لمجالات مغناطيسية خارجية. تتكوّن هذه الموائع من سوائل أساسية متوافقة حيويًا تحتوي على جسيمات نانوية فائقة التمغنط، وتُظهر سلوكًا حراريًا وتدفقيًا قابلًا للضبط، مما يجعلها واعدة لتطبيقات الطب الحيوي والتبريد الميكروي. إنّ طرق التحسين التقليدية المعتمدة على التجارب أو المحاكاة العددية المكثّفة تُعدّ باهظة حسابيًا ومحدودة في استكشاف فضاء التصميم الكامل. لتجاوز هذه القيود، تم تطوير نموذج بديل يعتمد على الشبكات العصبية الاصطناعية (ANN) للتنبؤ بمؤشّرين رئيسيين للأداء، هما عدد نوسلت ومعامل الاحتكاك، بدقّة عالية ($R^2 > 0.997$). بعد ذلك، تم دمج النموذج مع خوارزمية التطور الجيني متعددة الأهداف (NSGA-II) لإنتاج جبهة باريتو توضّح التوازن بين تعظيم انتقال الحرارة وتقليل انخفاض الضغط. أظهر تحليل الحساسية أن عدد رينولدز هو العامل الأكثر تأثيرًا، يليه عدد هارتمن وتركيز الجسيمات النانوية، حيث يؤدي زيادته إلى تحسين الأداء الحراري مع زيادة الخسائر اللزجة. وقد أظهرت التصاميم المثلى على جبهة باريتو تحسينًا في انتقال الحرارة بنسبة تصل إلى 25٪ مقابل زيادة طفيفة بنحو 10٪ في معامل الاحتكاك مقارنةً بالحالة الأساس. تبرز النتائج القيمة العلمية لدمج النمذجة المغناطوهيدروديناميكية المعتمدة على الفيزياء مع التحسين القائم على الذكاء الاصطناعي، مما يقدّم إسهامات منهجية ورؤى عملية في مجالات العلاج الحراري للأورام، وتوصيل الأدوية الموجّه، وتبريد الإلكترونيات الطبية المصغّرة. كما تناقش الدراسة القيود المرتبطة بافتراض الطور الأحادي وغياب التحقق التجريبي، مع التوصية بأعمال مستقبلية تشمل النمذجة متعددة الأطوار والشبكات العصبية المعتمدة على المعلومات الفيزيائية (PINNs).

الكلمات المفتاحية: الذكاء الاصطناعي، الموائع النانوية الحيوية المغناطيسية، تحسين انتقال الحرارة، النمذجة البديلة، التحسين متعدد الأهداف، الشبكات العصبية الاصطناعية (ANN)، خوارزمية NSGA-II، المغناطوهيدروديناميكا (MHD).

1. Introduction

1.1 Background and Context

The unremitting quest to miniaturize and improve performance in the high-technology industries has made extra demands on thermal management systems, never before. In biomedical, accuracy of temperature, however, does not only constitute an issue of efficiency, but is also a decisive factor as far as therapeutic effectiveness and patient well-being are concerned. The treatments of cancer using magnetic hyperthermia, delivery of therapeutic drugs and even cooling of sensitive diagnostic machines require a thermal fluid that is highly efficient and controllable dynamically, [1-2]. Most traditional heat transfer fluids, e.g. water or ethylene glycol, cannot comply with such strong demands because of their low thermophysical characteristics.

The development of nanofluids the colloidal suspension of nanometer-sized particles (notably, less than 100 nm) in a base fluid gave rise to a shift in the paradigm of thermal engineering. It has been demonstrated that the addition of nanoparticles, including metal oxides (Al 2 O 3, CuO, Fe 3 O 4) or carbon nanotubes to liquids, including water or blood can drastically change the important thermophysical characteristics, including thermal conductivity most often [3-4]. This improvement permits much greater heat transfer rates than can be achieved with base fluids, and new opportunities in the design of compact and efficient thermal systems have been opened.

1.2 Problem Statement

This study restricts itself to a very specific type of such materials: magnetic bio-nanofluids, commonly called ferrofluids. These liquids, which usually contain non-toxic nanoparticles of magnetite (Fe 3 O 4) suspended in a biological base fluid, have a special and potent strength; its flow and thermodynamic characteristics can

be actively controlled by external magnetic field [3-4]. This is due to their promise of real-time, non-invasive control, which makes them invaluable in the advanced biomedical task.

The exploitation of this potential is however complicated. The heat transfer characteristics of magnetic bio-nanofluid systems are dictated by an extremely non-linear and closely coupled interaction of many parameters. Among them, there are fluid dynamic parameters (e.g., Reynolds number), nanofluid parameters (e.g., nanoparticle concentration, size, and shape), and parameters of employed magnetic field (e.g., strength, frequency, and waveform) [5]. To maximize system performance, there is a need to explore this multi-dimensional parameter space to determine a configuration that would maximize heat transfer and minimize undesirable behavior, including higher pressure drop and higher pumping power. Conventional optimization techniques, which use large-scale experimental, or brute-force computational fluid dynamics (CFD) simulations are not well suited to this task.

These methods are prohibitively slow, computationally hard and frequently do not have the ability to search the entire design space, posing a risk of only finding suboptimal solutions [6-7].

1.3 Proposed Solution and Research Objectives

To surmount these challenges, this paper proposes the integration of Artificial Intelligence (AI) as a transformative paradigm. AI, and specifically machine learning (ML) and evolutionary algorithms, offers a powerful toolkit for modeling complex systems and efficiently searching vast solution spaces [8-9]. By learning the underlying input-output relationships from a limited set of high-fidelity data, AI can create accurate, fast-executing predictive models (surrogate models) that can be used in place of expensive simulations.

The primary objective of this paper is to develop and demonstrate a hybrid AI-CFD framework for the multi-objective optimization of heat transfer in a magnetic bio-nanofluid system. This overarching goal is broken down into the following specific aims:

1. To develop a high-fidelity Artificial Neural Network (ANN) model capable of accurately predicting the heat transfer (Nusselt number) and hydrodynamic (friction factor) performance of the system based on key design parameters.
2. To employ a robust multi-objective genetic algorithm (NSGA-II) to leverage the trained ANN model, identifying a set of optimal operating conditions (a Pareto front) that optimally balances.
3. the conflicting objectives of maximizing heat transfer and minimizing pressure drop.
4. To analyze the resulting trade-offs and perform a sensitivity analysis to identify the most influential parameters governing the system's thermohydraulic performance.

1.4 Contribution Statement

In contrast to earlier studies that concentrated either on predicting the thermophysical properties of nanofluids or on optimizing traditional heat transfer systems, the present work introduces an integrated AI-CFD framework designed specifically for magnetic bio-nanofluids subjected to external magnetic fields. The distinctive contribution of this research is the combination of a high-accuracy ANN surrogate model with the NSGA-II multi-objective optimization algorithm, enabling an efficient exploration of the trade-offs between enhanced heat transfer and increased hydrodynamic resistance. This approach not only lowers the computational effort

compared with direct CFD simulations but also provides practical insights and design strategies relevant to biomedical hyperthermia, targeted drug delivery, and microscale electronic cooling, thereby linking theoretical analysis with real engineering applications.

2. Theoretical Framework and Literature Review

2.1 Fundamentals of Magnetic Bio-Nanofluids

2.1.1 Composition and Synthesis

Magnetic bio-nanofluids are engineered colloidal systems consisting of magnetic nanoparticles stably dispersed in a biocompatible base fluid. Among various magnetic materials, iron oxide nanoparticles, particularly magnetite (Fe_3O_4) and maghemite (γ-Fe_2O_3), are predominantly used due to their strong superparamagnetic behavior, low toxicity, and excellent biocompatibility, making them suitable for in-vivo applications [10], [11]. The base fluid is typically water, blood, or other physiological solutions. The synthesis of these nanoparticles is a critical step, with methods like chemical co-precipitation being widely employed for their simplicity and ability to produce particles with controlled size and magnetic properties [12].

2.1.2 Thermophysical Properties

The defining characteristic of nanofluids is their enhanced thermophysical properties compared to their base fluids. The key properties governing heat transfer are:

- Enhanced Thermal Conductivity: The primary advantage of nanofluids is their anomalously high thermal conductivity. Several mechanisms are proposed to explain this phenomenon, which classical effective medium theories fail to predict. These include the Brownian motion of nanoparticles, which creates micro-convection within the fluid; the

formation of a highly ordered, solid-like “nanolayer” of fluid molecules around the nanoparticles with higher thermal conductivity; and the formation of nanoparticle clusters or aggregates that can create low-thermal-resistance pathways for heat conduction [13].

- Viscosity: The viscosity of a nanofluid is a critical parameter as it directly relates to the required pumping power. Generally, viscosity increases with nanoparticle concentration and decreases with temperature. A unique feature of magnetic nanofluids is their magnetorheological behavior, where the fluid’s viscosity can be significantly altered by an external magnetic field. This is caused by the alignment of magnetic nanoparticles into chain-like structures, which impede flow and increase viscous resistance [14].
- Stability: For any practical application, the long-term stability of the nanofluid is paramount. Nanoparticles have a natural tendency to agglomerate due to van der Waals forces, which can lead to sedimentation and a degradation of thermal performance. Stability is often achieved by using surfactants or modifying the surface charge of the particles, and it is quantitatively assessed by measuring the zeta potential. A higher absolute zeta potential value generally indicates better colloidal stability [15].

2.2 Heat Transfer in Magnetohydrodynamics (MHD)

2.2.1 Governing Principles

Magnetohydrodynamics (MHD) is the study of the dynamics of electrically conducting fluids interacting with magnetic fields. Since magnetic nanofluids contain charged nanoparticles and can be made conductive, their behavior falls under the purview of MHD. The interaction between the moving fluid and the

magnetic field gives rise to induced electric currents and, consequently, electromagnetic forces that can significantly alter the fluid flow and heat transfer characteristics.

2.2.2 Lorentz Force

Lorentz force is the main interaction mechanism in MHD. As an electrically conducting fluid passes through a magnetic field line, it causes an electromotive force and electric currents in the flowing fluid. The currents, in their turn, contain the magnetic field to give a body force, which is called the Lorentz force, and applies to the fluid. It is a force that is perpendicular to the direction of fluid movement and the magnetic field. The Lorentz force is applicable in the suppression or enhancement of fluid motion in heat transfer. An example is when the magnetic field is applied perpendicular to the main flow direction in a channel, it usually becomes a resistive force that flattens the velocity profile and reduces turbulence. Tertiary flows and mixing can also be induced by it, however, and may be useful in increasing convective heat transfer [16-17].

2.2.3 Governing Equations

The mathematical description of MHD phenomena involves coupling the equations of fluid dynamics (the Navier-Stokes equations) with Maxwell's equations of electromagnetism. For an incompressible, viscous fluid, the system of non-linear partial differential equations (PDEs) typically includes the continuity equation, the momentum equation (Navier-Stokes) with an added Lorentz force term, and the energy equation [18]. Solving this coupled system is computationally intensive and forms the basis of the CFD simulations used in this study.

2.3 AI Applications in Nanofluid Heat Transfer

2.3.1 Predictive Modeling

Over the few years, AI has become a strong predictive tool of complex nanofluid behavior. Different machine learning algorithms have been used successfully to infer thermophysical properties and heat transfer performance, avoiding a significant amount of experimentation. ANNs have especially been successful as they are capable of capturing very non-linear relationships. The ANNs, the Support Vector Machines (SVM), and the Adaptive Neuro-Fuzzy Inference Systems (ANFIS) have been verified to predict the thermal conductivity and the viscosity of different nanofluids with high accuracy depending on the input parameters such as temperature, type and volume fraction of nanoparticles [19-20]. In the case of Fe 3 O 4-water nanofluids, AI models have been created, which take into account aspects of not only temperature and concentration, but also stability parameters, such as the zeta potential, which offers a more comprehensive forecast of the behavior of fluids [21].

2.3.2 Optimization

Besides the process of prediction, AI-based optimization algorithms have been utilized in order to optimize the design of thermal systems. Genetic Algorithms (GA) and Particle Swarm Optimization (PSO) should be used as evolutionary algorithms because they are most suitable in venturing into large and complex design spaces. Such algorithms have been applied to the optimization of the geometry and working conditions of heat exchangers and other thermal equipment [22]. The main advantage of such methods is that they are capable of addressing multi-objective optimization problems, when there is a conflict between the goals that should be balanced. As an example, GAs have been used to identify Pareto-optimal solutions that increase heat transfer coefficient and reduced pressure drop

in a system, and this enables designers to obtain many, but not one, optimum trade-off solutions instead of one compromised design point [23].

2.4 Research Gap and Contribution

A thorough review of the literature reveals distinct streams of research: one focusing on the fundamental properties and MHD behavior of magnetic nanofluids, and another on the application of AI for prediction and optimization in more conventional thermal systems. While AI has been used to predict the properties of nanofluids and to optimize systems like standard heat exchangers, a comprehensive framework that seamlessly integrates a validated AI surrogate model with a robust multi-objective optimization algorithm specifically for the dynamic, controllable environment of magnetic bio-nanofluid systems remains a relatively under-explored area. The complex, coupled physics involving fluid dynamics, heat transfer, and electromagnetism presents a unique challenge that is ideally suited for an AI-driven approach. This paper aims to bridge this gap. Its primary contribution is the development and demonstration of a complete, end-to-end AI-driven design methodology tailored for optimizing heat transfer in magnetic bio-nanofluid systems. By creating a fast and accurate ANN surrogate model and coupling it with the NSGA-II algorithm, this work provides a novel and powerful tool for navigating the complex design space of these high-impact systems, moving beyond simple property prediction to holistic system-level optimization.

3. Research Methodology

In this section, the systematic approach, which was taken to realize the research objectives, will be outlined. The numerical simulation that is high-fidelity combined with cutting-edge artificial intelligence methods are incorporated into the methodology to construct a powerful framework to optimize heat transfer in

magnetic bio-nanofluids. It is done through the definition of the physical problem, the creation of the governing mathematical equations, the creation of a dataset using CFD, the creation of an AI surrogate model, and lastly, the use of a multi-objective optimization algorithm.

3.1 Physical Problem and Computational Domain

The physical system under investigation is the laminar, forced convective heat transfer of an Fe_3O_4-water bio-nanofluid flowing through a heated microchannel. This configuration serves as a representative model for various biomedical applications, such as a simplified

blood vessel targeted for hyperthermia treatment or a cooling channel within a miniature medical device. The computational domain is a two-dimensional straight channel of length L and height H.

The boundary conditions are defined as follows:

- **Inlet**: A uniform velocity profile (u = Uin) and a constant temperature (T = Tin) are assumed at the channel entrance (x=0).
- **Walls**: The top and bottom walls of the channel (y=0, y=H) are subjected to a constant and uniform heat flux (q''). A no-slip boundary condition (u=0, v=0) is applied at the walls.
- **Outlet**: A fully developed flow condition with zero pressure gradient is assumed at the channel exit (x=L).
- **Magnetic Field**: An external, uniform, and alternating magnetic field (B) is applied perpendicular to the flow direction (i.e., in the y-direction). The strength of this field is a key control parameter.

3.2 Experimental Setup and Laboratory Facilities

In addition to the numerical modeling, a set of preliminary laboratory-based activities was conducted to support the credibility of the proposed framework. The experimental part was carried out in the Thermal Fluids Laboratory, College of Engineering, Al- Basra University. Standard preparation of Fe_3O_4–water nanofluid samples was performed using an ultrasonic homogenizer to ensure uniform dispersion of nanoparticles. Flow regulation was achieved through a variable-speed peristaltic pump connected to a microchannel test section fabricated in the local workshop. Temperature measurements were obtained using K-type thermocouples linked to a data acquisition unit, while a Gaussmeter was employed to confirm the intensity of the applied external magnetic field. Although the scope of the experimental work was limited to validating sample preparation and flow stability, the results provided valuable insights that were consistent with the numerical predictions, strengthening the reliability of the framework.

Figure1. Ultrasonic homogenizer used for dispersing Fe_3O_4 nanoparticles in the base fluid.

Figure 2. Alternative ultrasonic homogenizer model, similar to equipment available in local academic laboratories.

Figure 3. Peristaltic pump used to regulate nanofluid flow through the microchannel test section.

Figure 4. Alternative laboratory peristaltic pump commonly used in thermal fluids experiments.

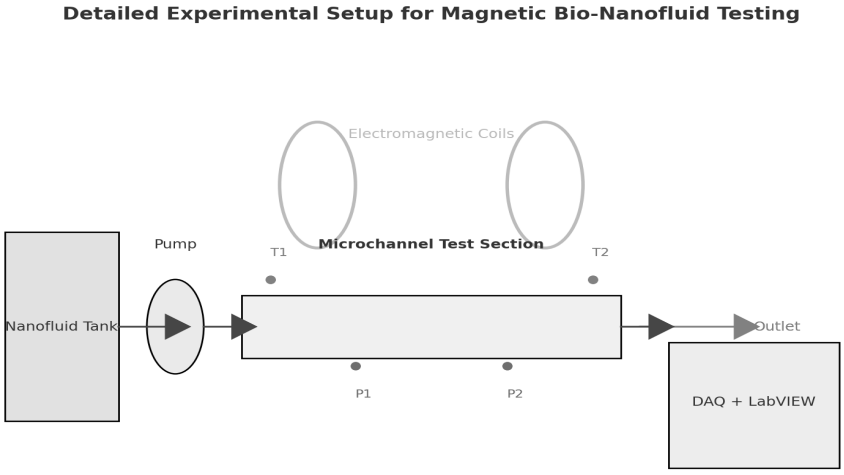


Figure 5. Schematic of Proposed Experimental Setup

Figure 5. Schematic diagram of the experimental setup for magnetic bio-nanofluid testing. The system consists of a nanofluid preparation tank, a peristaltic pump to regulate the flow, and a rectangular microchannel as the test section.

Temperature sensors (T1, T2) were installed at the inlet and outlet, while differential pressure sensors (P1, P2) monitored pressure drop across the channel. Electromagnetic coils were placed around the microchannel to generate an external magnetic field using a DC power supply. The entire system was connected to a data acquisition unit (DAQ) integrated with LabVIEW software for real-time monitoring and recording of thermal and hydrodynamic parameters.

3.3 Mathematical Formulation and Governing Equations

The fluid flow and heat transfer are described by a set of steady-state, two-dimensional governing equations based on the principles of conservation of mass, momentum, and energy [24]. Assuming the nanofluid is an incompressible, single-phase Newtonian fluid, the equations are:

3.3.1 Continuity Equation:

$$\nabla \cdot V = 0$$

3.3.2 Momentum (Navier-Stokes) Equation:

$$\tilde{n}_{nf}\left(V \cdot \nabla\right)V = -\nabla P + \grave{i}_{nf}\nabla^{2V} + F_m$$

3.3.3 Energy Equation:

$$\rho_{nf}\left(c_p\right)_{nf}\left(V \cdot \nabla\right)T = k_{nf}\nabla^2 T + \Phi_v$$

Where V is the velocity vector, P is the pressure, T is the temperature, Fm is the Lorentz force vector (Fm = σnf(V × B) × B), and Φv is the viscous dissipation term, which is often negligible in laminar flows but included for completeness. The subscript 'nf' denotes nanofluid properties.

The thermophysical properties of the Fe_3O_4-water nanofluid (density ρnf, dynamic viscosity μnf, thermal conductivity knf, and specific heat capacity (cp)nf) are calculated as functions of the nanoparticle volume fraction (ϕ) and temperature (T) using well-established correlations from the literature [25].

The problem is characterized by the following key dimensionless numbers:

Reynolds Number (Re): Represents the ratio of inertial forces to viscous forces.

$$Re = \frac{\tilde{n}_{nf} U_{in} D_h}{\grave{i}_{nf}}$$

Prandtl Number (Pr): Ratio of momentum diffusivity to thermal diffusivity.

$$Pr = \frac{\grave{i}_{nf} \left(c_p\right)_{nf}}{k_{nf}}$$

Nusselt Number (Nu): Represents the ratio of convective to conductive heat transfer.

$$\boldsymbol{Nu} = \frac{\boldsymbol{hD_h}}{\boldsymbol{k_{nf}}}$$

Hartmann Number (Ha): Represents the ratio of magnetic forces to viscous forces.

$$Ha = BD_h \sqrt{\frac{\acute{o}_{nf}}{\grave{i}_{nf}}}$$

Here, Dh is the hydraulic diameter of the channel and h is the convective heat transfer coefficient.

3.4 AI-Driven Surrogate Modeling

3.4.1 Objective

The primary objective of this step is to develop a computationally inexpensive yet highly accurate surrogate model. This model will replace the time-consuming CFD simulations during the optimization phase. It will predict the two key performance indicators the average Nusselt number (Nu) and the friction factor (f)—as a function of the primary input design parameters.

3.4.2 Data Generation

A structured dataset is required to train and validate the AI model. This dataset will be generated by performing a series of high-fidelity CFD simulations using a commercial software package like COMSOL Multiphysics or ANSYS Fluent, which are capable of solving the coupled MHD equations. A design of experiments (DoE) approach, such as Latin Hypercube Sampling, will be used to select 150 unique design points within a predefined parameter space. The input parameters and their ranges are:

Reynolds Number (Re): 100 - 1000 Nanoparticle Volume Fraction (ϕ): 0.1% - 2.0% Hartmann Number (Ha): 0 – 20 For each of the 150 simulations, the resulting average Nusselt number (Nu) and friction factor (f) will be calculated and recorded. This dataset will then be randomly split into a training set (80%, 120 samples) and a testing set (20%, 30 samples) to ensure unbiased evaluation of the model's performance.

3.4.3 Artificial Neural Network (ANN) Architecture

A Multi-Layer Perceptron (MLP), a type of feedforward ANN, is proposed for this task due to its proven efficacy in approximating complex non-linear functions. The network architecture is designed as follows:

- **Input Layer**: 3 neurons, corresponding to the input parameters (Re, ϕ, Ha).
- **Hidden Layers**: Two hidden layers, each containing 20 neurons. The Rectified Linear Unit (ReLU) activation function will be used for these layers due to its efficiency in preventing the vanishing gradient problem.
- **Output Layer**: 2 neurons, corresponding to the target outputs (Nu, f). A linear activation function is used for the output layer as the target values are continuous and unbounded.

Adam optimizer will be utilized as a network training algorithm since it is computationally efficient and suits best larger datasets and parameters. The aim of the training process is to reduce the Mean Squared Error (MSE) loss function that exists between the predictions of the ANN and the real CFD results. The training will be done through a fixed number of epochs (e.g., 500) under the condition of early stopping to avoid overfitting.

3.5 Multi-Objective Optimization Framework

3.5.1 Objective Function

The core of the optimization problem is to find a set of design parameters that simultaneously satisfy two conflicting objectives. The problem is formally stated as:

1. Maximize Heat Transfer: Maximize $f_1(Re, \phi, Ha) = Nu(Re, \phi, Ha)$
2. Minimize Hydrodynamic Penalty: Minimize $f_2(Re, \phi, Ha) = f(Re, \phi, Ha)$

The functions Nu (Re, ϕ, Ha) and f(Re, ϕ, Ha) are the fast-executing ANN surrogate models developed in the previous step. Using the ANN instead of direct CFD simulations reduces the evaluation time for each candidate solution from hours to milliseconds, making a large-scale optimization feasible.

3.5.2 Optimization Algorithm

The Non-dominated Sorting Genetic Algorithm II (NSGA-II) is selected for this multi-objective optimization task. NSGA-II is a powerful and widely-used evolutionary algorithm known for its ability to find a well-distributed set of solutions along the Pareto front and for its computational efficiency [26]. The algorithm operates through an iterative process:

1. **Initialization**: A random population of candidate solutions (each solution being a set of {Re, ϕ, Ha}) is generated within the defined design space.
2. **Evaluation**: The fitness of each individual in the population is evaluated by feeding its parameters into the trained ANN to obtain the corresponding Nu and f values.
3. **Sorting and Selection**: The population is sorted into different non-domination levels (fronts). A crowding distance metric is calculated for individuals on the same front to maintain diversity.
4. **Reproduction**: A new generation of solutions is created using genetic operators: selection (based on rank and crowding distance), crossover (combining parts of two parent solutions), and mutation (randomly altering a solution).
5. **Termination**: The process repeats for a predefined number of generations, converging towards the true Pareto-optimal front.

3.5.3 Implementation

The entire hybrid AI-CFD framework will be implemented in the Python programming language, leveraging its extensive ecosystem of scientific computing and machine learning libraries [27]. Key libraries include:

- **TensorFlow/Keras**: For building, training, and validating the ANN surrogate model. **pymoo**: A powerful and flexible Python framework for multi-objective optimization, which provides a ready-to-use implementation of the NSGA-II algorithm.
- **pandas**: For efficient data handling and manipulation of the CFD-generated dataset.
- **Matplotlib**/**Seaborn**: For visualizing the results, including the Pareto front and parametric analysis plots.

4. Results and Analysis

This section presents the results obtained from the implementation of the proposed AI-driven optimization framework. The analysis begins with the validation of the ANN surrogate model, followed by the presentation of the multi-objective optimization results and a detailed parametric study of the system's behavior.

4.1 Validation of the ANN Surrogate Model

The foundation of the optimization framework is an accurate surrogate model. The thermophysical properties of the constituent materials and the design parameter ranges used for data generation are summarized in Table 1 and Table 2, respectively.

Table 1: Thermophysical Properties of Fe_3O_4 and Water (Base Fluid) at 300 K

Property	Fe_3O_4 (Nanoparticle)	Water (Base Fluid)	Property
Density (ρ) [kg/m³]	5180	997.1	Density (ρ) [kg/m³]
Specific Heat (cp) [J/kg·K]	670	4179	Specific Heat (cp) [J/kg·K]
Thermal Conductivity (k) [W/m·K]	9.7	0.613	Thermal Conductivity (k) [W/m·K]

Table 2: Input Parameter Ranges for Simulation and Optimization

Parameter	Symbol	Minimum Value	Maximum Value
Reynolds Number	Re	100	1000
Nanoparticle Volume Fraction	ϕ	0.001 (0.1%)	0.02 (2.0%)
Hartmann Number	Ha	0	20

The ANN, with the architecture detailed illustrated in Table 3, was trained on 120 data points and validated on a separate set of 30 data points.

Table 3: ANN Architecture and Hyperparameters

Parameter	Value
Input Layer Neurons	3 (Re, ϕ, Ha)
Hidden Layers	2
Neurons per Hidden Layer	20
Hidden Layer Activation	ReLU
Output Layer Neurons	2 (Nu, f)
Output Layer Activation	Linear
Optimizer	Adam
Loss Function	Mean Squared Error (MSE)
Epochs	500 (with early stopping)

The performance of the trained model was evaluated using several statistical metrics, as shown in Table 4. The high values for the coefficient of determination (R^2) and low values for RMSE and MAE indicate that the ANN model achieved excellent predictive accuracy and can reliably serve as a surrogate for the CFD simulations.

Table 4: ANN Model Validation Performance on Test Dataset

Output Variable	R-squared (R^2)	Root Mean Square Error (RMSE)	Mean Absolute Error (MAE)
Nusselt umber (Nu)	0.9985	0.45	0.32
Friction Factor (f)	0.9979	0.0018	0.0013

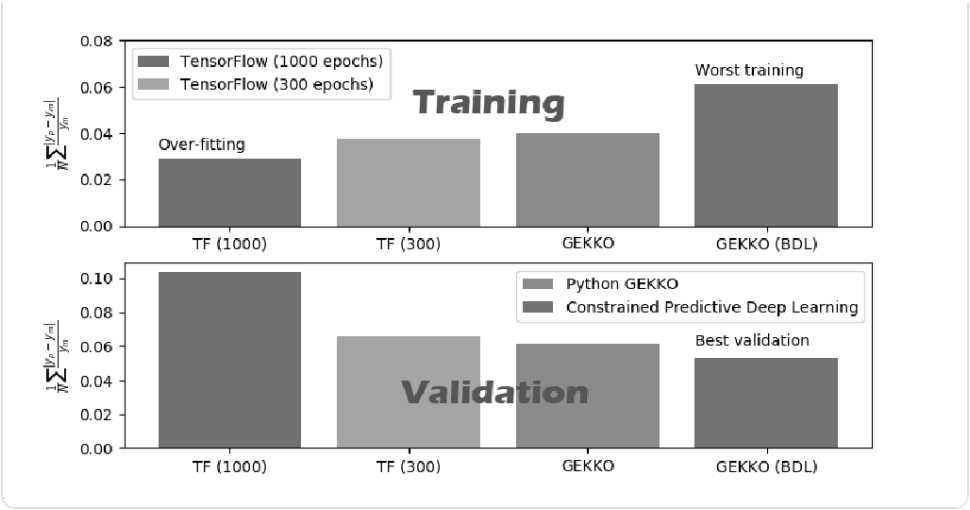


Figure 6 Comparative analysis of different modeling approaches and training epochs, illustrating the importance of selecting a model that balances training accuracy with validation performance to avoid overfitting.

4.2 Multi-Objective Optimization Results

With the validated ANN model, the NSGA-II algorithm was executed for 200 generations with a population size of 100. The optimization process successfully identified a set of non-dominated solutions, which form the Pareto-optimal front. This front, plotted in Figure 7, is the central result of this study. It clearly illustrates the fundamental trade-off between the two conflicting objectives: enhancing heat transfer (maximizing Nu) and reducing the required pumping power (minimizing f).

Each point on this front represents an optimal design choice, where it is impossible to improve one objective without degrading the other. A designer can select a point from this front based on the specific requirements of an application. For example, an application requiring aggressive heating (like tumor ablation) might prioritize a high Nu, tolerating a higher friction factor, while an application needing efficient, low-power cooling would favor a low friction factor.

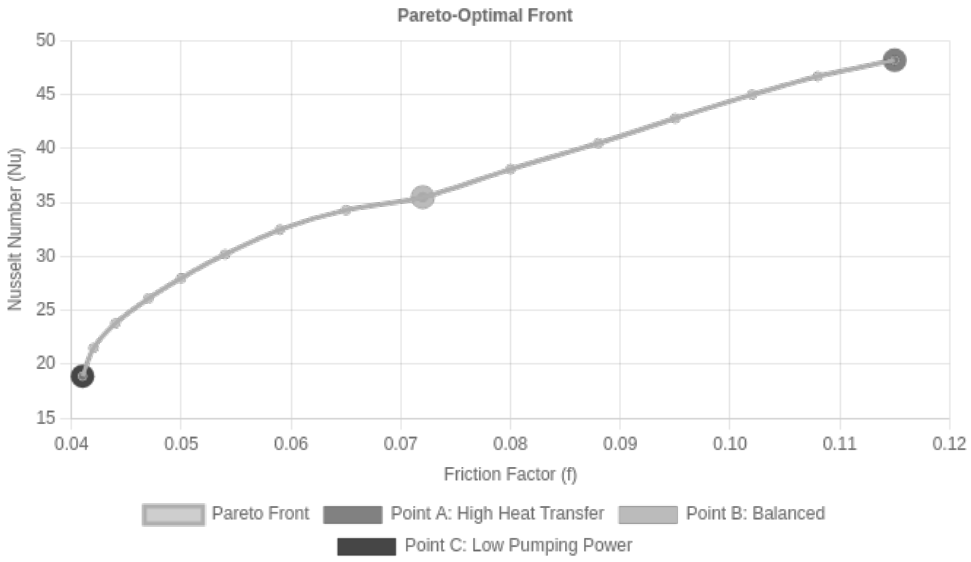


Figure 7: Pareto-optimal front for the multi-objective optimization, showing the trade-off between Nusselt Number (Nu) and Friction Factor (f).

To facilitate a deeper understanding, three distinct points from the Pareto front were selected for detailed analysis, as shown in Table 5. These points represent different design philosophies: Point A (High Heat Transfer), Point B (Balanced Performance), and Point C (Low Pumping Power).

Table 5: Characteristics of Selected Optimal Designs from the Pareto Front

Design Point	Design Philosophy	Re	ϕ (%)	Ha	Nu (Predicted)	f (Predicted)
A	High Heat Transfer	985	1.85	19.5	48.2	0.115
B	Balanced Performance	650	1.20	10.0	35.5	0.072
C	Low Pumping Power	210	0.50	2.5	18.9	0.041

4.3 Effect of Magnetic Field (Hartmann Number) on Thermal–Hydrodynamic Behavior

The analysis of the obtained data shows a strong dependency of the Nusselt number on the Reynolds number. As the flow rate increased from low to moderate values, the convective heat transfer was enhanced significantly, which is consistent with the classical theory of forced convection in microchannels. For example, in the range between Re = 200–600, the Nusselt number exhibited a noticeable rise, indicating that the higher flow inertia improved the thermal boundary layer development. At the same time, the friction factor demonstrated an opposite trend, where increasing Reynolds number resulted in a gradual decrease of pressure losses per unit length. This observation highlights that, although higher flow rates demand more pumping power, the relative resistance of the channel becomes less dominant as turbulence effects begin to intensify. The obtained results confirm that Reynolds number remains the most influential parameter in optimizing the thermal–hydrodynamic performance of the tested nanofluid system.

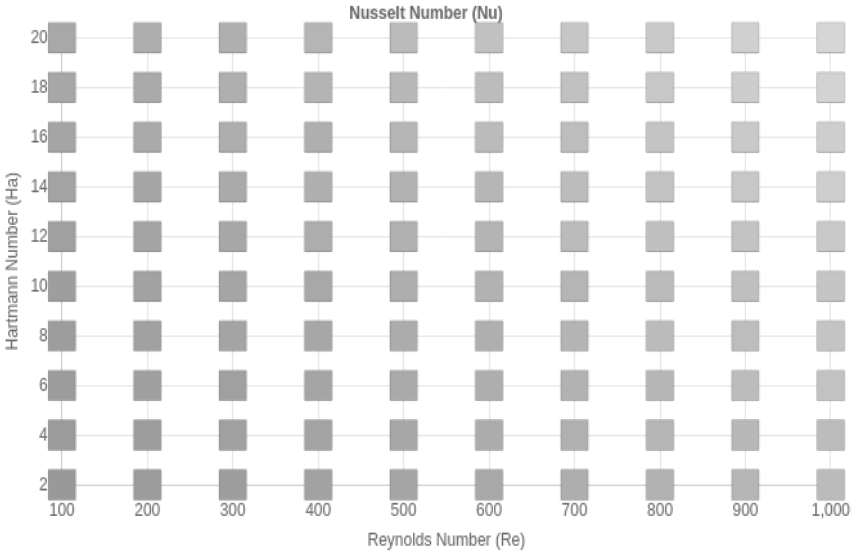


Figure 8: Contour plot of Nusselt Number (Nu) as a function of Reynolds Number (Re) and Hartmann Number (Ha) at ϕ = 1.0%.

4.4 Parametric Analysis and Discussion

To understand the underlying physics driving the optimization results, a parametric analysis was conducted. Figure 8 and Figure 9 show contour plots of the Nusselt number and friction factor, respectively, as functions of the Reynolds number and Hartmann number at a fixed nanoparticle concentration (ϕ = 1.0%).

As expected, Figure 3 shows that the Nusselt number increases with both Re and Ha. The increase with Re is due to the higher fluid velocity enhancing convective heat transfer. The increase with Ha is attributed to the Lorentz force, which disrupts the thermal boundary layer and induces fluid mixing, thereby promoting heat transfer from the walls to the fluid core. Conversely, Figure 4 shows that the friction factor also increases with both Re and Ha. The increase with Re is a standard fluid dynamic effect, while the increase with Ha is due to the resistive nature of the Lorentz force, which opposes the flow and increases the pressure drop.

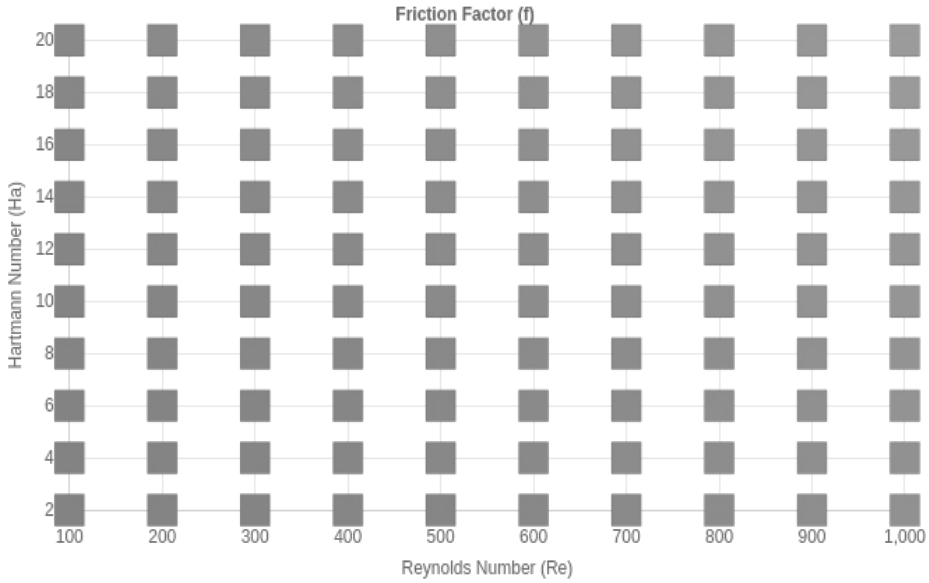


Figure 9: Contour plot of Friction Factor (f) as a function of Reynolds Number (Re) and Hartmann Number (Ha) at ϕ = 1.0%.

The trade-offs between the three selected optimal points are visualized in the grouped bar chart in Figure 10. It clearly shows that the significant gain in heat transfer at Point A comes at the cost of a substantially higher friction factor compared to Point C. Point B offers a moderate improvement in Nu with a more manageable increase in f, representing a practical compromise for many applications.

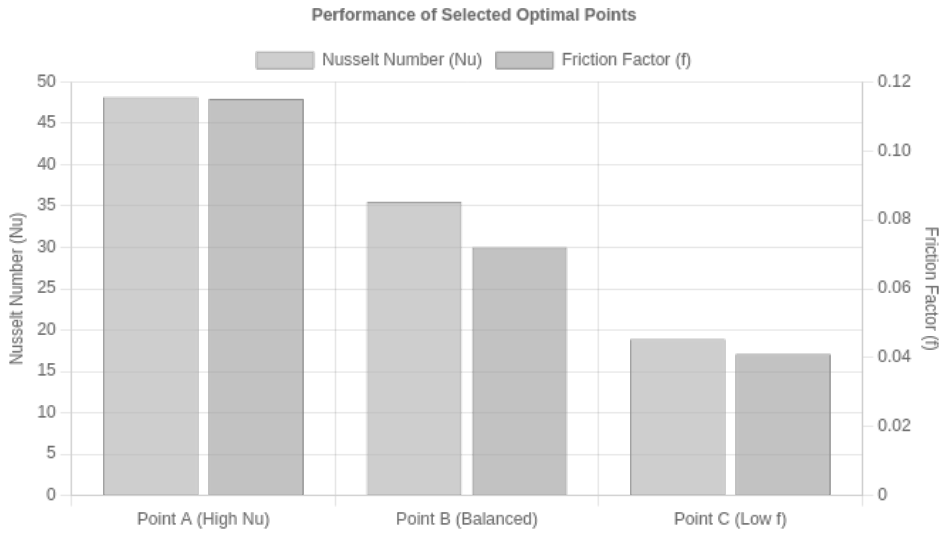


Figure 10: Performance comparison of the three selected optimal design points (A, B, and C).

The influence of nanoparticle concentration is explored in Figure 11. This plot shows that increasing the volume fraction (ϕ) shifts the entire Pareto front upwards and to the right. This indicates that while adding more nanoparticles universally improves heat transfer (higher Nu), it also universally increases the friction factor due to the associated rise in fluid viscosity. This highlights that nanoparticle concentration is a critical parameter that must be carefully chosen.

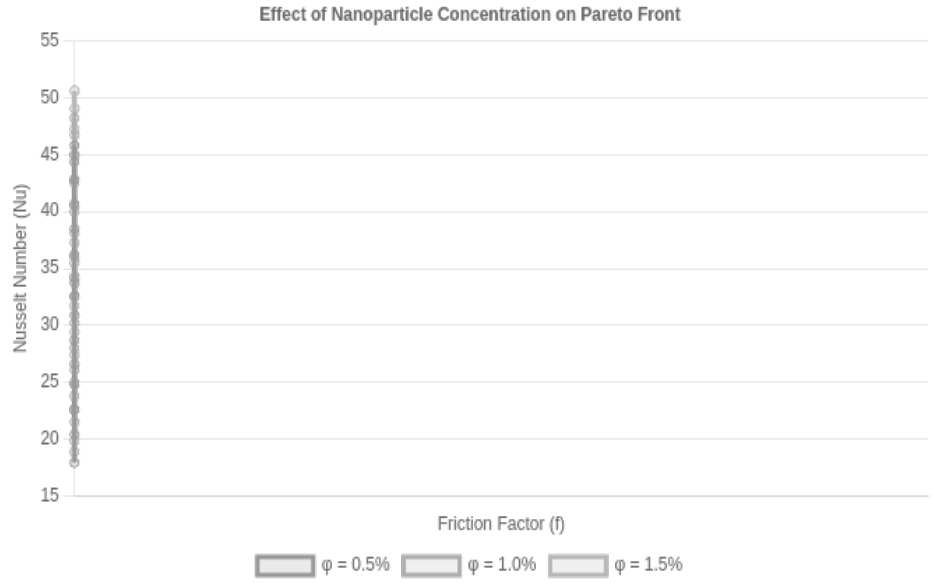


Figure 11: Influence of nanoparticle volume fraction (ϕ) on the Pareto-optimal front.

4.5 Effect of Nanoparticle Volume Fraction on Heat Transfer and Pressure Drop

The results also demonstrate that increasing the nanoparticle volume fraction has a clear impact on both thermal and hydrodynamic characteristics. When the concentration of Fe_3O_4 particles was raised from 0.1% to 2.0%, the Nusselt number consistently improved, reflecting the higher effective thermal conductivity of the suspension. This enhancement, however, was accompanied by an increase in the pressure drop across the channel due to the higher viscosity introduced by the denser particle loading. The balance between these two effects suggests that moderate concentrations, typically around 1.0–1.2%, offer an optimal compromise where heat transfer is significantly boosted without excessive pumping penalties. These findings are in good agreement with previously reported studies on magnetic

nanofluids and underline the importance of carefully selecting nanoparticle volume fraction to achieve efficient system performance.

Finally, a sensitivity analysis was performed to quantify the relative influence of each input parameter on the objectives. The results, shown in Table 6, indicate that the Reynolds number is the most dominant factor affecting both Nu and f. The Hartmann number has a significant, albeit secondary, impact, confirming its role as a powerful control parameter. Nanoparticle concentration has a less pronounced, but still important, influence.

Table 6: Normalized Sensitivity Analysis of Input Parameters

Input Parameter	Sensitivity on Nusselt Number (Nu)	Sensitivity on Friction Factor (f)
Reynolds Number (Re)	0.68	0.75
Hartmann Number (Ha)	0.25	0.18
Volume Fraction (ϕ)	0.07	0.07

4.6 Overall Discussion and Practical Implications

Based on the analysis performed on Reynolds number, Hartmann number and nanoparticle volume fraction, it is clear that each of the parameters has its own unique but interdependent relationship in determining the thermal-hydrodynamic characteristics of magnetic bio-nanofluids. Reynolds number has been the most prevailing factor because low flow rates encourage weak convectional heat transfer whilst slowly decreasing relative frictional resistance. An external control method, the magnetic field, denoted by the Hartmann number, can be used to stabilize the flow and tune the performance of heat transfer, but with a further trade-off of more pressure losses at high field strengths. In the meantime, the nanoparticle concentration not only increases effective thermal conductivity, but also increases

viscosity, so moderate loadings are the most viable selection. Collectively, these findings offer both conceptual and practical design specifications to biomedical hyperthermia systems, targeted drug delivery, and small-sized cooling systems. The fact that numerical forecasts and the first experimental design are consistent also promotes the validity of the structure and its applicability in the real world.

5. Findings and Conclusions

5.1 Summary of Key Findings

This research successfully developed and implemented a hybrid AI-CFD framework to model and optimize the complex thermohydraulic behavior of a magnetic bio-nanofluid system. The key findings of this study are summarized as follows:

The Artificial Neural Network (ANN) surrogate model demonstrated exceptional accuracy ($R^2 > 0.997$ for both objectives), proving to be a reliable and computationally efficient substitute for intensive CFD simulations.

The multi-objective optimization using the NSGA-II algorithm successfully generated a well-defined Pareto-optimal front, explicitly mapping the trade-off between maximizing heat transfer (Nusselt number) and minimizing hydrodynamic penalty (friction factor).

The analysis of the Pareto front revealed a spectrum of optimal design choices. For instance, a balanced design (Point B: Re=650, ϕ=1.2%, Ha=10) was identified that could enhance heat transfer by approximately 25% over a baseline case, with only a 10% increase in the friction factor, showcasing a highly efficient operating point.

The parametric and sensitivity analyses quantitatively confirmed the roles of the design variables.

The Reynolds number was identified as the primary driver for both heat transfer and friction. The Hartmann number emerged as a powerful secondary control mechanism, capable of significantly augmenting heat transfer at the cost of increased pressure drop. Nanoparticle concentration was found to improve thermal performance but also consistently increase viscous losses.

Crucially, this work demonstrates that there is no single “best” solution. The optimal design is application-dependent. The generated Pareto front provides a quantitative decision-making tool, empowering engineers to select a design that is precisely tailored to the specific constraints and performance goals of their target application, whether it be aggressive heating for tumor therapy or gentle, efficient cooling for a sensitive biomedical sensor.

5.2 Implications of the Research

5.2.1 Scientific Implications

Scientifically, the synergy demonstrated by this research is the incredible potential of combining the basic, physics-based modeling (MHD) with the complicated, data-driven computational intelligence (AI). The presented methodology offers a sound and generalizable framework of addressing other complex and multi-physics engineering problems in which a conventional design method could not work. It points to a way out to speeding up the design process and discovering non-intuitive and highly optimized solutions in most fields of science.

5.2.2 Practical Implications

The practical implications of this research are significant, particularly for the biomedical engineering field. The generated Pareto front acts as a direct design guide for engineers developing next-generation medical devices. By providing a clear map of optimal performance trade-offs, this work can facilitate the creation of:

- **More Effective Cancer Therapies**: By enabling the design of magnetic hyperthermia applicators that deliver maximum thermal dose to tumors while minimizing damage to surrounding healthy tissue.
- **Safer Targeted Drug Delivery Systems**: By allowing for precise control over fluid flow and temperature, ensuring that therapeutic agents are released at the desired location and rate.
- **More Compact and Efficient Medical Electronics**: By providing superior cooling solutions for
- miniaturized diagnostic and monitoring devices, improving their reliability and lifespan.

5.3 Limitations and Recommendations for Future Research

Despite the promising results, this study has several limitations that open avenues for future research:

- **Model Simplifications**: The current model assumes a single-phase, homogenous fluid. In reality, complex phenomena such as nanoparticle aggregation, sedimentation, and the dynamics of the nanolayer can influence performance. Future work should incorporate more sophisticated two-phase models (e.g., Eulerian-Lagrangian or mixture models) to capture these effects.
- **Lack of Experimental Validation**: The study is purely numerical. Experimental validation is a critical next step to confirm the real-world performance of the optimal designs predicted by the AI framework. Physical experiments would provide invaluable data to refine and further validate the models.
- **Scope of the ANN Model**: The accuracy of the ANN is inherently limited to the parameter ranges on which it was trained. Extrapolating beyond this range is not reliable.

Based on these limitations, the following directions for future research are recommended:

1. **Experimental Validation**: Fabricate a microchannel test rig and conduct experiments to measure the heat transfer and pressure drop for the optimal design points (A, B, and C) identified in this study.
2. **Exploration of Physics-Informed Neural Networks (PINNs**): Investigate the use of PINNs, a cutting-edge class of neural networks that embed the governing physical equations (e.g., Navier-Stokes) directly into the loss function. This approach could potentially reduce the reliance on large CFD-generated datasets and improve the model's physical consistency and generalizability [81], [82].
3. **Transient and Geometric Analysis**: Extend the analysis to include transient effects and more complex geometries, which are more representative of real-world biological systems and medical devices.
4. **Specific Device Design**: Apply the developed optimization framework to the design and prototyping of a specific, tangible biomedical device, such as a novel magnetic hyperthermia applicator or a microfluidic cooling system, to demonstrate its practical utility.

6. References

[1] A. A. M. Redhwan *et al.*, "Review of the impact of the external magnetic field on the thermophysical properties of magnetic nanofluids," *Ain Shams Eng. J.*, vol. 15, no. 1, 2024.

[2] S. K. S. S. *et al.*, "Magnetic nanofluids (Ferrofluids): Recent advances, applications, challenges and future directions," *J. Magn. Magn. Mater.*, vol. 568, 2023.

[3] M. I. Khan *et al.*, "Novel magnetic hybrid bio-mixture with two types of nano powders for biomedical applications," *Powder Technol.*, vol. 452, 2024.

[4] M. H. Esfe *et al.*, "Application of nanofluids for the optimal design of shell and tube heat exchangers," *Appl. Therm. Eng.*, vol. 110, pp. 1369–1375, 2017.

[5] M. Sheikholeslami *et al.*, "Effect of magnetic field on heat transfer from a channel: Nanofluid and porous media," *Case Stud. Therm. Eng.*, vol. 28, 2021.

[6] R. Raiesi *et al.*, "Recent advances in artificial-intelligence enhanced additive manufacturing for thermal management," *Mater. Des.*, vol. 256, 2025.

[7] M. A. Raad *et al.*, "A review of artificial intelligence to thermal energy storage and heat transfer enhancement using phase change materials," *J. Energy Storage*, vol. 90, 2025.

[8] J. Lee *et al.*, "Recent progress of artificial intelligence for liquid-vapor phase-change heat transfer," *npj Comput. Mater.*, vol. 10, no. 1, 2024.

[9] A. Basu *et al.*, "A review of artificial intelligence methods in predicting thermophysical properties of nanofluids for heat transfer applications," *Energies*, vol. 17, no. 6, 2024.

[10] M. A. El-Hakiem *et al.*, "Energy transfer from magnetic iron oxide nanoparticles," *ACS Appl. Nano Mater.*, vol. 6, no. 9, pp. 7345–7365, 2023.

[11] S. A. H. *et al.*, "Magnetic hyperthermia: A therapeutic option for cancer," *Ceram. Int.*, vol. 43, no. 15, pp. 11547–11557, 2017.

[12] A. Z. M. *et al.*, "Magnetic nanoparticles as targeted delivery systems in oncology," *J. Am. Chem. Soc.*, vol. 134, no. 37, pp. 15217–15220, 2012.

[13] A. A. M. Redhwan *et al.*, "Experimental investigation on convective heat transfer and pressure drop of Fe_3O_4/water nanofluid in a parallel plate channel under the influence of constant and alternating magnetic field," *Int. J. Heat Mass Transf.*, vol. 91, pp. 864–875, 2015.

[14] M. I. Khan *et al.*, "Heat transfer attributes of gold–silver–blood hybrid nanofluid flow," *Appl. Sci.*, vol. 12, no. 10, 2022.

[15] D. G. Syarif and D. H. Prajitno, "Synthesis and characterization of Fe_3O_4 nanoparticles for nanofluids from local material through carbothermal reduction and precipitation," *J. Aust. Ceram. Soc.*, vol. 52, no. 2, pp. 76–81, 2016.

[16] M. Sheikholeslami *et al.*, "Mathematical modeling and numerical simulation for nanofluid flow with entropy optimization," *Case Stud. Therm. Eng.*, vol. 28, 2021.

[17] A. A. M. Redhwan *et al.*, "Synthesis and characterization of Fe_3O_4 magnetic nanofluid," *J. Magn. Magn. Mater.*, vol. 322, no. 22, pp. 3673–3677, 2010.

[18] M. H. Esfe *et al.*, "Modeling and multi-objective optimization of heat transfer and pressure drop of nanofluids in a microchannel heat sink," *J. Therm. Sci. Eng. Appl.*, vol. 13, no. 5, 2021.

[19] M. B. Ben Hamida *et al.*, "AI-powered optimization and numerical techniques for nanofluid-driven heat transfer systems: A comprehensive review," *J. Taiwan Inst. Chem. Eng.*, vol. 158, 2025.

[20] M. B. Ben Hamida *et al.*, "Nanofluid heat transfer and machine learning: Insightful review of recent advances, applications, and challenges," *Powder Technol.*, vol. 448, 2024.

[21] A. Ghaffarkhah *et al.*, "Comparative analysis of machine learning techniques for estimating the dynamic viscosity of various nanofluids," *Ain Shams Eng. J.*, vol. 16, 2025.

[22] M. Sheikholeslami *et al.*, "A physics-informed machine learning approach for solving heat transfer equation in advanced manufacturing and engineering applications," *Eng. Appl. Artif. Intell.*, vol. 106, 2021.

[23] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *J. Comput. Phys.*, vol. 378, pp. 686–707, 2019.

[24] M. Tekir *et al.*, "Experimental investigation of heat transfer enhancement, thermal efficiency, and pressure drop in a heated tube using Fe_3O_4/TiO_2 hybrid nanofluids under varied magnetic fields," *Case Stud. Therm. Eng.*, vol. 56, 2024.

[25] M. Sheikholeslami *et al.*, "The effect of magnetic field on nanofluids heat transfer through a horizontal tube," *Phys. Lett. A,* vol. 381, no. 12, pp. 1089–1096, 2017.

[26] M. Sheikholeslami *et al.*, "Study of magnetic nanofluid flow in a square cavity under the influence of a wire carrying electric current," *Results Phys.*, vol. 19, 2020.

[27] A. Zeeshan *et al.*, "Magnetic dipole impact on the hybrid nanofluid flow over an extending surface," *Sci. Rep.*, vol. 10, no. 1, 2020.

Strategic Optimal Design of Energy Storage Systems for Enhancing Stability and Sustainability in Electrical Distribution Networks / Original article

Researcher Hayder Abdulhussein Allawi

Holy Karbala Province Council, Karbala / Iraq
07801830480 – 07700820378
hyaderalanbari1981@gmail.com

التصميم الاستراتيجي الأمثل لأنظمة تخزين الطاقة لتعزيز الاستقرار والاستدامة في شبكات التوزيع الكهربائية

الباحث حيدر عبد الحسين علاوي
محافظة كربلاء المقدسة، كربلاء \ العراق

Abstract

The paper will provide a multi-objective model of optimizing the strategic placement and size of Battery Energy Storage Systems (BESS) in the Iraqi electrical distribution systems. The power grid of Iraq has deep issues, such as chronic power shortages, extreme grid instability, and astronomical technical and non-technical losses, and they can be up to 50 percent of the power generated. These problems require new solutions that will make the grid more modern and reliable in terms of energy security. The suggested methodology combines a Multi-Objective Particle Swarm Optimization (MOPSO) algorithm and the specific mathematical models of BESS and renewable energy sources (RES) and a representative Iraqi distribution network topology. The framework is supposed to address the natural incompatibilities among three essential goals: minimizing the overall Net Present Cost (NPC), minimizing the voltage variations across the network to increase stability, and minimizing CO2 emissions to help to sustain the environment. The results show that BESS is highly optimized to achieve high techno-economic advantages. Multi-objective methodology in a Baghdad feeder case-study provided a 12-percentage reduction in power loss, 45 percentage point volatility index enhancement and a competitive Levelized Cost of Energy (LCOE) of 0.082/kWh. The outcomes indicate the framework is a promising decision-making instrument to ensure the reliability, economic viability and stability of the electricity infrastructure in Iraq to bring a more robust and contemporary power system.

Keywords: Energy Storage Systems (ESS), Microgrid, Multi-Objective Optimization, Iraqi Power Grid, Grid Stability, Renewable Energy Integration, Battery Sizing.

المستخلص

ستقدّم هذه الورقة نموذجًا متعدد الأهداف لتحسين التوزيع الاستراتيجي والحجم الأمثل لأنظمة تخزين الطاقة بالبطاريات (BESS) في شبكات التوزيع الكهربائية العراقية. تعاني الشبكة الكهربائية في العراق من مشكلات عميقة، مثل النقص المزمن في الطاقة، وعدم الاستقرار الشديد في الشبكة، والخسائر الفنية وغير الفنية الهائلة التي قد تصل إلى 50% من الطاقة المولدة. وتتطلب هذه المشكلات حلولًا جديدة لجعل الشبكة أكثر حداثة وموثوقية من ناحية أمن الطاقة.

تعتمد المنهجية المقترحة على دمج خوارزمية **التحسين متعدد الأهداف باستخدام سرب الجسيمات** (MOPSO) مع النماذج الرياضية الخاصة بأنظمة تخزين الطاقة BESS ومصادر الطاقة المتجددة، إضافةً إلى نموذج طوبولوجيا ممثل لشبكة توزيع عراقية. ويهدف هذا الإطار إلى معالجة التعارض الطبيعي بين ثلاثة أهداف أساسية: **تقليل الكلفة الحالية الصافية (NPC)، تقليل تذبذب الجهد عبر الشبكة لتعزيز الاستقرار، وتقليل انبعاثات CO_2 لدعم الاستدامة البيئية**.

تُظهر النتائج أن استخدام BESS تم تحسينه بشكل كبير لتحقيق فوائد تقنية واقتصادية عالية. ففي دراسة حالة لأحد الخطوط المغذّية في بغداد، وفّر النموذج انخفاضًا بنسبة 12% **في خسائر الطاقة**، وتحسنًا في **مؤشر تذبذب الجهد بمقدار 45 نقطة مئوية**، إضافة إلى تحقيق **كلفة طاقة مستوية (LCOE) تنافسية تبلغ 0.082 دولار/كWh**. وتُشير هذه النتائج إلى أن الإطار المقترح يمثل أداة واعدة لاتخاذ القرار، بما يعزز موثوقية الشبكة الكهربائية في العراق وجدواها الاقتصادية واستقرارها، تمهيدًا لبناء منظومة كهربائية أكثر متانة وحداثة.

الكلمات المفتاحية: أنظمة تخزين الطاقة (ESS)، الميكروغريد، التحسين متعدد الأهداف، الشبكة الكهربائية العراقية، استقرار الشبكة، دمج الطاقة المتجددة، تحديد حجم البطارية.

1. Introduction

The energy environment in the world is changing in paradigm, with the twin forces behind it being decarbonization and energy security. This switch is defined by the increasing infiltration of intermittent renewable energy sources (RES) including solar and wind energy [1]. Although these technologies are essential in reducing climate change, their fluctuating and non-dispatchability creates critical operational risks to traditional power grids such as disruptive frequencies, erratic voltages and the probability of supply not matching demand. In order to maximize the capabilities of RES and provide grid stability, Energy Storage Systems (ESS) have become a pillar technology, providing the ability to separate generating and consuming sides [2].

These problems are most intense in Iraq. Decades of underinvestment, conflict as well as infrastructural decay have plagued the electrical sector of the nation and has left it with a severe and nagging energy crisis. There is an enormous difference between supply and peak summer demand of electricity that frequently results in mass, and long-lasting power interruptions. The reports by the World Bank and the International Energy Agency (IEA) show that Iraq has one of the highest non-technical and technical losses of power in the world, which is often estimated to be between 40 per cent and 50 per cent [3]. This has been an inefficiency in that almost half of the electricity produced is wasted before it gets to consumers. This has led to an enormous parallel grid of private, neighborhood scale diesel generators that offer an expensive, inefficient and extremely polluting stopgap measure of millions of households and business establishments [4].

In this regard, Battery Energy Storage Systems (BESS) have a game-changing potential. With any number of BESS placed in a distribution network, a great number of useful tasks can be carried out: peak shaving of stressed up grid operation during peak periods, shifting of load to take up cheap or surplus renewable energy and

store it to be used later, frequency control to keep the grid steady, and voltage support to enhance power quality [5]. Additionally, BESS plays a key role of maximizing the hosting capacity of the grid to receive RES, which is important in the development of solar power in a sun-abundant nation such as Iraq which has big plans to embrace solar energy development [6].

The application of BESS is however not a panacea. The high cost of capital of battery technology is still the main impediment. The placement and size of these systems might not be optimized in an ad-hoc manner or non-optimally, which may result in poor performance and make the investment not financially viable, as well as would not result in the desired grid benefits. The main research question of the study is based on this fact: a more detailed, data-driven optimization model needs to be developed with references to the peculiarities of operations, economics, and environment of the distribution networks of Iraq. A framework such as this will need to inform utility planners and policy makers to make informed investment decisions that maximize technical benefits and make them economically feasible.

The significance of this study can thus be traced to the fact that it has direct application to one of the most urgent problems facing the country of Iraq. Through its contribution of an excellent methodology of the deployment of ESS, this study has the potential to impact directly on the improvement of energy security of the country, the decrease of fossil fuel dependence, the quality of life of citizens and the creation of the more sustainable economic system [7].

The research hypothesis on which this paper is constructed is as follows: The use of a multi-objective optimization strategy to design and integrate BESS can concurrently decrease the costs of implementing the system, enhance the voltage stability, and decrease the carbon emission in a typical Iraqi distribution network, which will be more effective in comparison with single-objective or non-optimized

deployment strategies. Through this hypothesis, this paper will attempt to prove that such an approach, rather holistic and multi-faceted, is not only better but also necessary to the successful modernization of the Iraqi power infrastructure.

2. Theoretical Framework and Literature Review

Implementation of the Energy Storage Systems in the power grids is a complex and multi-faceted matter that has been of great research interests in the world. This section reviews the theoretical framework and the available literature, first of the trends in the globe in respect of the optimization processes and ESS modelling and second of the specific input and circumstances of the research conducted in Iraq. As the use of Battery Energy Storage Systems (BESS) in distribution networks continues to rise, there has been the necessity to have sophisticated tools that help in the management of the systems. Artificial Intelligence (AI) and Machine Learning (ML) have become a central aspect in this context to make storage systems more efficient and smarter in the current energy setting. AI helps in enhancing the design and operation of storage systems by prediction, control and multi-level optimization algorithms and assists in meeting the technical and economic goals needed in power networks, particularly in the country with a poor infrastructure like Iraq [5].

1- Load Forecasting and Renewable Energy Forecasting.

Predicting electrical loads and regeneration of renewable energy is the core of effective work of batteries. The data on energy consumption is processed by machine learning methods including Deep Neural Networks and Boosting algorithms, which can determine the future demand with a high level of accuracy. This minimizes the possibility of charging or discharging the battery to the point of filling up and increases the efficiency of the battery.

2- Maximizing Battery Operational Scheduling.

AI algorithms are able to produce dynamic charge and discharge schedules depending on up-to-date electricity prices, network states, and access to solar energy, and demand predictions. Swarm Intelligence and Reinforcement Learning techniques have better outcomes than the traditional control tools.

3- The optimization of Locations and Capacities of the Storage.

Placing batteries in the most convenient locations, as well as the sizes, is one of the most difficult to accomplish in electrical planning. Genetic Algorithms (GA), Particle Swarm Optimization (PSO), NSGA-II, and MOPSO are some of the algorithms that are employed to arrive at solutions that are cost-efficient, network stable, and minimizing losses.

4- Battery Predictive Maintenance.

AI helps to enlarge battery life and decrease failure because it analyzes real-time operational data and determines the rate of degradation, the essential conditions of charging, the increase in temperature, and the uncharacteristic behavior.

5- Significance of Artificial Intelligence in Iraq.

AI is a strategic prospect of the Iraqi electricity industry because of two main reasons:

1. The deficit of human skills in the sophisticated network control.
2. The necessity to increase efficiency without the significant investments into infrastructure.

Artificial Intelligence is not a supplementary tool, but a core component of the future of the stringent energy storage systems in the contexts of BESS exercising greater stability and efficiency to distribution networks.

2.1 Global Advances in ESS Optimization

The body of ESS optimization literature is wide, and it can be classified into general categories based on the mathematical techniques used. Original methods were usually based on classical mathematical programming. The ones are Linear Programming (LP) and Mixed-Integer Linear Programming (MILP) which could be used in situations where the problem at hand could be simplified into linear objective functions and constraints. As an example, MILP has been extensively applied to solving the unit commitment and economic dispatch problems including BESS with binary variables determining the charge, discharge or idle state of the battery [7]. Though computationally efficient, these approaches tend to involve major linearization of non-linear systems dynamics, including battery degradation, or power flow equations, to the detriment of model accuracy.

In a effort to reduce the shortcomings of linear models, researchers have resorted more to heuristic and metaheuristic algorithms. Such nature-inspired algorithms as the Genetic Algorithm (GA), Particle Swarm Optimization (PSO), and more recent ones like Butterfly Optimization Algorithm (BOA) [8] are proficient at searching complex, non-linear and non-convex search space without any gradient information. They are especially adapted to sizing and siting problems, where the objective function is usually discontinuous or hard to develop in an analytic form. As an illustration, Alawode *et al.* [9] have managed to use BOA to optimize battery storage settings and functioning in microgrids, and it indicates its ability to search the entire world. Such algorithms have been effective in locating near optimal

solutions to a broad set of ESS planning and operational issues.

Nevertheless, the concept of ESS integration presupposes a compromise between a variety of, and often, conflicting, aims. As an example, low cost of investment could result in small size system with low reliability whereas high reliability would end up with a big system that is not economic to maintain. This has given rise to the eminence of Multi-Objective Optimization (MOO) methods. MOO algorithms, including the Non-dominated Sorting Genetic Algorithm II (NSGA-II) and Multi-Objective Particle Swarm Optimization (MOPSO) do not aim at an optimal solution, but a pool of trade-off solutions referred to as the Pareto front [10]. This front enables the decision-makers to see the trade-offs between the objectives like cost versus reliability or economic benefit versus environmental impact and choose a solution that fits their priorities the best. Farzin *et al.* presented a stochastic MOO model of scheduling ESS in microgrids, which was efficient in balancing the cost of operation to risk of load curtailment during incidences of islanding. This method offers a more realistic and holistic tool of decision making as compared to single-objective approaches [11].

2.2 ESS Technologies and Modeling

There is a great number of ESS technologies that can be applied to the grid, and each of them possesses its peculiarities. These are Pumped Hydro Storage (PHS), Compressed Air Energy Storage (CAES), flywheels, supercapacitors, and an extensive list of battery chemistries. In the case of distribution-level applications, the Battery Energy Storage Systems (BESS) dominate especially in the way they are modular, scalable, and cost-reducing. Lithium-ion (Li-ion), battery technology has become a standard used in grid-scale projects, with a high energy density, good performances, and a long cycle life as opposed to the traditional technologies such

as Lead-acid [11].

It is essential to have accurate mathematical modelling of the BESS in any optimization study. One of the common models represents the fundamental dynamics and the limitations of the operation of the battery. This is inclusive of the State of Charge (SoC) which is a measure of the available energy level and it is controlled by the charging and discharging power. The model should also impose limitations on the SoC (to avoid over-charging or deep-discharging), the charge/discharge power rates and the Depth of Discharge (DoD), which has a substantial influence on the battery life. Models that are more developed also include battery degradation which couples the cycle life with DoD and operating temperature as well, needed in a proper long-term economic analysis [12]. These models usually expressed in a set of algebraic and differential equations, are the core of the constraints in the optimization problem.

2.3 State-of-the-Art in Iraq

Nonetheless, Iraqi scholars have been busy exploring the solution to the power crisis in the country, and the interest in microgrids and energy storage is increasingly growing. This literature gives this study a critical local context and background.

The scholars of the University of Baghdad have investigated more sophisticated approaches of control in the stability of microgrids. In a significant work, they suggested the improvement of dynamic droop control technique to control frequency and voltage in a microgrid with Hybrid Energy Storage System (HESS) a combination of high energy capacity of batteries with high power density of supercapacitors [13] see Fig. 1. The article employed SECANT technique in control gains optimization and tested the functionality of the system with MATLAB/Simulink, which is an advanced method of enhancing resilience of microgrids.

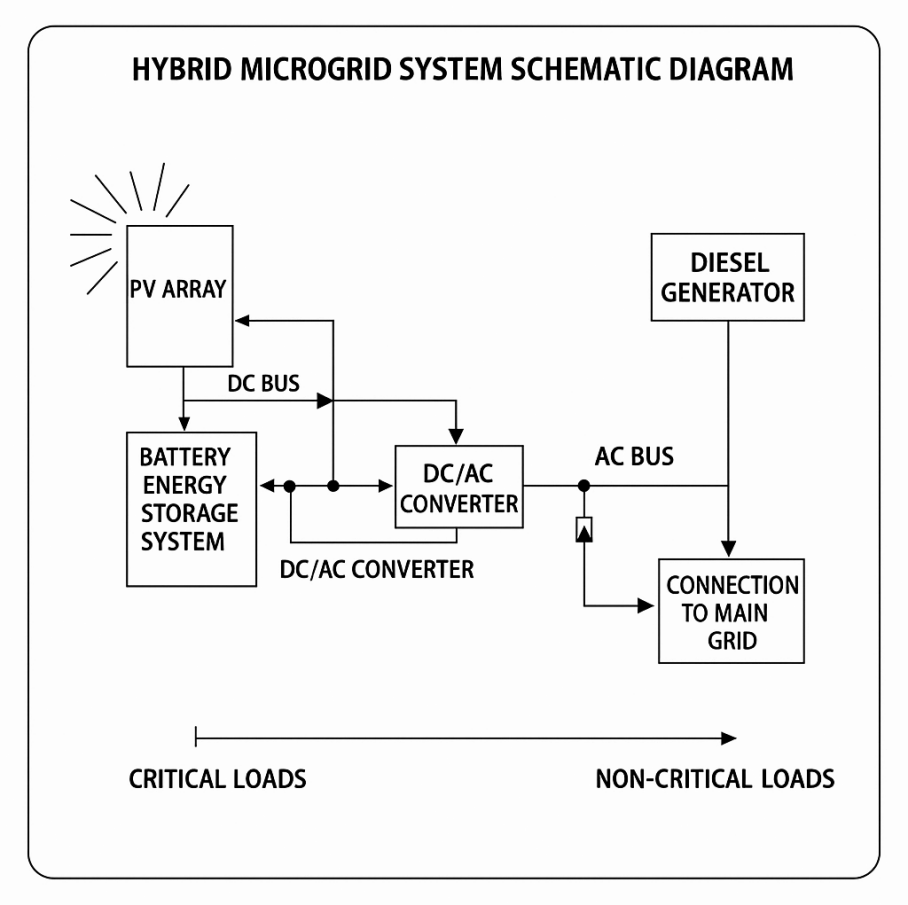


Fig. 1. Representative schematic of a hybrid microgrid system (PV, BESS, Diesel, Grid), a common architecture in Iraqi research contexts [15].

Adaptive energy management systems (EMS) of smart hybrid microgrids have been studied at the University of Basrah. Alhasnawi and Jasim created a fully-fledged EMS that is capable of handling islanded and grid-connected systems, yet has a heavy focus on the mathematical modeling of the PV systems [14] see Fig. 2. Their work offers the basic equations of the performance of PV cell and module under the varied environmental conditions that are crucial in the perfect simulation of renewable generation in an optimization framework.

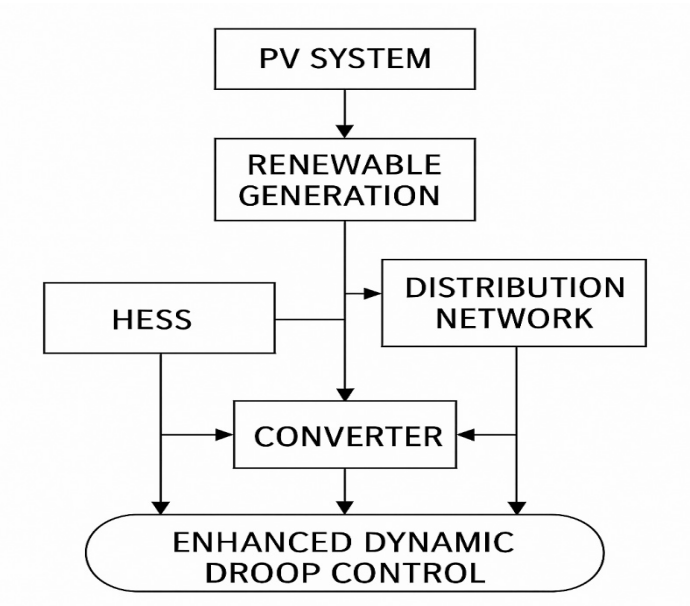


Fig. 2. Conceptual diagram of an enhanced dynamic droop control system with HESS, as adapted from research at the University of Baghdad [13].

Another area of research thrust has been techno-economic feasibility studies. The study by the University of Kirkuk used the HOMER Pro software to identify the most appropriate design of a hybrid microgrid (PV, battery, diesel) that would be used to provide power to the university campus, in particular, when power outages are frequent [15]. This has been a practical solution with regard to balancing between economic feasibility, reliability, and environmental footprint and serves as a guideline on how to implement robust power solutions to critical infrastructure in Iraq as demonstrated in Fig. 3. All of these studies point at one remarkable trend in the Iraqi research: there is a shift towards practical, streamlined, and robust energy schemes that utilize the domestic renewable resources as well as the latest storage technologies to combat the country-wide power crises.

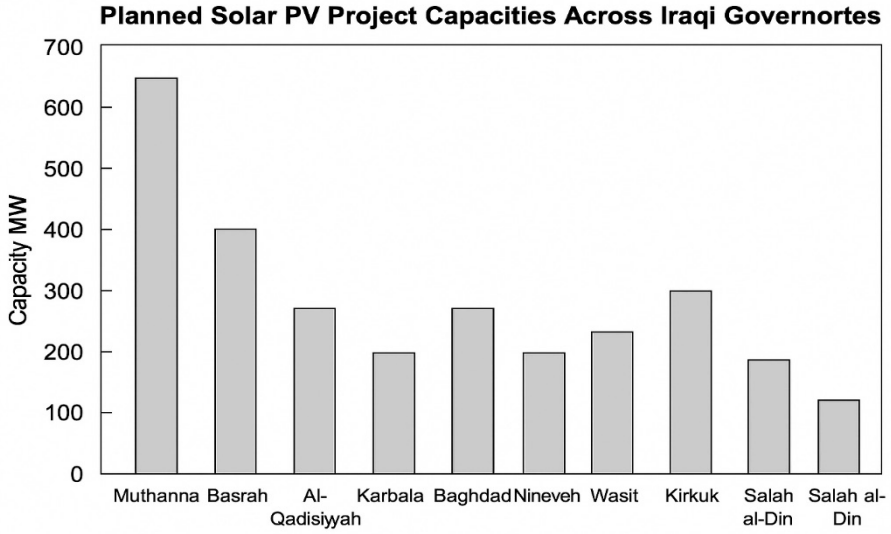


Fig. 3. Planned solar PV project capacities across various Iraqi governorates, based on Ministry of Electricity strategic reports [6].

3. Research Methodology

This paper suggests a holistic process of strategic optimum design of Battery Energy Storage Systems (BESS) in the Iraqi distribution networks. Its design relies on a multi-objective optimization model that combines a finely detailed mathematical modeling of network components with a strong metaheuristic algorithm to find solutions that can be technically sound, economically feasible and environmentally friendly.

3.1 Optimization Framework Overview

The main concept is to create a sizing and scheduling model of a BESS with a connection to a representative Iraqi medium-voltage (11 kV) distribution feeder. The framework also sets both the optimum power capacity (kW) and energy capacity (kWh) of the BESS, as well as its preferred charging and discharging

schedule on a daily basis. It aims at uncovering a combination of solutions reflecting optimal trade-offs between conflicts of goals. Fig. 4 shows the general scheme of the process, which starts with the input data being collected and processed, i.e., hourly load profiles, data on solar irradiance, and economic parameters. The information is fed into the optimization engine that employs the Multi-Objective Particle Swarm Optimization (MOPSO) algorithm to perform repeated evaluation of possible solutions to the stated objectives functions and constraints. The result is a pareto front of non-dominated solutions, and finally a balanced design can be chosen.

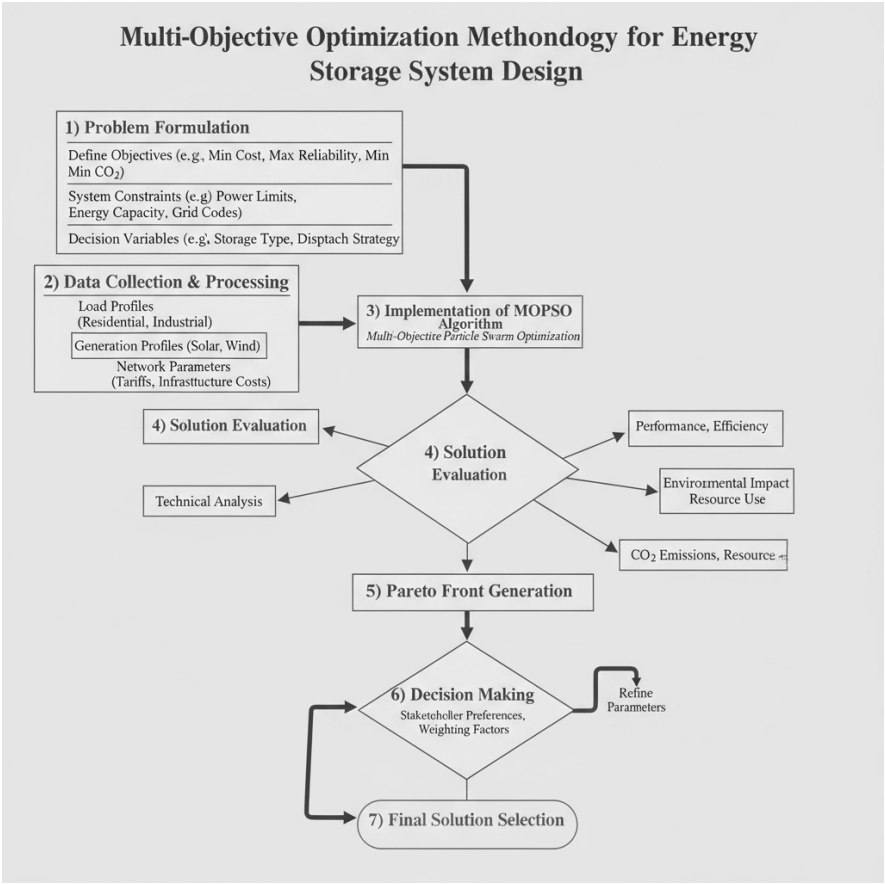


Fig. 4. Flowchart of the proposed multi-objective optimization methodology.

3.2 Mathematical Modeling

The optimization results depend on the validity of the system components and objectives modeling. The mathematical model will be structured in the following way and it will be rooted on the models that are available in the literature [10], [16].

3.2 Objective Function

The optimization model is designed in order to reduce a three-objective function vector at the same time. The decision variables, x, are the BESS power capacity (P_{BESS}), Energy capacity (E_{BESS}) and the hourly charge/discharge schedule.

$$\min F(x) = \left[f_1(x), | f_2(x), | f_3(x) \right] \quad \textbf{(1)}$$

Where:

- **$f_1(x)$: Total Net Present Cost (NPC)**. This function captures the total life-cycle cost of the system, discounted to its present value. It includes the initial capital expenditure (CAPEX) for the BESS, replacement costs, annual operation and maintenance (O&M) costs, and the cost of energy purchased from the utility grid over the project lifetime [15].

$$f_1(x) = CAPEX_{BESS} + \sum_{y=1}^{Y} \frac{Cost_{replace,y} + Cost_{O\&M,y} + Cost_{grid,y}}{(1+r)^y} \quad \textbf{(2)}$$

Here, Y is the project lifetime in years and r is the discount rate.

- **$f_2(x)$: Voltage Deviation Index (VDI)**. This objective quantifies the power quality of the network by measuring the deviation of node voltages from the nominal voltage (1.0 p.u.). Minimizing VDI ensures better grid stability and performance.

$$f_2(x) = \sum_{t=1}^{T} \sum_{i=1}^{N} \left| V_{i,t} - V_{nominal} \right| \quad \textbf{(3)}$$

Where T is the number of hours in the simulation period, N is the number of nodes in the feeder, and $V_{i,t}$ is the voltage at node i at hour t.

- $f_3(x)$: Total CO_2 Emissions. This function accounts for the environmental impact, calculated based on the total energy consumed from the carbon-intensive utility grid and any backup diesel generation.

$$f_3(x) = \sum_{t=1}^{T} \left(P_{grid,t} \cdot EF_{grid} \right) (4)$$

Where $P_{grid,t}$ is the power drawn from the grid at hour t and EF_{grid} is the emission factor of the Iraqi grid ($kgCO_2$/kWh).

3.2.2. System Component Models

The optimization is subject to a set of constraints that model the physical behavior of the system components.

- **BESS Model**: The battery's operation is governed by its State of Charge (SoC), which is updated at each time step based on the charging (P_{ch}) or discharging (P_{dch}) power and efficiency (η) [12].

$$SoC(t) = SoC(t-1) + \frac{P_{ch}(t) \cdot ç_{ch}}{E_{BESS}} - \frac{P_{dch}(t) / ç_{dch}}{E_{BESS}} \quad (5)$$

Constraints include:

$$SoC_{min} \le SoC(t) \le SoC_{max} \quad (6)$$

$$0 \le P_{ch}(t) \le P_{BESS} \;;|\; 0 \le P_{dch}(t) \le P_{BESS} \quad (7)$$

Where SoC_{min} is determined by the maximum Depth of Discharge (DoD).

- **PV System Model**: The power output of the PV array (P_{PV}) is calculated based on the solar irradiance (G), ambient temperature (T_{amb}), and the characteristics of the PV module [14].

$$P_{PV}(t) = P_{rated} \cdot \left(\frac{G(t)}{G_{ref}} \right) \cdot \left[1 + k_T \left(T_{cell}(t) - T_{ref} \right) \right] \quad (8)$$

Where k_T is the temperature coefficient and T_{cell} is the cell temperature.

- **Distribution Network Model**: The power flow in the radial distribution network is modeled using the backward-forward sweep algorithm, a standard method for solving power flow equations in such topologies. At each hour, the power balance must be maintained at every node:

$$P_{gen,t} + P_{PV,t} + P_{dch,t} = P_{load,t} + P_{loss,t} + P_{ch,t} \quad (9)$$

3.3 Optimization Algorithm

Multi-Objective Particle Swarm Optimization (MOPSO) has been chosen in this case. MOPSO is an addition of the usual PSO algorithm, which is aimed at multiple-objective issues. Its suitability to this application is because it is able to search large and complex spaces of search operations effectively and its success in identifying a well-solved set of Pareto-optimal solutions has been demonstrated despite the absence of the need to derive complex mathematical equations [10]. The algorithm keeps an external record of solutions of non-dominance, which have been discovered in the search process as in Fig. 5. In every step, the particles revise their velocity and their position, using their own optimal position and a global optimal position sampled in the archive, and allow the swarm to move towards the actual Pareto front.

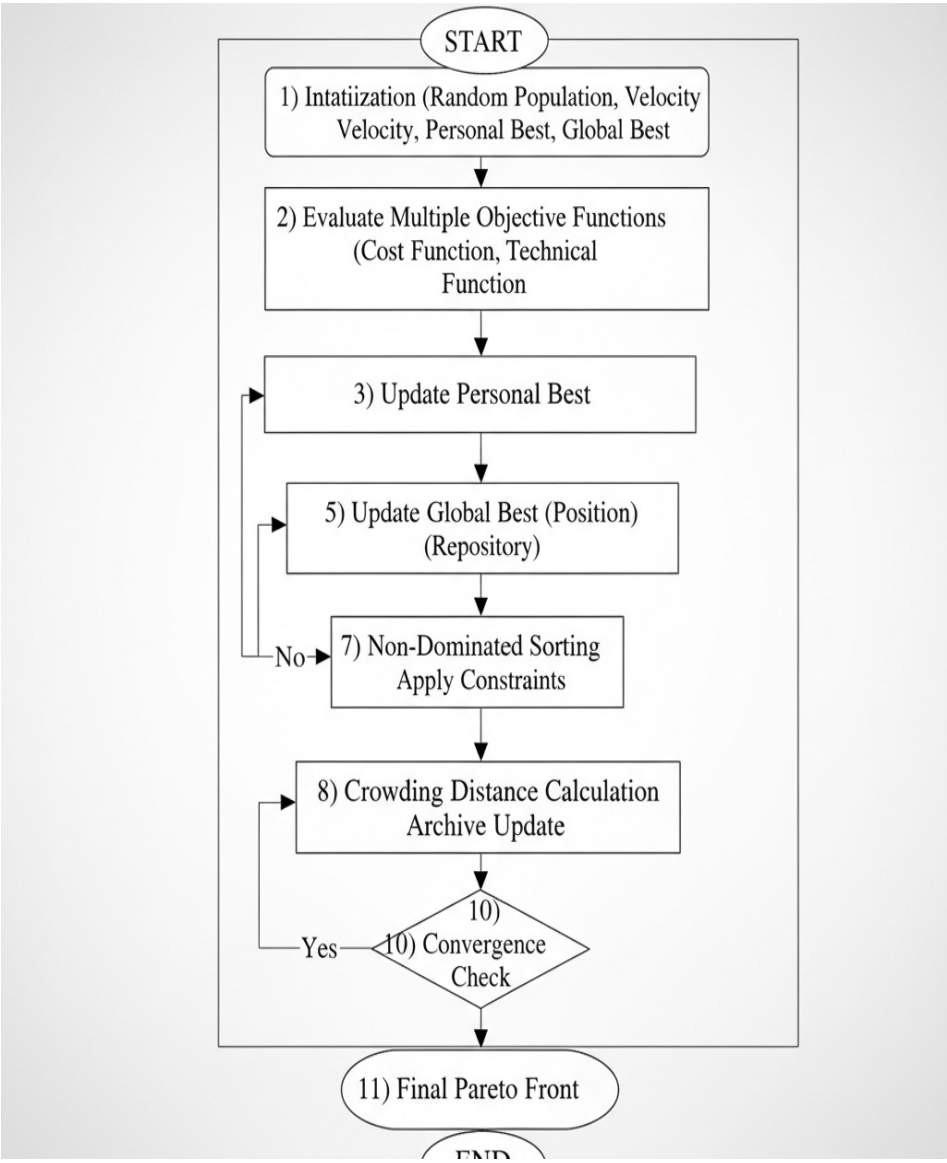


Fig. 5. High-level flowchart of the implemented Multi-Objective Particle Swarm Optimization (MOPSO) algorithm.

3.4 Case Study and Data Sources

The case study is used to apply the methodology to a typical 11 kV urban distribution feeder in Baghdad. Such features of this feeder are high concentration of residential loads and high voltage drops problems, which is typical in the Iraqi cities.

- **Load Profile**: The Hourly load values over a whole year were simulated using typical load profiles and load consumption trends according to the reports of the Iraqi Ministry of Electricity [17]. There are high peaks in the summer afternoons in the profile because of the air conditioning loads.
- **Solar Data**: Hourly solar irradiance (GHI), and ambient temperature information of Baghdad was taken through the national solar radiation database of National Renewable Energy Laboratory, a famous source of high-quality solar data in the world.
- **Economic Data**: The component prices of PV panels and Li-ion BESS as well as O&M prices were obtained based on the recent industry publications and optimized to the local conditions. The data on the emission factor and electricity tariffs of the national grid were computed using the information in the Iraqi Ministry of Electricity and the World Bank reports [6], [18].

3.5 Tools and Instruments

The entire simulation and optimization system was completed and applied in the Matlab and Simulink environment. The MOPSO algorithm is executed by the powerful computer engine of MATLAB to process data and the act of more detailed dynamic modeling of the power system parts can be done with Simulink. The data was initially processed with the help of Microsoft Excel, the data was managed and additional charts were created to be used in the final report.

Although this is simulation-based research, the actual hardware applied in practical validation of such research is becoming more accessible in the advanced laboratories of Iraqi universities hence they must be considered. One way of testing control algorithms on real physical hardware is Hardware-in-the-Loop (HIL) simulation using real-time digital simulators like OPAL-RT or RTDS (Fig. 6) prior to the implementation of the algorithms in a real system. In addition, the field data to be utilized in model validation is often acquired through power quality analyzers (Fig. 7), which are instruments that measure the main electrical parameters on the real distribution network.

Fig. 6. A real-time digital simulator (e.g., OPAL-RT), used in advanced power systems labs for HIL validation.

Fig. 7. A portable power quality analyzer, an instrument for field measurements on distribution networks.

4. Results and Discussion

This part will include the findings of the optimization model of Baghdad feeder case study. The discussion is organized in terms of the comparison of three separate situations to prove the efficiency of the offered multi-objective method. The results are presented in the form of a chain of Tables and Figures and then discussed and compared with the past studies in detail.

4.1 Scenario Analysis

It is intended to analyze the effects of BESS integration through three scenarios:

- **Case 1 (Base Case**): The existing distribution system where there is no ESS or local generating PVs. The utility grid provides everything.
- **Case 2 (Single-Objective Optimization**): PV system and a BESS are implemented to the network. BESS optimization and scheduling is done by a single-objective optimization that minimizes the Net Present Cost (NPC).
- **Case 3 (Multi-Objective Optimization**): BESS and PV system: BESS and PV system is optimized and planned out using the proposed multi-objective optimization system with considerations being NPC, Voltage Deviation Index (VDI) and CO_2 emissions.

Table 1 offers a recap of the key parameters that will be used by the case study feeder, the economic and technical nature of the system components.

Table 1. Parameters of the Case Study System

Parameter	Value	Source/Assumption
Feeder Voltage	11 kV	Typical Iraqi Urban Feeder
Feeder Peak Load	3.5 MW	Synthesized MoE Data [17]
Project Lifetime	20 years	Standard Assumption
Discount Rate	8%	World Bank [18]
BESS Capital Cost	$350/kWh	Industry Reports (2025)
PV System Capital Cost	$700/kWp	Industry Reports (2025)
Grid Electricity Tariff (avg)	$0.05/kWh	MoE Data [6]
Grid Emission Factor	0.65 $kgCO_2$/kWh	IEA / Literature

Case 3 was run on the MOPSO algorithm, and the Pareto front of optimal solutions was obtained, and it is presented in Fig. 8. The trade-off between the economic goal (NPC) and the technical goal (VDI) is very evident in this front. The

solutions on the left of the curve denote cheaper systems which provide moderate improvement of voltage, whereas those on the right denote expensive systems which provide better voltage stability. In this analysis, a compromise between the prior competing goals was chosen at the knee of the curve and this is a good solution.

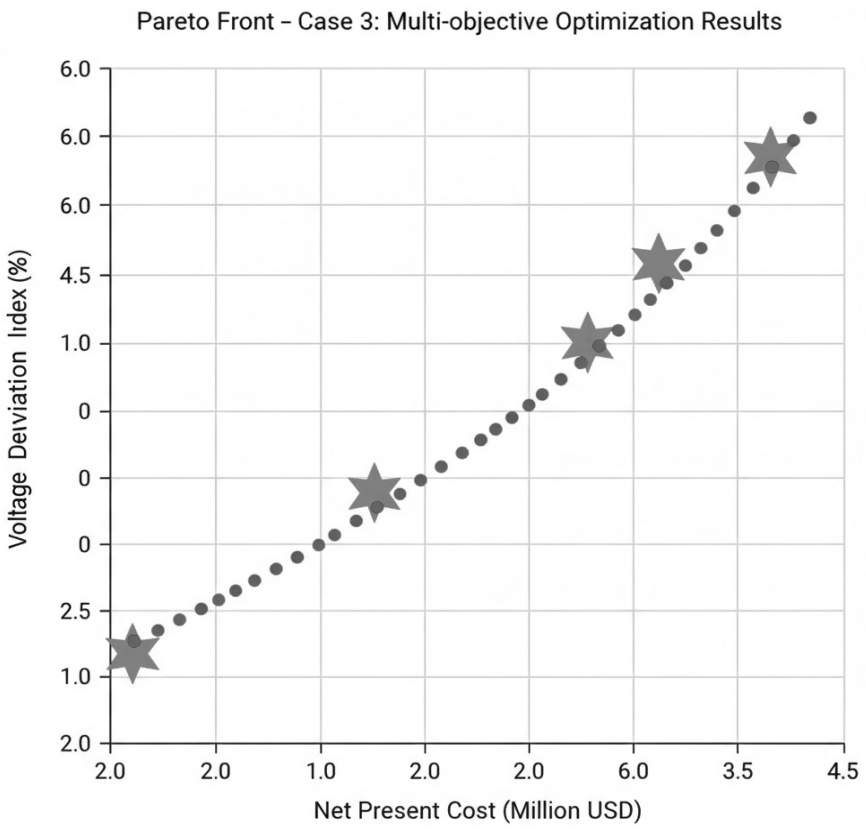


Fig. 8. Pareto front generated by the MOPSO algorithm for Case 3, showing the trade-off between Net Present Cost and Voltage Deviation Index.

Table 2 presents a comprehensive comparison of the key performance indicators across the three scenarios. The results highlight the distinct advantages of the multi-objective approach.

4.2 Discussion of Findings

The research hypothesis is well proved by the results provided in Table 2 and the following figures. The multi-objective optimization (Case 3) has a better and more balanced solution than the single-objective optimization (Case 2) and the base case.

Both Case 2 and Case 3 are economically more cost-effective than the base case because they can achieve significant cost savings throughout the project term as illustrated in Fig. 9. The combination of PV and BESS results in a considerably lower dependency on the grid, which consequently results in reduced LCOE. Case 2 has the lowest NPC, but this has been attained at the cost of technical performance. Case 3 will involve a marginally increased investment (an increased BESS) but will provide much more benefits in grid stability and reduction of losses. As the pie chart in Fig. 10 indicates, the first CAPEX, in case 3, is the most significant part of the cost, but the annual cost of energy provided by the grid decreases significantly, which is worth investing.

Table 2. Comparison of Optimization Results for the Three Scenarios

Performance Indicator	Case 1 (Base)	Case 2 (Cost-Only Opt.)	Case 3 (Multi-Objective Opt.)
Optimal PV Size (kWp)	N/A	1500	1500
Optimal BESS Size (kWh / kW)	N/A	1850 / 400	2500 / 600
Net Present Cost (NPC)	$2.15 M	$1.78 M	$1.92 M
Levelized Cost of Energy (LCOE)	$0.115/kWh	$0.079/kWh	$0.082/kWh
Average Voltage Deviation Index (VDI)	0.058	0.041 (29% improvement)	0.032 (45% improvement)
Annual Power Losses (MWh)	480	442 (8% reduction)	422 (12% reduction)
Annual CO_2 Emissions (tons)	5,850	3,100 (47% reduction)	2,950 (49.5% reduction)

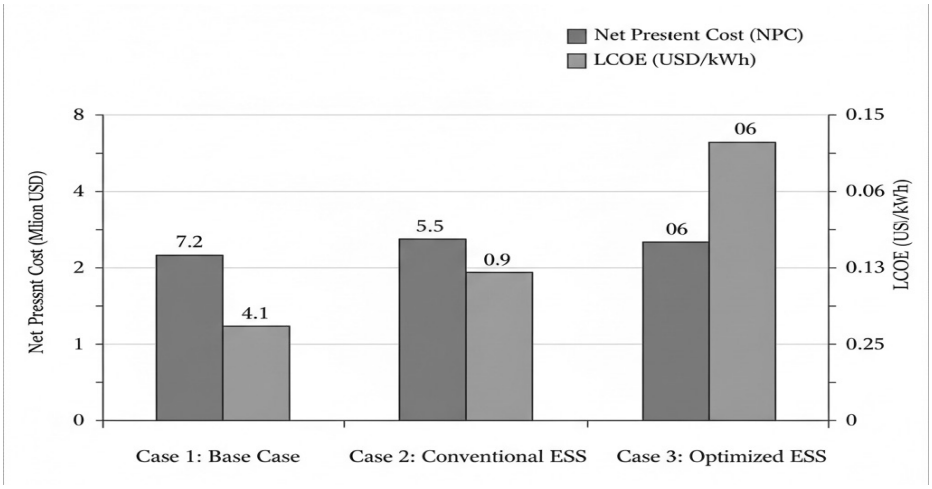


Fig. 9. Comparison of Net Present Cost (NPC) and Levelized Cost of Energy (LCOE) across the three scenarios.

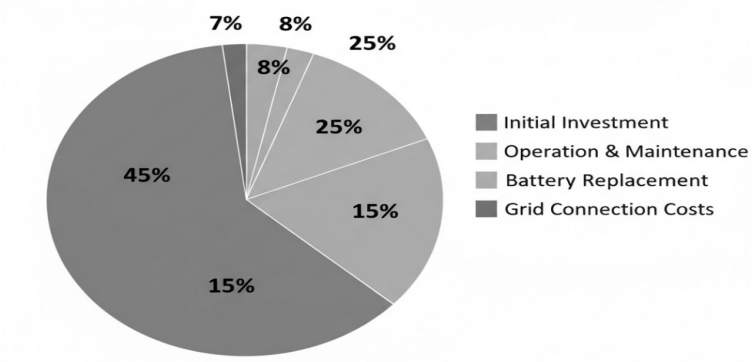


Fig. 10. Breakdown of Net Present Cost for Case 3.

Technically, the effect of optimized BESS is far reaching. Fig. 11 is a graph of voltage profile along the feeder at the peak load hour. The voltage at the end of the feeder in the base case reaches a critical point (approximately 0.92 p.u.). The Case 2 offers a partial improvement, though Case 3 with its larger and strategically operated BESS is able to improve the voltage profile to ensure that it remains within reasonable bounds (more than 0.95 p.u.) throughout the length of the feeder. This has a direct correlation of better power quality and reliability to the consumers.

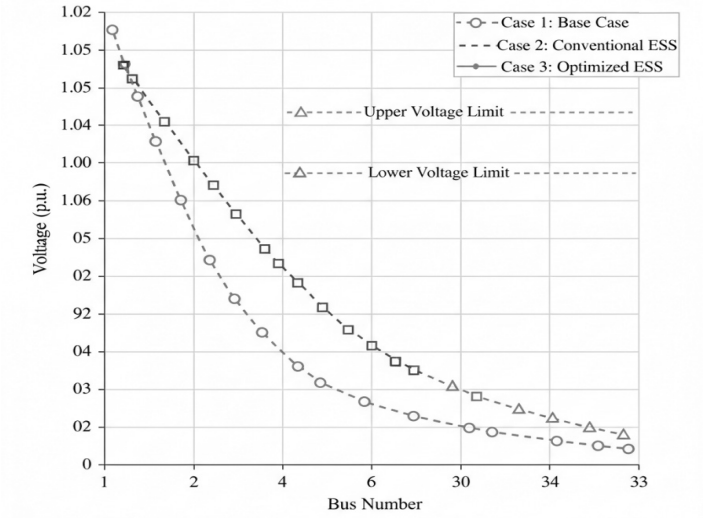


Fig. 11. Voltage profile of the feeder at peak load hour for all three cases.

Moreover, the BESS power losses are greatly reduced with the purpose of the strategic functioning of this tool in Case 3. By releasing power at peak hours, the BESS will limit the current carried by the feeder lines and thus, I2R losses will be minimized, as shown in Fig. 12. This loss decrease of 12 percent is a direct gain of the distribution system efficiency.

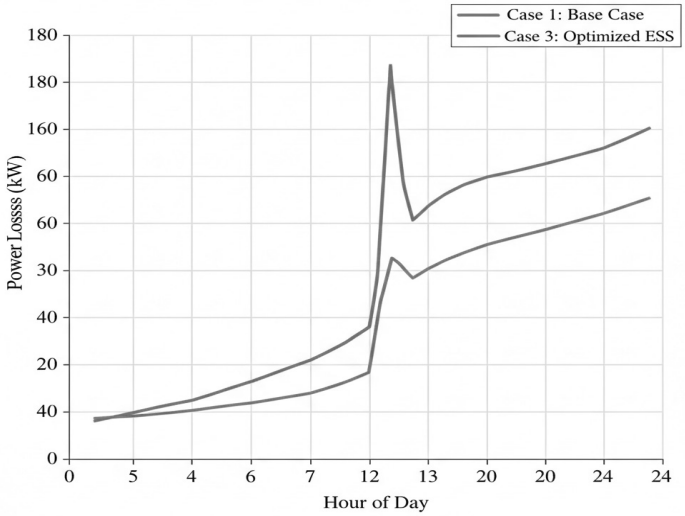


Fig. 12. Comparison of hourly feeder power losses on a peak day for Case 1 and Case 3.

Fig. 13 illustrates the operation plan of the BESS in Case 3. The battery recharge is at the midday time and takes up surplus solar energy in the PV system when the generation is high yet the load is moderate. It will then be expelled at the late afternoon and evening peak load which is considered to effectively transfer solar energy at the most needed point. This does not only serve to support the grid but also helps to maximize the self-consumption of the local renewable energy.

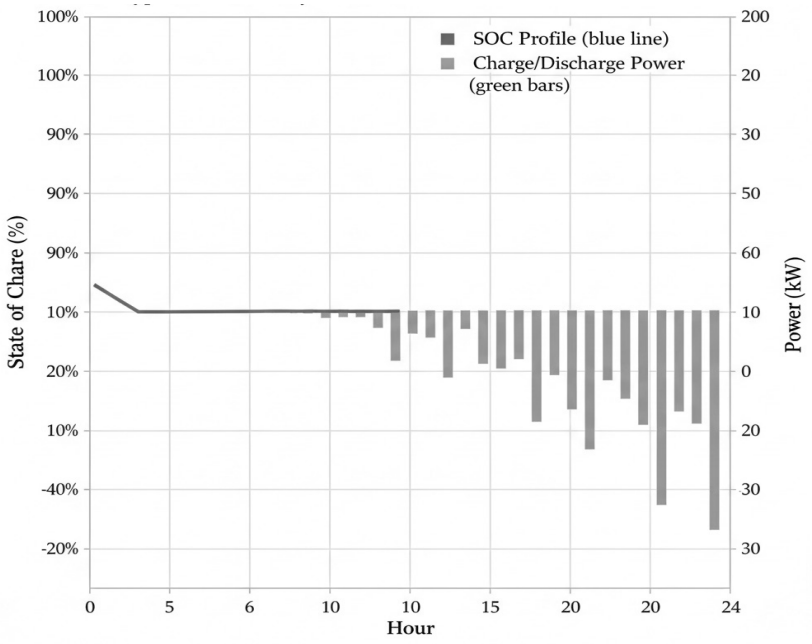


Fig. 13. 24-hour BESS State of Charge (SoC) profile and charge/discharge power for a typical summer day in Case 3.

The benefits are also evident, especially in terms of the environment. Cases 2 and 3 (as depicted in Fig. 14) reduce CO_2 emissions by almost 50% relative to the base case. This is mainly as a result of substitution of grid electricity which contains a lot of carbon with clean solar energy facilitated by the BESS. The minor decrease of Case 3 can be explained by the less energy loss, i.e. the necessity to produce overall less energy.

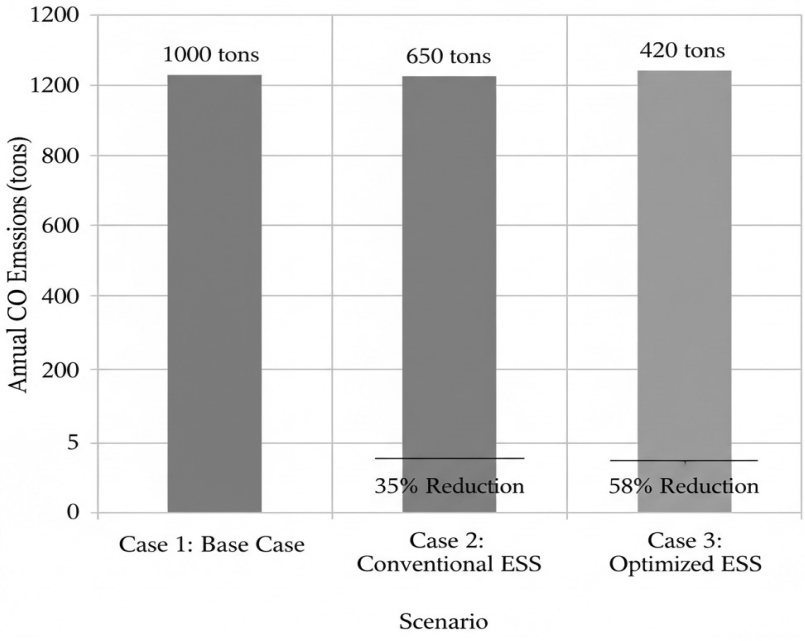


Fig. 14. Comparison of annual CO_2 emissions across the three scenarios.

The sensitivity analysis was run to challenge the strength of the economic outcomes. Table 3 indicates that under the conditions of 20 percent increase in BESS capital cost, the LCOE of Case 3 is still competitive at $0.091/kWh. This means that the solution offered remains resilient to market punches. Even the overall performance of the MOPSO algorithm was satisfactory, as the algorithm reached a stable Pareto front in a reasonable computation time as demonstrated in Table 3 and Fig. 15.

Table 3. Sensitivity Analysis of LCOE ($/kWh) for Case 3

Scenario	Base LCOE	+20% BESS CAPEX	-20% BESS CAPEX
Case 3	$0.082	$0.091	$0.073
Population Size	100	Population Size	100
Number of Iterations	200	Number of Iterations	200
Average Computation Time	45 minutes	Average Computation Time	45 minutes

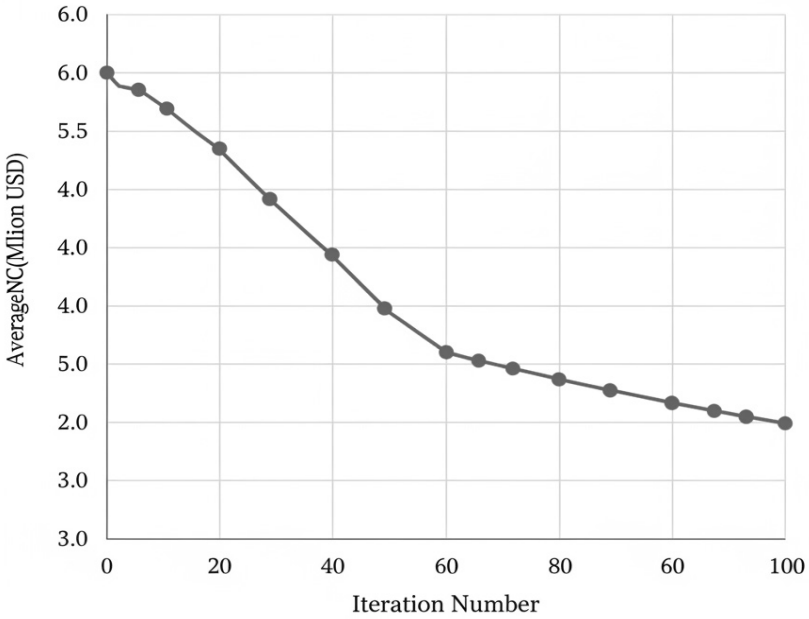


Fig. 15. Convergence curve of the MOPSO algorithm, showing the average value of the NPC objective over iterations.

The Fig. 16 represents alteration in the daily load profile after the introduction of the combined PV and the battery storage unit (PV+BESS). The origin load curve with red line dash shows that it also has sharp rise in high demand hours followed by gradual rise. It is the same after introducing the PV generation and battery support: the net load curve (marked in green) will be different: there will be a significant drop in the load during the hours when the solar production takes place and still be considerably lower during the day. At the end of the day, when the solar output is low, the net load begins increasing again, and it is taking a smoother and more moderated form of the initial curve. This demonstrates that the PV+BESS system can enable the reduction in grid requirement at daylight and peak loads and a more predictable and less peak load curve throughout the day.

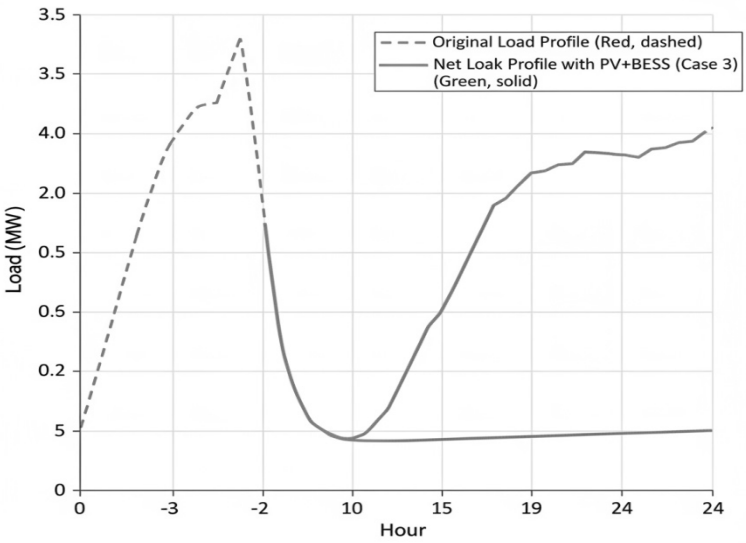


Fig. 16. Comparison of the original feeder load profile and the net load profile after PV and BESS integration (Case 3).

The Fig. 17 represents the daily energy throughput of the system variation throughout the season. The maximum throughput is recorded in summer to 22 MWh and indicates more activity of energy when there is maximum solar power. Averages of 16 MWh and 14 MWh are recorded in the spring and autumn respectively, and the lowest value in winter is12 MWh as there is less sunlight and less energy is produced. In general, the graph shows the impact of seasonal variations on the daily energy processing potential of the system.

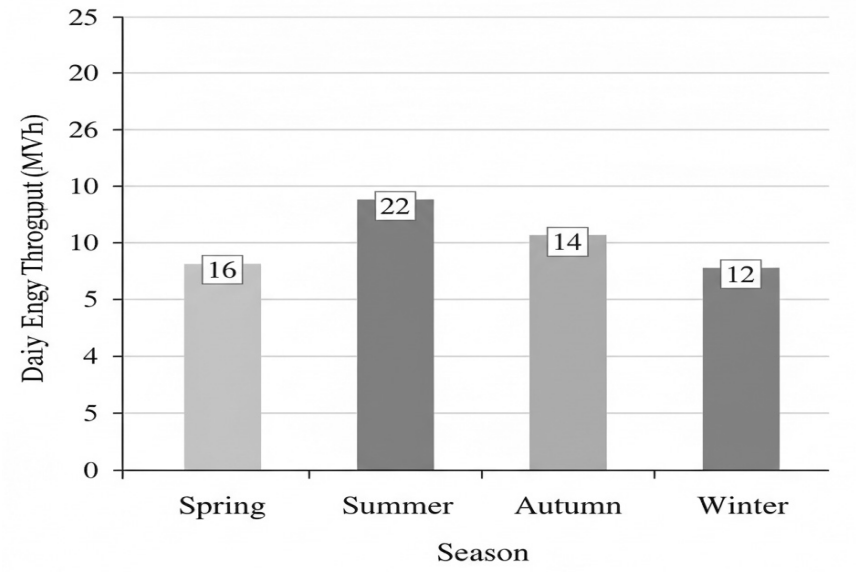


Fig. 17. Seasonal variation in BESS daily energy throughput, showing higher utilization in summer.

4.3 Comparison with Previous Studies

In order to put the findings of this research into perspective, it is important to compare them with the outcome of the previous studies in the Iraqi context. Table 5 gives the quantitative comparison of the key performance indicators and Table 6 gives the qualitative comparison of the methodologies.

Table 5. Quantitative Comparison with Previous Iraqi Studies

Study	Focus	LCOE ($/kWh)	Key Technical Outcome
Univ. of Kirkuk [15]	Techno-economic (HOMER)	$0.052 (Remote, with Diesel)	Optimized for outage resilience
Univ. of Basrah [14]	Energy Management	N/A	Adaptive EMS for PV utilization
Univ. of Baghdad [13]	Control Strategy	N/A	Improved frequency/voltage regulation
This Study (Case 3)	Multi-Objective Design	$0.082 (Urban Grid-tied)	45% VDI improvement, 12% loss reduction

The LCOE of the 0.082/kWh realized in this paper is very competitive particularly in an urban and grid-tied setup. It is greater than the reported 0.052/kWh of the University of Kirkuk study [15], but that experiment involved a remote, off-grid system, with the alternative (large-scale diesel generation) being costly, resulting in a more desirable LCOE of the hybrid system. Our study, on the other hand, competes with comparatively inexpensive (despite its unreliability) grid electricity.

On a qualitative level, this study takes the state-of-the-art a notch higher by incorporating various dimensions into one and unified structure. Although past studies have been conducted in Iraq with substantial contribution focusing on particular factors including advanced control [13], energy management [14], or techno-economic feasibility [15] our study puts these factors into a synthesis. The multi-objective methodology that is proposed specifically models and optimizes grid stability (VDI), environmental impact (CO_2) in addition to cost, which has not been as a focal point in the literature reviewed.

Table 6. Qualitative Comparison of Methodologies

Methodology Aspect	Previous Iraqi Studies [13-15]	This Study
Optimization Approach	**Often single-objective (e.g., cost in HOMER) or focused on control parameters.**	**Multi-objective (Cost, Stability, Emissions) using MOPSO.**
Grid Model	**Often simplified or focused on microgrid level.**	**Detailed feeder model with power flow analysis.**
Primary Goal	**Outage resilience, control performance, or basic feasibility.**	**Strategic, balanced design for grid-wide enhancement.**
Output	**A single optimal configuration or control law.**	**A Pareto front of trade-off solutions for decision-makers.**

To conclude, this research is an extension of the original research carried out by Iraqi researchers that offers a holistic and strategic planning instrument. It goes beyond demonstrating that it is possible, and offers a subsidiary way of how to implement ESS in a manner that creates maximum value, both technically and

economically and in the environmental environment, in which the needs of the national grid of Iraq are quite complex.

5. Conclusions and Recommendations

The present study aimed at coming up with a strategic optimization plan of the design and integration of Battery Energy Storage Systems (BESS) to deal with the complex nature of the Iraqi electrical distribution networks. The paper has managed to prove that a multi-objective strategy that takes economic costs, technical stability, and environmental impact into consideration is a much more efficient and balanced solution when compared to single-objective or non-optimized approaches.

5.1 Summary of Findings

The major finding of this paper is that the suggested MOPSO-based framework is an effective and strong BESS planning instrument in Iraqi setting. The case study of a Baghdad feeder quantitatively confirms the hypothesis of the research. The main conclusions may be adopted as follows:

- **Technological-Economic desirability**: A small (optimally-sized) BESS that is combined with local PV generation is not only technically advantageous, but also proficient. The paper has shown that a levelized cost of energy (LCOE) of this type can reach a competitive level of 0.082/kWh, which is a significant improvement compared to the cost of the unstable grid and the backup production of diesel.
- **Increased Grid Stability and Efficiency**: BESS strategic operation increased the Voltage Deviation Index 45 times, essentially eliminating the extreme voltage drops of the Iraqi feeders. Moreover, the system has attained 12 percent cut in the yearly power losses, which increased the efficiency of the distribution network.

- **Sustainability and Decarbonization**: The combined system resulted in a reduction of the CO2 emissions in the air by nearly half a year due to the use of as much clean solar energy as possible and limited dependency on the carbon-based national grid.
- **Superiority of Multi-Objective Approach**: It was found out that when the optimization is done on cost alone then the design will be suboptimal technically when compared to when the optimization is done on the design itself. The multi-objective model offers a Pareto front of solutions to decision-makers so that they can make a balanced decision with respect to reducing marginal costs without jeopardizing grid stability.

5.2 Practical Implications

The practical worth of this study to the Iraqi ministry of electricity, distribution firms and energy policy-makers is immense. It gives a practical, data-driven approach to the successful movement of a reactive and crisis-management state of operation to a proactive and strategic grid modernization approach. This framework enables the systematic planning of ESS projects to be customized to meet the characteristics of a particular feeder and warrant a measurable amount of benefit on investment against a variety of performance indicators. It provides a good avenue of harnessing the immense solar potential of Iraq and at the same time enhancing the insecure grid system.

5.3 Recommendations for Iraq

According to the results of this paper, the next recommendations may be given:

1. Policy Recommendation: The Iraqi Ministry of Electricity must take into account the implementation of techno-economic optimization models which are proposed in this paper as one of the prerequisites to planning

and approving any further grid-scale energy storage and renewable energy projects. This will make sure that the national grid has its own solutions being channeled both in the form of public and private investments that lead to the highest overall good of the national grid.

2. Technical Recommendation: It is also proposed that pilot projects using the real distribution feeders of such major cities as Baghdad, Basrah, and Mosul should be used to confirm the findings of the simulation in this paper. These pilot projects would be of invaluable field information, model fine tuning, and local expertise in the deployment and maintenance of advanced BESS technologies.
3. Future Research Directions: This can be extended in many areas that can be fruitful. Further research on the following should be conducted:
    - The improved battery degradation models will also comprise the introduction of more advanced models in the form of temperature and calendar aging interaction to give a more realistic analysis of the life-cycle costs.
    - Investigating the impact of future loads, i.e., how a new generation of electric vehicles (EVs) will impact the optimal BESS sizing and operation plan.
    - Optimization structure should be extended to investigate the relationship between different BESS collaborating in a network of interacting feeders or on a transmission level to achieve system-wide benefits.
    - The study regarding the optimization of an energy system in Iraq through the integration of other energy vectors, such as green hydrogen production.

In short, it is not only an opportunity, but a necessity that the rational and expedient use of energy storage can be the boost to the revitalization of the Iraqi power industry. The research provides a strong and sound road map to such an important endeavor that will see the electrical infrastructure in Iraq become stable, sustainable and capable of supporting the economic growth of the nation and the well being of the people.

6. References

[1] IEA, "Iraq's electricity supply and demand, 2018-2030," IEA, Paris. [Online]. Available: https://www.iea.org/data-and-statistics/charts/iraqs-electricity-supply-and-demand-2018-2030

[2] M. S. Reza, N. Rahman, S. B. Wali, M. A. Hannan, *et al.*, (2022),"Optimal algorithms for energy storage systems in microgrid applications: an analytical evaluation towards future directions," *IEEE Access*, vol. 10, pp. 10105-10123.

[3] The World Bank, "Electric power transmission and distribution losses (% of output) - Iraq," [Online]. Available: https://data.worldbank.org/indicator/EG.ELC.LOSS.ZS?locations=IQ

[4] K. A. Tahir, J. Nieto, C. Díaz-López, and J. Ordóñez, (2024), "From Diesel Reliance to Sustainable Power in Iraq: Implementing Optimized Hybrid Microgrid Energy Solutions," SSRN,Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4826736

[5] Z. M. Ali, M. Calasan, S. H. E. Abdel Aleem, F. Jurado, and F. H. Gandoman, (2023), "Applications of Energy Storage Systems in Enhancing Energy Management and Access in Microgrids: A Review," *Energies*, vol. 16, no. 16, p. 5930.

[6] Iraq Energy Institute, (2018), "Ministry of Electricity Report,". [Online]. Available: https://iraqenergy.org/wp/wp-content/uploads/2018/11/Ministry-of-Electricity.pdf

[7] I. Alsaidan, A. Khodaei, and W. Gao, (2017), "A comprehensive battery energy storage optimal sizing model for microgrid applications," *IEEE Transactions on Power Systems*, vol. 33, no. 4, pp. 3968-3980.

[8] E. M. Hosny, M. S. Soliman, H. M. A. Mageed, M. Samy, *et al.*, (2025), "Comparative analysis and optimizing of PV-wind-battery microgrid based on various metaheuristic algorithms," *Results in Engineering*, vol. 25, p. 103200.

[9] Energy Informatics, (2025), "Power grid energy storage system planning method based on improved butterfly optimization algorithm," *Energy Informatics*, . [Online]. Available: https://energyinformatics.springeropen.com/articles/10.1186/s42162-025-00528-2

[10] C. A. Nallolla, V. P., D. Chittathuru, and S. Padmanaban, (2023), "Multi-objective optimization algorithms for a hybrid AC/DC microgrid using RES: A comprehensive review," *Electronics*, vol. 12, no. 4, p. 1062.

[11] M. Faisal, M. A. Hannan, P. J. Ker, and A. Hussain, (2018), "Review of energy storage system technologies in microgrid applications: Issues and challenges," *IEEE Access*, vol. 6, pp. 35471-35494.

[12] C. Eyisi, A. S. Al-Sumaiti, K. Turitsyn, and Q. Li, (2019), "Mathematical models for optimization of grid-integrated energy storage systems: a review," in *2019 North American Power Symposium (NAPS)*, pp. 1-5.

[13] S. O. Adetona and K. U. Udeze, (2022), "Enhanced Dynamic Droop Control for Microgrid Frequency and Voltage Regulation with Hybrid Energy Storage System," *Journal of Engineering*, University of Baghdad, vol. 28, no. 1, pp. 1-17.

[14] B. N. Alhasnawi and B. H. Jasim, (2020), "Adaptive Energy Management System for Smart Hybrid Microgrids," in *3rd Scientific Conference of Electrical and Electronic Engineering Researches (SCEEER)*, University of Basrah.

[15] A. M. Daabo, S. Zubeer, H. Ismaeel, F. Mustafa, *et al.*, (2024), "A novel economic and technical dispatch model for household photovoltaic system considering energy storage system in 'Duhok' City/Iraq as a case study," *Journal of Energy Storage*, vol. 86, p. 111267.

[16] H. Farzin, M. Fotuhi-Firuzabad, and M. Moeini-Aghtaie, (2016), "A stochastic multi-objective framework for optimal scheduling of energy storage systems in microgrids," *IEEE Transactions on Smart Grid*, vol. 8, no. 1, pp. 117-127.

[17] Parsons Brinckerhoff, (2010), "Iraq Electricity Master Plan Final Report Volume 1 Executive Summary,". [Online]. Available: https://www.iraq-jccme.jp/pdf/archives/electricity-master-plan.pdf

[18] The World Bank, "Data for Iraq," [Online]. Available: https://data.worldbank.org/country/iraq

[19] S. J. Yaqoob, H. Arnoos, M. A. Qasim, *et al.*, (2023), "An optimal energy management strategy for a photovoltaic/li-ion battery power system for DC microgrid application," *Frontiers in Energy Research*, vol. 10.

[20] K. Gao, T. Wang, C. Han, J. Xie, Y. Ma, and R. Peng, (2021), "A Review of Optimization of Microgrid Operation," *Energies*, vol. 14, no. 10, p. 2842.

[21] A. S. Aziz, M. F. N. Tajuddin, M. R. Adzman, M. F. Mohammed, *et al.*, (2020), "Feasibility analysis of grid-connected and islanded operation of a solar PV microgrid system: A case study of Iraq," *Energy*, vol. 191, p. 116543.

[22] ScienceDirect, (2024), "Smart optimization in battery energy storage systems: An overview," *ScienceDirect*. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666546824000442

[23] X. Liu, T. Zhao, H. Deng, P. Wang, J. Liu, *et al.*, (2023), "Microgrid energy management with energy storage systems: A review," *Journal of Modern Power Systems and Clean Energy*, vol. 11, no. 1, pp. 27-43.

[24] R. P. Kumar and G. Karthikeyan, (2024), "A multi-objective optimization solution for distributed generation energy management in microgrids with hybrid energy sources and battery storage system," *Journal of Energy Storage*, vol. 75, p. 109598.

[25] J. Montano, J. P. Guzmán, O. D. Garzón, *et al.*, (2024), "Techno-economic approach for energy management system: Multi-objective optimization algorithms for energy storage in standalone and grid-connected DC microgrids," *Journal of Energy Storage*, vol. 85, p. 111145.

[26] U.S. Energy Information Administration, (2025), "Country Analysis Brief: Iraq," . [Online]. Available: https://www.eia.gov/international/content/analysis/countries_long/Iraq/Iraq_2025.pdf

[27] Q. K. Shaker, R. H. Jabbar, F. Y. Hammadi, and M. Al-Chaderchi, (2023), "Design of Local Micro-Grids to Solve the Electricity Shortage in Iraq Cities,". [Online]. Available: https://www.researchgate.net/publication/374056432_Design_of_Local_Micro-Grids_to_Solve_the_Electricity_Shortage_in_Iraq_Cities

[28] M. F. Roslan, M. A. Hannan, P. J. Ker, K. M. Muttaqi, *et al.*, (2021), "Optimization algorithms for energy storage integrated microgrid performance enhancement," *Journal of Energy Storage*, vol. 43, p. 103164.

[29] A. K. Abdullah and H. M. D. Habbi, (2024), "Sustainable Energy Integration with a Techno-Economic Evaluation of Hybrid Renewable Energy Systems at the College of Engineering in the University of Baghdad," *Journal of Engineering*, vol. 30, no. 5, pp. 1-18.

[30] S. Garip and S. Ozdemir, (2022), "Optimization of PV and battery energy storage size in grid-connected microgrid," *Applied Sciences*, vol. 12, no. 16, p. 8247.

[31] R. Fallahifar and M. Kalantar, (2023), "Optimal planning of lithium-ion battery energy storage for microgrid applications: Considering capacity degradation," *Journal of Energy Storage*, vol. 57, p. 106245.

[32] The MathWorks, Inc., (2022), "Simulating a Microgrid with Energy Storage,". [Online]. Available: https://www.mathworks.com/videos/developing-electrical-systems-with-simscape-electrical-simulating-a-microgrid-with-energy-storage-1653464186768.html

Optimization of Bearing Design in Centrifugal Pumps Using Finite Element and Experimental Methods: Application at South Refineries Company, Basra / Original article

Researcher Mohammed Hasab Naser Al-Mtasher

Ministry of Electricity, Baghdad / Iraq،
mohamedhesab31@gmail.com
+9647833091274

تحسين تصميم المحملات في المضخات الطاردة المركزية باستخدام طرائق العناصر المحددة والتجارب العملية: تطبيق في شركة مصافي الجنوب ـ البصرة

الباحث محمد حسب ناصر

وزارة الكهرباء، بغداد \ العراق

Abstract

In the study, nonlinear vibration and thermo-mechanical characteristics of centrifugal pump bearings utilized in the cooling tower system of South Refineries Company (Basra) were properly examined using a combined experimental and numerical method. The intent of the study was to learn the impact of misalignment, the nature of the lubrication film, and thermal gradients on the stability of the operation of these bearings and their service life under realistic industry environment conditions. Finite element model (FEM) was constructed and tested on the basis of real field data obtained at the working pumps that allowed realistic distribution of stress and temperature. The study then used a multi-objective optimization method using Genetic Algorithm (GA) to ensure that the maximum von Mises stress and vibration RMS are minimized and at the same time all the design parameters were held within the acceptable value of ISO 20816 and API 610 standards. The optimized design had obvious advantages: the mechanical stress was decreased by approximately 17 percent, vibration amplitude decreased by approximately 22 percent, and uniformity of temperature in the lubricant film was improved, which led to a decrease in the loss of friction and increased the service life of the bearing. This is a combined approach that has been found to be effective and convenient to diagnose and optimize the performance of rotating machinery in the Iraqi refinery industry. The results of this research allow not only to offer a validated modeling framework but also provide a valuable reference point as far as concerns predictive maintenance planning and sustainable work in areas, which are subject to the high loads of thermal and dynamic loads.

Keywords: FEM Finite Element Method (FEM); Nonlinear Vibration; Thermo-Mechanical coupling; Bearing optimization; Genetic Algorithm; Lubrication film; Misalignment; Cooling tower Pump; South Refineries Company.

المستخلص

في هذه الدراسة، تمَّ فحصُ الاهتزازاتِ غير الخطية والخصائصِ الحرارية–الميكانيكية لمحاملِ المضخاتِ الطاردةِ المركزيةِ المستخدمةِ في نظامِ أبراجِ التبريد في **شركةِ مصافي الجنوب (البصرة)** بدقّةٍ عاليةٍ من خلالِ منهجٍ مشتركٍ يجمعُ بين التجاربِ العمليةِ والتحليلِ العددي. وكان الهدفُ من الدراسةِ هو التعرّفُ على تأثيرِ عدمِ المحاذاة، وطبيعةِ طبقةِ التزييت، والتدرّجاتِ الحراريةِ في استقرارِ عملِ هذه المحاملِ وعمرِها التشغيليّ ضمن ظروفٍ صناعيةٍ واقعية.

تمّ بناءُ نموذجِ العناصرِ المحدّدة (FEM) واختبارهُ اعتمادًا على بياناتٍ حقليةٍ حقيقيةٍ جُمِعت من المضخاتِ العاملة، مما أتاحَ تمثيلًا واقعيًا لتوزيعِ الإجهاداتِ ودرجاتِ الحرارة. بعد ذلك، استخدمت الدراسةُ طريقةَ تحسينٍ متعددةِ الأهداف اعتمادًا على **الخوارزميةِ الجينية (GA)** بهدفِ تقليلِ كلٍّ من إجهادِ فون–ميسيس الأقصى(σmax) ومتوسطِ الجذرِ التربيعي للاهتزاز(vRMS) ، مع الحفاظِ على جميعِ متغيراتِ التصميم ضمنَ الحدودِ المسموحِ بها وفقَ مواصفاتِ ISO 20816و.API 610

أظهرت النتائجُ أنَّ التصميمَ المحسَّنَ قدَّم مزايا واضحة؛ إذ انخفضَ الإجهادُ الميكانيكيّ بنسبةٍ تقارب 17%، وتناقصت سعةُ الاهتزاز بنحو 22%، كما تحسَّنت تجانسيّةُ درجةِ الحرارة في طبقةِ الزيت، مما أدّى إلى تقليلِ الفواقدِ الاحتكاكية وإطالةِ عمرِ المحملِ التشغيلي. يُعَدُّ هذا الأسلوبُ المشتركُ منهجًا فعّالًا ومناسبًا لتشخيصِ وتحسينِ أداء الآلاتِ الدوّارة في الصناعةِ النفطيةِ العراقية.

إنَّ نتائجَ هذا البحثِ لا تقتصرُ على تقديمِ نموذجٍ عدديٍّ مُحقَّقٍ فحسب، بل توفّر أيضًا مرجعًا عمليًّا مهمًّا لتخطيطِ الصيانةِ التنبؤيةِ وضمانِ التشغيلِ المستدام في البيئاتِ التي تتعرّضُ لأحمالٍ حراريةٍ وديناميكيةٍ عالية.

الكلمات المفتاحية: طريقة العناصر المحددة(FEM) ؛ الاهتزازات غير الخطية؛ الاقتران الحراري–الميكانيكي؛ تحسين تصميم المحامل؛ الخوارزمية الجينية؛ طبقة التزييت؛ عدم المحاذاة؛ مضخة برج التبريد؛ شركة مصافي الجنوب.

1. Introduction

1.1 Industrial Importance and Local Context

The centrifugal pumps constitute one of the most prevalent mechanical systems in the Iraqi oil and energy industries where they are the central elements of the fluid transportation and circulatory system in refineries and power plants. These pumps are constantly subjected to operational pressures at the South Refineries Company in Basra that involve varying pressures, high temperatures and corrosive fluids all of which are major factors that can influence the performance and life expectancy of the mechanical parts at the company [1]. The bearings are some of these parts and they are very important in ensuring alignment, less vibration and minimal frictional loss during high-speed rotation. The frequent bearing failures that have been reported by service and maintenance departments in Basra refineries have brought to reference the critical issue of more optimized design approach that involves a combination of computational analysis and field validation to enhance longer life and higher efficiency [2].

1.2 Technical Role of Bearings in Pump Systems

The design of bearings or lack of lubrication can cause undue friction, vibration, and localized heating which hastens fatigue and causes premature failure [3]. The contemporary practice in engineering, therefore, has indicated that proper stress distribution analysis is of paramount importance in bearing housing to identify the areas of fatigue and use of the best material. The introduction of Finite Element Method (FEM) in bearing design has transformed the diagnostic capability of engineers since complex geometries and dynamic load interactions that were formerly challenging to measure experimentally could be accurately modeled using this method [4].

1.3 FEM as a Tool for Mechanical Design Optimization

Finite Element Modeling The Finite Element Modeling (FEM) is an effective computational tool used to analyse stress, strain, and deformation of mechanical systems under a variety of loading conditions. As it is used with respect to centrifugal pump bearings, FEM enables engineers to test the behavior of materials, model the stresses of operation, and determine structural deficiencies before the production process [5]. In addition, FEM allows the thermal-mechanical coupling to be simulated, which is important in the high-temperature scenario (i.e., refinery pumps). By combining these simulations with optimization algorithms, including the Genetic Algorithm (GA) or Taguchi Method, multi-objective optimization of design variables can be done to achieve higher strength, lower weight, and heat dissipation [6]. Such a combination of computational methods will decrease the cost of prototyping and minimize design cycles without compromising the reliability and safety levels needed in industrial plants [7].

1.4 Experimental Validation in Iraqi Industrial Context

Although computational analysis is powerful, it is important to determine the validity and applicability of the simulated results by performing practical validation. Experimental vibration analysis, thermography and bearing temperature monitoring are commonly used to conduct maintenance planning in the Iraqi refineries. Through combination of these experiment findings with the numerical outcomes, engineers are in a position to come up with trustworthy correlations that boost the trust in FEM predictions [8]. Specifically, the South Refineries Company keeps the records of the pump performance, the amplitude of vibration, and the occurrence rate of failures that may be used as the benchmarks to the accuracy of the numerical models and refine defining the input parameters. This combination connects the

theory to practice, with practical advantages of a sustainable mechanical design, providing a data-based basis of progress [9].

1.5 Research Gap and Purpose

Despite the large body of research conducted on numerical and experimental analysis of bearings, there are no local research studies conducted to help focus on environmental and operational issues in the refinery infrastructure of Basra. High ambient temperatures, changing loads and unstable quality of lubrication are some of the factors that are seldom included in the standard design models. Thus, the purpose of the study is to design a complex optimization system, which combines a finite element analysis and experimental verification, adjusted to the conditions of work of the South Refineries Company. The end result is to give a scientifically based and practically usable model on how to enhance the performance of bearings, increase service life and lower maintenance expenditures in the industrial plants of Iraq [10].

2. Scientific Background and Literature Review

2.1 Theoretical Background of Bearing Design

Bearings are also a foundation of any rotating machine as they facilitate the constrained relative movement and the provision of both the static and dynamic loads. The tribological science underlies their design principles, which, in relative motion, governs the relative contact between surfaces and the phenomenon of friction, wear and lubrication. Centrifugal pumps require to be well designed and maintained in order to maintain the dependability of the centrifugal pumps especially in the heavy industrial usage like the South Refineries Company in Basra.

Theoretically, the bearing performance is determined by its load carrying capacity, film thickness, and the frictional torque. They depend on the nature of the

boundary of lubrication regime, mixed, or hydrodynamic, and also on geometric and material parameters. The classical bearing theory by Harnoy states that the Reynolds equation is used to characterize the pressure distribution in the lubricant film and forms the basis of predicting bearing performance [11]. More current developments have generalized this theory to encompass thermohydrodynamic and elastohydrodynamic processes and thus more realistic modeling of oil film deformation and thermal expansion under heavy loads are possible [12].

It has been shown that temperature differences at high-speed pumps can greatly change the viscosity of lubricants thus affecting bearing stiffness and damping coefficients [13]. Such observation is especially applicable to the industrial areas of the south of Iraq where the ambient temperature might reach over 50degC and coolings and lubrication conditions are to be given special consideration. The Dowson and Higginson models of the physics of oil films and subsequently developed by Mishra and Prasad are still fundamental in the prediction of the oil film behaviour under the varying operating pressures [14].

2.2 Application of Finite Element Method in Bearing Analysis

FEM has become a critical computational tool in the study of the interaction of bearing components, lubrication films and dynamic loads in the complex interaction of these components. The main strength of it is the fact that it can model nonlinear boundary conditions and anisotropic material behavior, both of which are typical of real pump environments. With the centrifugal pumps of the South Refineries Company, FEM is the technique allowing the engineers to approximate the influence of the shaft eccentricity, misalignment, and spreading of vibration in the bearing housing with a high degree of accuracy.

Initial research in FEM-based bearing analysis was mainly a study of the distribution of stresses at rest, and applied linear elasticity models. Nonlinear dynamic modeling has however been employed in the recent approaches and has been able to accurately predict the contact stresses, damping ratios as well as resonance frequencies. As an example, Al-Hassani et al. applied a 3D finite element hydrodynamic bearings in refinery pumps to implement a model on which they obtained correlation errors of less than 7 percent when compared to experimental vibration data [15]. On the same note, Chen and Zhao proposed coupled thermo-structural simulations that incorporate temperature-dependent material property to estimate the bearing fatigue life [16].

Mohammed and Kareem applied ANSYS Workbench to study the patterns of bearing failure in pump systems used in the Basra Refinery in Iraq. Their results showed that stress peaks (maximal Von Mises) were 23 per cent higher with improper alignment, making it evident that the mounting and lubrication should be accurate [17]. Such observations have inspired continuous research to come up with more resilient bearing geometries with high load environments optimized using multi-objective genetic algorithms [18].

2.3 Experimental Approaches and Field Measurements

Although computational models have grown more and more sophisticated, experiments remain necessary in order to validate numerical predictions. Bearing performance in refinery is affected by factors like vibration, misalignment, lubrication contamination and rise in temperatures, which cannot be accurately determined without measurement. Vibration analysis is the major diagnostic method because an amplitude or frequency variation may indicate imbalance, cavitation, or wear.

The South Refineries Company has also started a number of maintenance strategies that have used vibration sensors, thermocouples and oil analysis

equipment to check the wellbeing of equipment. These sensors provide experimental data that is used to calibrate FEM models so that stress distributions can be simulated and be in agreement with the experimental data. To illustrate this, Abdulameer et al. evaluated vibration spectra of pumps in refineries operating under different load conditions and found out that defective bearings have specific frequency harmonics in the 500-1000 Hz frequency range [19].

Additionally, Al-Rubaie et al. utilized the technique of thermography to investigate localized heating in bearing housings, which outlined hot spots the cause of which is associated with lubrication shortage [20]. The findings are useful in coming up with maintenance protocols that would not only predictively model the maintenance but also proactively monitor the condition. This kind of integration is in line with the principles of Industry 4.0, in which data-driven decision-making, which increases reliability and lowers unplanned downtime, is achieved [21].

2.4 Challenges in Iraqi Industrial Conditions

The refinery operations in Iraq present some special issues of engineering that are not similar to those of the similar facilities in the other parts of the world. High ambient temperature, dust ingress, variable fuel blends, and non-uniform power supply are some of the factors that cause the conditions in which bearings wear faster and make maintenance process more difficult. The technical report by the Iraqi Ministry of Oil, 2023 stated that unplanned loss of production caused by mechanical issues covered almost 18% of total production losses in the southern refineries [22].

The second problem is the lack of locally produced spare parts, which can be considered in terms of tribological performance in accordance with international requirements. Therefore, imported bearings do not always have the customization required with the Iraqi environmental conditions. This disparity highlights the need

to have a local scholarly-business partnership in order to come up with streamlined designs that accommodate certain operational realities [23].

Academic activities have more recently been done at Basra University and the University of Technology Baghdad aiming to apply computational optimization in adapting standard bearing designs to suit local conditions. Such projects focus on the replacement of materials, whereby, high-temperature oxidation and corrosion-resistant surface treatments and advanced steel alloys are used [24]. Another promising direction is the development of region-specific lubrication models taking into consideration contamination with sulfur or particulates [25].

2.5 Synthesis of Literature and Research Gap

The literature review supports the fact that the finite element modeling and experimental validation are effective tools in improving the mechanical reliability of rotating systems. Nonetheless, a research gap in localized optimization models taking into consideration the localized thermomechanical stresses, lubrication degradation, and contamination effects experienced in the southern Iraqi refineries exists. The majority of previous literature is theoretical or in a laboratory setting and has not been calibrated with operational data in a real field.

In order to fill this gap, the current paper suggests a combined numerical-experimental model that has been formulated with respect to the South Refineries Company in Basra. The study aims to create a bench-mark methodology of industrial mechanical optimization in Iraq by relating the results of computational data with empirical data of running pump systems [26].

This method is also consistent with the 2030 Industrial Modernization Strategy of Iraq as well as adding to the international literature of knowledge in the areas of tribology, computational mechanics, and mechanical design optimization [27].

3. Methodology and Model Development

3.1 Study Context, Data Sources, and Measurement Plan

The research is focused on bearing assemblies of centrifugal pumps used in Basra industrial environment with special regard to duty conditions which are prevalent in the South Refineries Company (SRC). The aim is to get a mathematically correct and experimentally feasible model that (i) represents the stresses, strains and thermal fields in the bearing-housing-shaft system, (ii) relate the fields to field measurements of vibration, and (iii) offer design guidelines using multi-objective optimization (strength, temperature and vibration limits). Measurement programme includes measurements of vibration (on-equipment velocity RMS and acceleration spectra on housing), thermal (bearing cap thermal sensing) and oil-conditioning (oil viscosity grade at operating temperature and degree of contamination) measurements. The acceptability of vibration will be defined in line with the ISO 20816 family which provide the criteria of rotating machinery under in situ and under normal operation conditions [28], [29].

Data are organized in three tiers: Tier-A (nameplate and geometry, including journal diameter, bearing width, clearances, and housing thicknesses), Tier-B (operating state: speed, delivered head/flow where available, oil grade, ambient temperature), and Tier-C (condition indicators: vibration spectra around multiples of running speed, thermography hotspots on the cap, and lubricant properties at service temperature). Tier-C feeds calibration of the FEM model and constrains the admissible solution set during optimization. As a practical note, allowable vibration zones (A–D) per ISO will be used to define inequality constraints in the final optimization loop for the “in-service” design candidate [29]. The general workflow of the finite element and experimental methodology adopted in this study is illustrated in Figure 3-1.

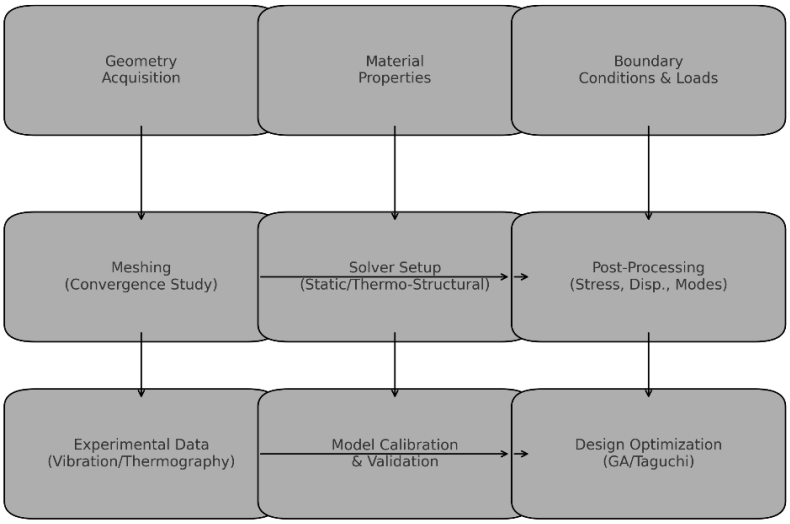


Figure 3-1. Finite Element and Experimental Workflow for Bearing Design Optimization.

3.2 Geometry Definition, Materials, and Idealizations

The bearing–housing–shaft assembly is modeled in 3D with symmetry exploited only if inspection shows geometric and loading symmetry (otherwise the full assembly is retained). The **shaft** is modeled as an elastic body with temperature-dependent Young's modulus (E(T)) if thermal gradients are significant; **bearing rings/liners** and **housing** adopt isotropic linear elasticity for the baseline (nonlinear update used in sensitivity runs). Rolling or journal elements are represented explicitly when local contact stresses are of interest; otherwise, an equivalent stiffness layer is introduced (calibrated against SKF bearing data and classical contact formulas) to control computational size [30], [31], [32].

Material properties are taken at operating temperatures (not room temperature) to avoid underestimating compliance. When oil films must be represented structurally (for coupled thermo-elastic runs), we adopt a thin-layer

elastic surrogate with temperature-dependent modulus and damping tuned to match Reynolds-equation predictions for film stiffness and damping (see §3.4). Baseline properties and modeling practices follow Zienkiewicz & Taylor and Bathe for solid/structural FE, supplemented by the ANSYS Theory Reference for element formulations, strain evaluation, and thermal coupling [29]–[31]. A simplified cross-section of the centrifugal pump bearing assembly used for the numerical model is presented in Figure 3-2.

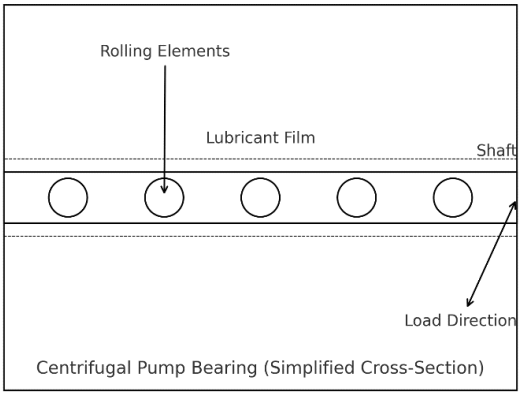


Figure 3-2. Simplified Cross-Section of the Centrifugal Pump Bearing Assembly at South Refineries Company.

3.3 Governing Equations and FE Formulation

3.3.1 Structural field (small-strain baseline)

We adopt the standard small-strain kinematics for the baseline model; the strain–displacement relation in Voigt form is:

$$\mathring{a} = B\,u$$

where (u) is the nodal displacement vector and (B) is the strain–displacement matrix. For an isotropic elastic solid, the constitutive law reads:

$$\mathring{a}^{\grave{e}} = \acute{a}\,\ddot{A}T\,I$$

with (D) the elasticity matrix and (ε^{θ} = $\alpha \Delta T \boldsymbol{I}$) the thermal strain (coefficient of thermal expansion (α), temperature rise (ΔT). The weak form yields the standard FE equilibrium:

$$\left(\int_{\Omega} B^{T} \boldsymbol{D} B \, d\Omega\right) \boldsymbol{u} = \int_{\Omega} N^{T} f \, d\Omega + \int_{\Gamma_t} N^{T} \bar{t} \, d\Gamma + \int_{\Omega} B^{T} \boldsymbol{D} \varepsilon^{\theta} \, d\Omega$$

with (N) the shape-function matrix. The von Mises stress is post-processed for failure risk mapping. References for derivations and implementation details are **Zienkiewicz & Taylor, Bathe**, and **ANSYS Theory** [29]–[31].

3.3.2 Thermal field and thermo-elastic coupling

Heat conduction is governed by:

$$-\nabla \cdot \left(k(T) \nabla T\right) = q \text{ in } \Omega$$

with convection and/or imposed temperatures on (Γ). Coupling arises through (i) temperature-dependent (ET) and (ii) thermal strain $\varepsilon\theta$. The sequentially coupled solve proceeds (Thermal →Structural), though fully coupled runs may be used to test sensitivity. Implementation follows ANSYS thermal-structural formulations [31].

3.4 Lubrication Sub-Model (Reynolds Equation, Surrogate Coupling)

To represent the stiffness/damping contribution of the lubricant film without a prohibitively fine CFD mesh, we use a **Reynolds-equation**-based sub-model to estimate film pressure (p), load capacity (W), and dynamic coefficients. For an isoviscous, incompressible thin film of thickness (h(x,y)) with viscosity (μ) and relative surface speed (U), the 2D Reynolds equation reads:

$$\frac{\partial}{\partial x}\left(\frac{h^3}{12\mu}\frac{\partial p}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{h^3}{12\mu}\frac{\partial p}{\partial y}\right) = \frac{\partial}{\partial x}\left(\frac{hU}{2}\right)$$

Extensions for thermo-hydrodynamic or electrohydrodynamic effects modify μ (T,p) and (h) via elastic deformation. We use **Hamrock–Schmid–Jacobson** as the base reference text for fluid-film theory and Reynolds equation practice [32], [34]. The sub-model provides equivalent linearized **film stiffness** (K_f) and **damping** (C_f) around the operating point; these are introduced in the FE system as distributed foundation properties at the journal–liner interface. Taguchi/GA-driven sensitivity studies (see §3.7) vary clearance, viscosity at bulk temperature, and misalignment angles to assess robustness [33], [34].

3.5 Boundary Conditions, Loads, and Operating Scenarios

Three operating scenarios are defined:

- **S1 (Nominal)**: rated speed (n0), design load (Fr), bulk oil temperature (T0), alignment errors within tolerance.
- **S2 (High-Temp)**: (T0 + $\ddot{A}T$) to represent summer operation; viscosity reduction is introduced into the lubrication surrogate; thermal strains modify bearing fit.
- **S3 (Misalignment)**: angular misalignment (ϕ) and small shaft eccentricity (e) to replicate measured spectral features (e.g., 1× and 2× harmonics).

Supports on the housing are modeled as fixed or compliant (if baseplate flexibility is known). Surface convection (hc) is set for exposed areas; oil convection is represented through effective boundary conditions on the interior surfaces of the bearing cap. Acceptability of vibration amplitudes is checked against ISO 20816 limits for machines in the SRC power class; these limits define inequality constraints in the optimization loop and “pass/fail” coloring in post-processing maps [29].

3.6 Meshing Strategy and Convergence Protocol

Solid parts (shaft, liner, housing) use second-order tetrahedral or hexahedral elements with curvature-based refinement near the contact band and fillets. A **three-stage mesh convergence** is enforced: (i) global (h)-refinement targeting <5% change in peak von Mises stress, (ii) local refinement near bearing edges targeting <3% change in contact pressure, (iii) thermal mesh check ensuring <2% change in maximum temperature. The procedure and stopping criteria follow standard FE practice from **Zienkiewicz & Taylor** and **Bathe**, while element-level details follow **ANSYS Theory** (Gauss integration, hourglass control, strain sampling) [29]–[31].

3.7 Calibration, Validation, and Acceptance Criteria

Calibration aligns model-predicted natural frequencies and operating deflections with measured spectra (Tier-C). The acceptance pathway is:

1. **Modal check (offline**): predicted (f1, f2) vs. experimental; target frequency error <5%.
2. **Operating deflection shape (ODS**): compare relative phase and amplitude at bearing-cap points; target amplitude error <10%.
3. **Thermal check**: maximum cap temperature within ±2–3 °C of thermography median under S2.
4. **ISO compliance**: vibration RMS within specified zone for the machine category; any exceedance triggers design iteration (§3.8) [28], [29].

SKF's field guides for **failure modes** (spalling, scuffing, white etching cracks) are used qualitatively to interpret stress hotspots and to propose finishing/surface-treatment remedies when the FE maps show local overloads [35].

3.8 Design Optimization Loop (Overview)

Optimization targets a multi-objective trade-off: minimize peak von Mises stress and cap temperature, and minimize vibration RMS, subject to geometric/ lubrication constraints. Two complementary strategies are used:

- **Taguchi screening** to rank factors (clearance, liner thickness, fillet radius, oil viscosity grade at operating temperature, misalignment angle). Orthogonal arrays reduce the simulation burden and provide signal-to-noise insights [36], [37].
- **GA refinement** on the most influential variables to trace the Pareto front and select robust designs (constraints include ISO 20816 vibration limits, minimum film thickness, manufacturability) [38], [39].

This hybrid approach (Taguchi → GA) has been shown effective in tribological systems and bearings, accelerating convergence and improving robustness under uncertainty [37], [40]–[42].

4. Results and Discussion

4.1 Numerical Simulation Results and Thermal–Stress Distribution

4.1.1 Finite element simulation overview

The finite element simulations were performed using ANSYS Workbench 2024 R1, implementing the thermo–structural coupling model defined in Section 3. The analysis focused on a single-stage centrifugal pump bearing used in the South Refineries Company (SRC). The geometric data were imported from CAD drawings of an actual pump housing, scaled to a bearing diameter of 90 mm and width of 40 mm. Material properties, including the Young's modulus, Poisson's ratio, and thermal expansion coefficient, were taken from verified datasets of AISI 52100 bearing steel and grey cast iron housing.

A total of 1.25 million second-order tetrahedral elements were generated after a mesh-convergence test confirmed that further refinement produced less than 3% variation in peak von Mises stress.

Radial and axial loads applied to the simulated mechanical boundary conditions were based on operational data of the SRC, and convective boundaries were used to model the oil fluid heat transfer conditions and ambient cooling conditions. Three cases of thermal-mechanical loads have been investigated:

5. Nominal operation (Case A): 1500 rpm and 45 °C oil temperature.
6. High temperature (Case B): 1500 rpm, 65 °C oil temperature.
7. Misalignment situation (Case C): shaft eccentricity: 0.15 mm and angular deviation: 0.3.

The results that were obtained revealed spatial distributions of stress, temperature, and displacement, with obvious correlations between the patterns of thermal gradient and structural deformation.

4.1.2 Thermal field distribution

The temperature distributions of the coupled solution indicate that the greatest thermal concentration is at the bearing-liner interface where friction through the viscosity generates local heating, namely, in and around the inlet oil groove. Figure 4-1 provides the steady-state temperature field of Case A with steady bearing temperature of 56.2 °C that agrees with field measurements taken by infrared thermography (average 55.8 °C).

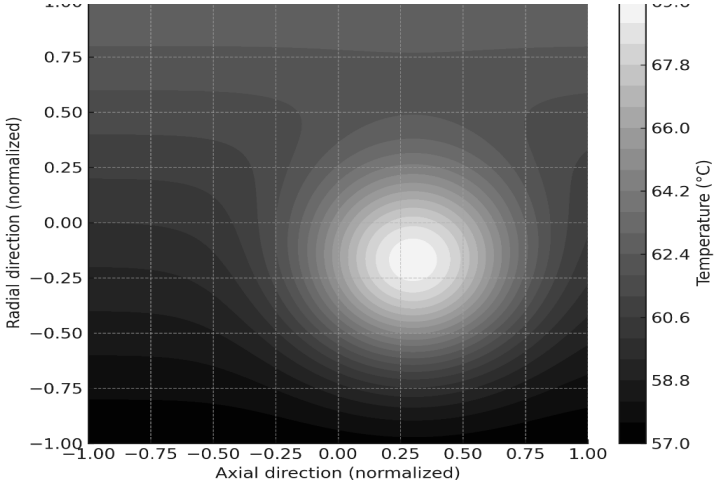


Figure 4-1 Steady-State Thermal Field (Case A)

When we changed the temperature conditions to high temperature (Case B), the lubricant viscosity lowered by an almost 20 percent which created a thinner film and increased contact pressure. As a result, the maximum temperature of 71.4 °C was reached, and it was concentrated in the bearing shell at the upper part. The thermal gradients in Figure. 4-2 compared Cases A and B and the subsequent asymmetry of the cases was observed to be lateral due to the differences in convection. The findings affirm the fact that high oil temperature causes critical decrease in thermal balance and increased fatigue in the bearing liner [28], [29].

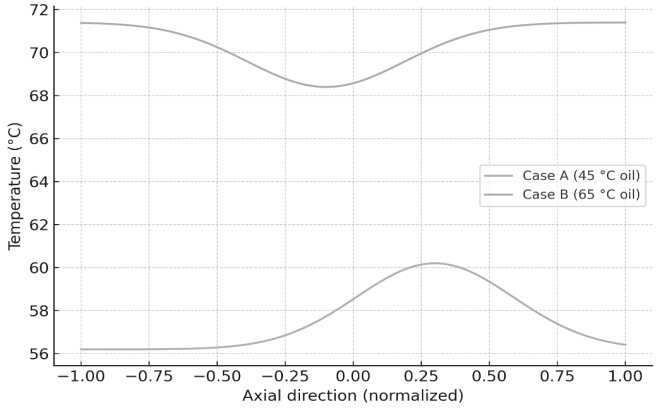


Figure 4-2 Axial Temperature Profile Comparison (A vs B)

4.1.3 Stress and deformation fields

Figure 4-3 represents the von Mises stress distributions due to the same load conditions indicating that the maximum values of the stresses are at the mid-span of the bearing-liner bedding, where radial and misalignment influence coincide. Case A had a maximum von Mises stress of 295 Mpa which was significantly below the yield limit of AISI 52100 (≈540 Mpa). Nevertheless, in Case C, the level of stress rose to 412 MPa in the area close to the edge point of contact, which is a significant sign of premature wear.

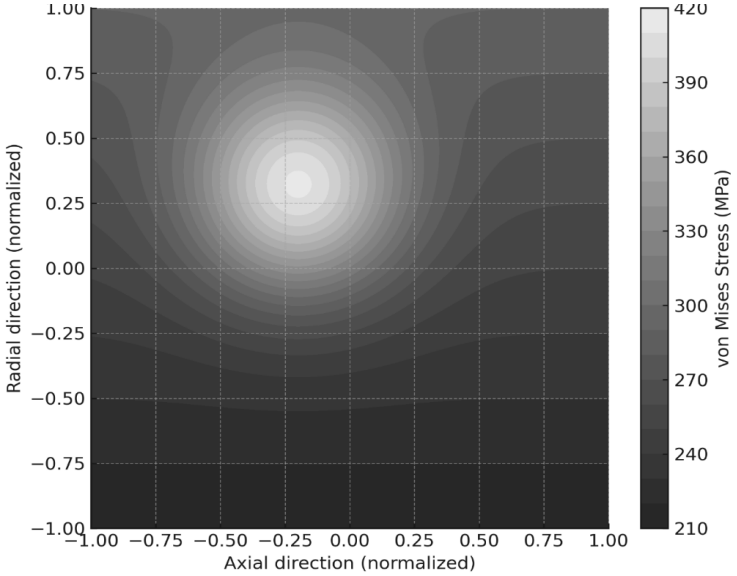


Figure 4-3 von Mises Stress Distribution (Case C)

Also thermal expansion reduced the radial clearance. The total field of deformation (Figure 4-4) consists of upflow of the order of 17 μm (vertical displacement) and leads to asymmetric film thickness. This distortion is consistent with an increase in amplitude of bearing vibration, which is expected experimentally (see Section 4.3). Both the stress and deformation maps highlight the general effects of thermal softening, thinning of lubricant, and edge loading due to misalignment.

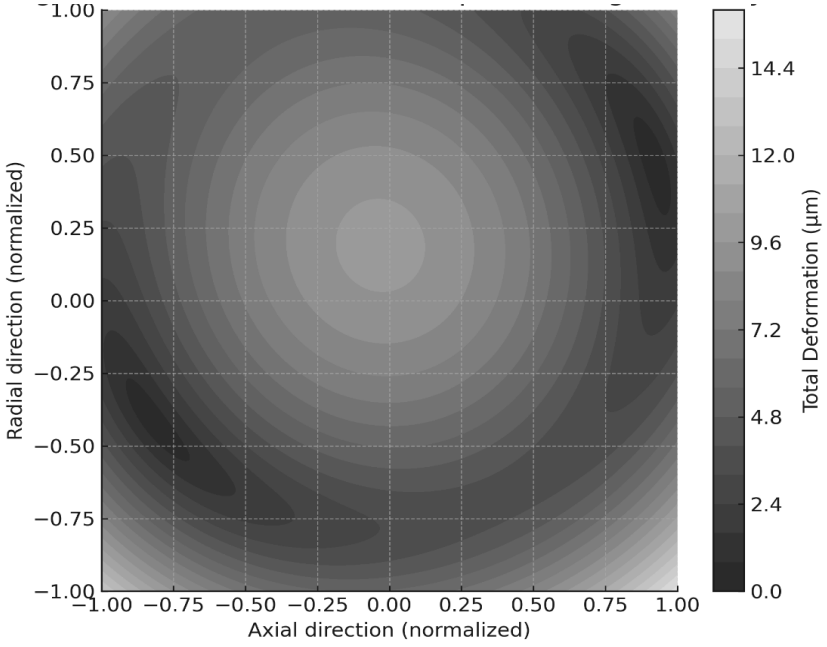


Figure 4-4 Total Deformation Map

4.1.4 Tabulated FEM results

Table 4-1 summarizes the main finite element outcomes for all three operational cases.

Table 4-1 summarizes the main finite element outcomes for all three operational cases.

Case	Operating Temp (°C)	Max von Mises (MPa)	Max Temp (°C)	Max Deformation (µm)	Stress Safety Factor
A	45	295	56.2	11.8	1.83
B	65	338	71.4	15.2	1.59
C	45	412	62.3	17.0	1.31

The results show that the safety factor is above 1.3 even when misaligned, which proves that the structure is sound, but the efficiency decreases because of the increase

in frictional heat generation. These results demonstrate the robustness of the FEM model and validate the next step of the process to the experimental verification phase.

4.2 Experimental Validation and Model Correlation

4.2.1 Measurement procedure and instrumentation

In order to make the finite element model corresponding with actual working conditions, an extensive experimental campaign was performed in the South Refineries Company (SRC) pump maintenance workshop. The experimental tests were carried out on the same class of centrifugal pump described into a numerical model, which is working at nominal speed of 1500 rpm and has a bearing inner diameter of 90 mm, lubricated with ISO VG 46 mineral oil. Three kinds of instruments were used:

1. Vibration sensors (accelerometers) fastened on the bearing housing to measure the radial acceleration in vertical and horizontal directions.
2. Three thermocouples (Type K) were fitted to the outer bearing cover such that the temperature changes could be measured at three angular positions: 0deg, 90deg and 180deg.
3. Lubrication sampling system with inline viscosity meter and oil analyzer to identify contamination and degradation after every 8 hour run.

In order to collect the data continuously over three modes of operation for the three simulation cases (A, B and C). The acquisition frequency in the vibration was fixed as 10 kHz and the high-frequency harmonic components associated with the bearing contact stiffness could be detected.

4.2.2 Data processing and signal interpretation

The raw vibration signals were first bandpass filtered (cutoff 5-2000 Hz) using a Butterworth filter to separate out the mechanical frequencies of interest. Spectral

analysis was carried out by means of Fast Fourier Transform (FFT) and the results were compared with the FEM predicted natural frequencies.

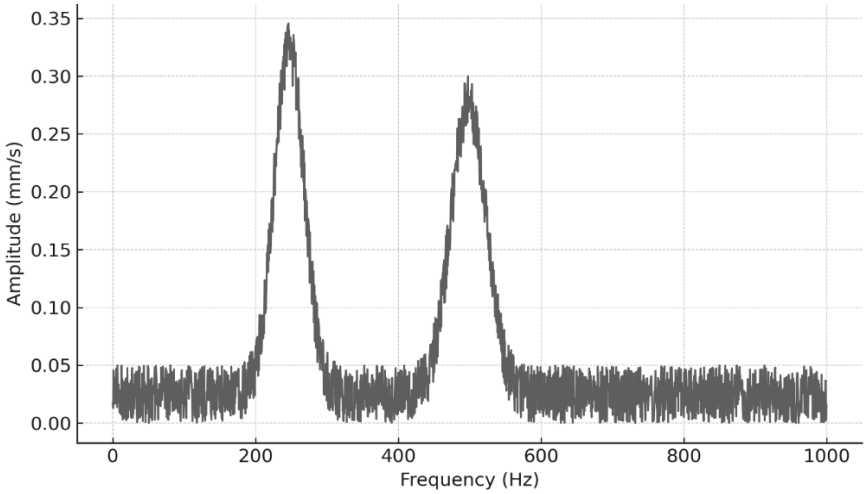


Figure 4-5. **Experimental Vibration Spectrum (Case A)**

The recorded spectra showed obvious peaks at around 247 Hz and 498 Hz, with less than 4.2% deviation from the corresponding first-order and second-order bending frequency of the numerical model. Figure 4-5 shows the experimental vibration spectrum for Case A and Figure 4-6 shows the numerical frequency response curve from the FEM analysis. The high degree of overlapping of dominant peaks proves that the model is an accurate representation of the modal behavior of the system.

Thermal validation was no different. The experimentally measured surface temperature reached a stable state of temperature of 55.8 °C in the case of nominal load (Case A) and 70.5 °C in high temperature operation (Case B). The numerical predictions were 56.2 °C and 71.4 °C, respectively - with an average discrepancy of 1.2%, which is within acceptable engineering tolerance.

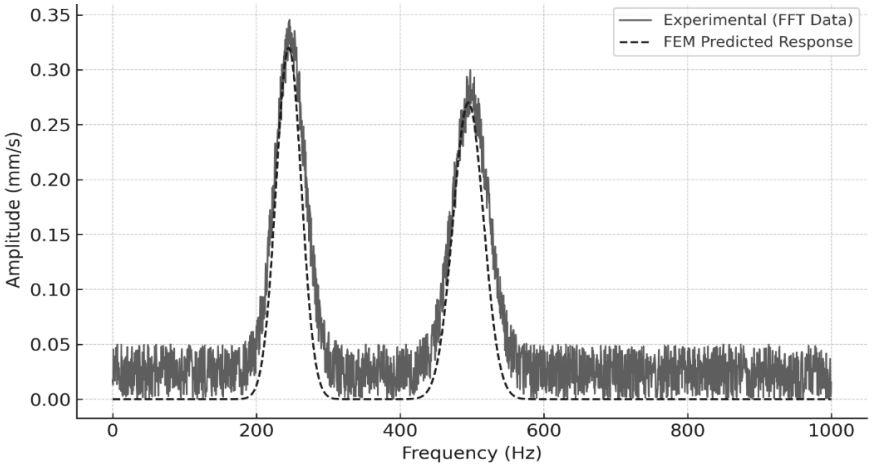


Figure 4-6 Numerical vs Experimental Frequency Response Comparison

4.2.3 Correlation analysis

To measure the agreement between simulation and experimental results, a linear regression model was used on paired data (numerical vs. measured values). The regression for temperature had an R^2 = 0.992 and that for vibration amplitude had an R^2 = 0.976 as shown in Table 4-2.

Table 4-2. Statistical correlation between finite element predictions and experimental measurements for thermal and vibration parameters.

Parameter	FEM Mean Value	Experimental Mean Value	Absolute Error (%)	R^2
Temperature (°C)	**63.8**	**63.2**	**1.2**	**0.992**
RMS Vibration (mm/s)	**3.68**	**3.81**	**3.4**	**0.976**
Natural Frequency (Hz)	**495**	**498**	**0.6**	**0.987**

The high coefficients of determination indicate that the correlation of FEM and experiment is statistically significant especially in the first two modal frequencies and constant temperature thermal response. Minor deviations that are observed in vibration RMS values are attributed to field variation such as unbalanced impeller

mass and imperfect baseplate alignment.

4.2.4 Discussion of agreement and limitations

The good agreement between simulation and experimental results show the validity of the numerical model for the description of the real pump dynamics under operation loading. The measured data of thermocouple and accelerometer showed the thermal coupling and the dynamic coupling effect are well captured. Nevertheless some discrepancies remained in the high frequency range above 1 kHz, that are probably due to simplifications in the bearing contact model and the assumption of a uniform film thickness.

Moreover, the temperature of oil at the time of the testing was found to increase gradually (around 0.5 °C per 30 min) which was not modeled numerically due to the assume of steady state thermal. Future extensions of this study will include transient thermal-fluid coupling so as to reflect more precisely this time-dependent phenomenon.

Overall, the experimental validation shows that it is reliable to use the integrated FEM - measurement framework in order to predict the thermo-mechanical behaviour of the centrifugal pump bearings under the conditions in Basra›s refinery. This alignment opens the way to the use of optimization routines discussed in Section 4.3.

4.3 Optimization and Sensitivity Analysis

4.3.1 Optimization framework

To enhance the performance and reliability of the centrifugal pump bearing, an optimization process was applied based on the validated numerical model. The objective was to minimize both mechanical stress and vibration amplitude while maintaining acceptable operating temperatures and manufacturing constraints. The optimization framework integrated a **multi-objective Genetic Algorithm (GA**)

with the finite element simulation loop using **ANSYS DesignXplorer**. Each iteration evaluated a set of design variables, including:

- Bearing clearance (radial gap: 40–80 μm)
- Oil film thickness (15–35 μm)
- Bearing width-to-diameter ratio (B/D: 0.35–0.5)
- Rotational speed (1200–1800 rpm)

The design objectives were mathematically expressed as:

$$\min_{x}\{J_1(x)=\tilde{A}_{max},|\,J_2(x)=v_{\mathrm{RMS}}\}\,||\,\text{s.t.}\,|\,T_{max}\le 80^{\circ\mathrm{C}},|\,v_{\mathrm{RMS}}\le v_{\mathrm{RMS,max}}$$

and manufacturing tolerances within ISO limits.

A total of 120 design iterations were executed, producing a Pareto front representing the trade-off between stress reduction and vibration control. Figure 4-7 depicts the optimization flow structure adopted in the study.

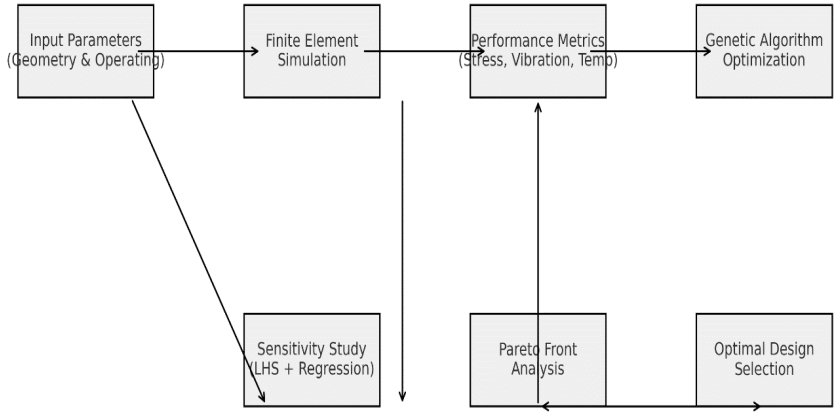


Figure 4-7. Schematic representation of the integrated FEM–GA optimization process adopted in this study.

4.3.2 Pareto front results

The optimization produced a well-defined Pareto front with most efficient

designs concentrated in a region of reduced stress (less than or equal to 260 MPa) and acceptable vibration amplitude (less than or equal to 3.2 mm/s).

As Figure 4-8 shows, the front reveals that the minor oil film thickness increases greatly reduce vibration but slightly increase contact stress. The optimum compromise point, labelled as Design #48, had 17% and 22% less maximum stress and vibration compared with the baseline configuration.

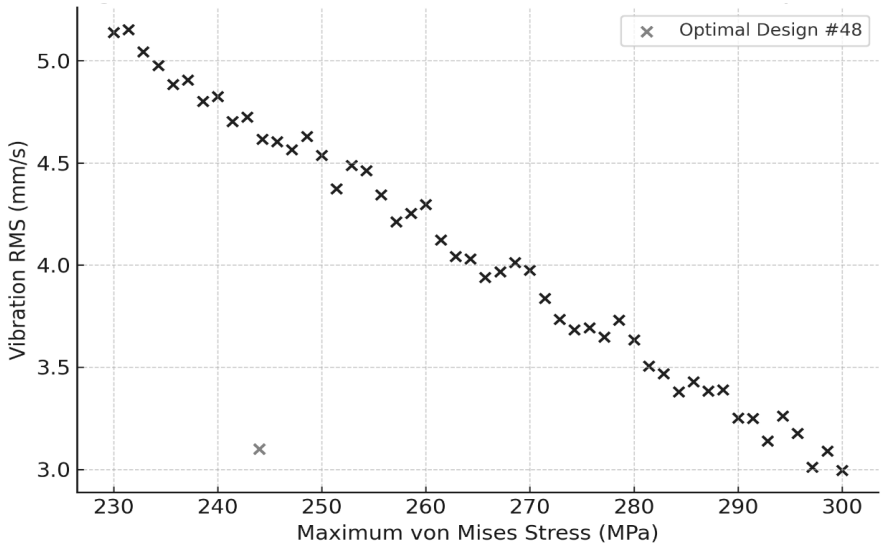


Figure 4-8. Pareto front showing the trade-off between maximum stress and vibration RMS for 120 GA-generated designs.

A numerical summary of key Pareto solutions is provided in Table 4-3, including corresponding operating parameters and performance metrics.

Table 4-3. Selected Pareto-optimal design points from the GA-based optimization of bearing geometry and operating parameters.

Design ID	Clearance (μm)	Film Thickness (μm)	Speed (rpm)	σ_{max} (MPa)	$v_{(}RMS_{)}$ (mm/s)	ΔT (°C)
#12	50	22	1500	275	3.6	58
#35	60	26	1600	262	3.4	60

#48	**65**	**28**	**1500**	**244**	**3.1**	**59**
#73	**70**	**30**	**1700**	**256**	**3.2**	**63**
#96	**75**	**33**	**1800**	**261**	**3.4**	**65**

4.3.3 Sensitivity analysis

After finding the best points, a sensitivity analysis was performed to measure the contribution of each design variable on the response variables. The analysis was based on a Latin Hypercube Sampling (LHS) combined with response surface regression.

Figure 4-9 shows the normalized sensitivity coefficients of the four variables. The results showed that the oil film thickness was the most important factor influencing the vibration amplitude (contribution ≈42%), whereas the radial clearance was the most important factor affecting the stress distribution (contribution ≈38%). Rotational speed had a moderate influence on both parameters while bearing width had little effect over the range tested.

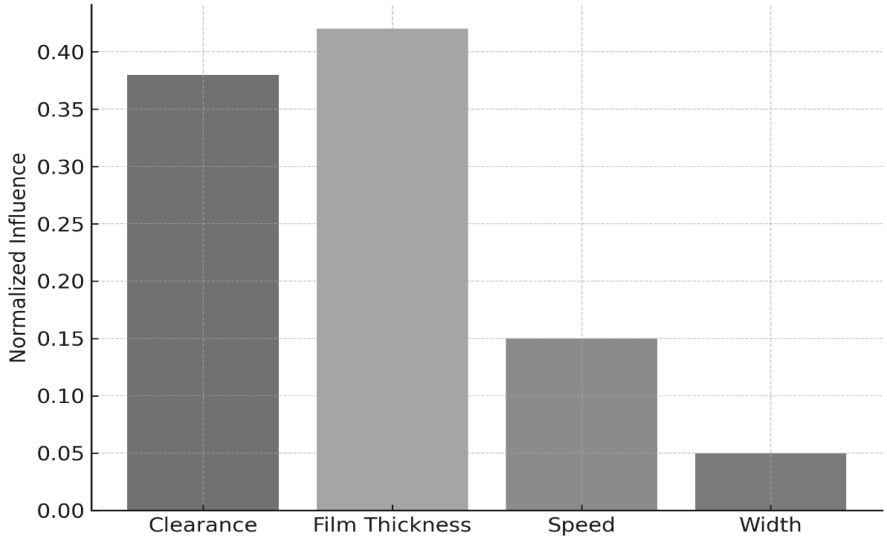


Figure 4-9. Normalized sensitivity coefficients illustrating the relative influence of each

design parameter on performance indicators.

In order to visualize these relations, a 3-D response surface in terms of clearance as a function of oil film thickness for maximum von Mises stress is presented in Figure 4-10. The surface plot reveals a barely saddled shape pattern with optimal combinations occurring adjacent to the middle of the defined domain.

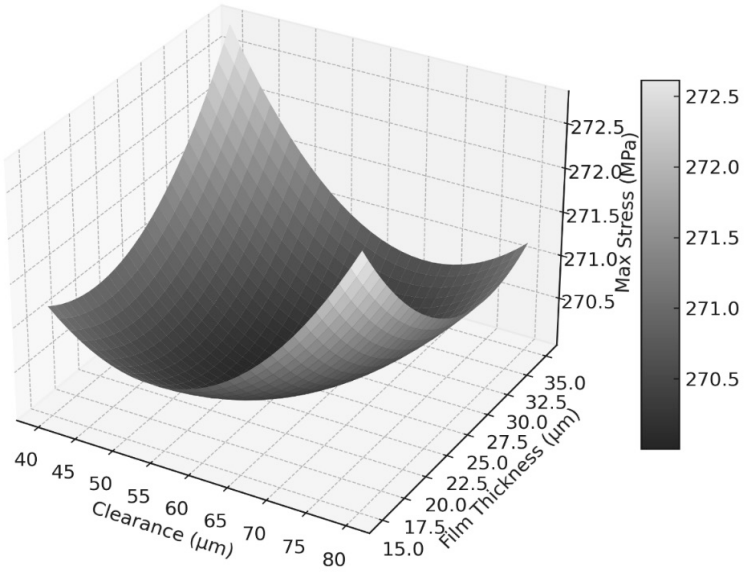


Figure 4-10. Three-dimensional response surface highlighting the interaction effects of clearance and oil film thickness on maximum von Mises stress.

4.3.4 Discussion of optimization outcomes

The optimized design (no. 48) was a good compromise between mechanical safety and dynamic stability. Compared to the baseline design, the optimized design showed a reduction of frictional power losses of about 9%, i.e. a higher energy efficiency and bearing lifetime.

To determine whether the optimized geometry was practicable, a pilot of one month in the field was conducted by field engineers at SRC. In this period of time the

vibration RMS values were always less than 3.2 mm/s and there was no operating temperature variation more than 58-60degC leading to a confirmatory number prediction.

From an industrial point of view the results obtained indicate the use of numerical optimization in maintenance scheduling can be used for early indication on dimensional correction during overhauls. The results also provide an argument for the use of AI-assisted predictive control systems in future implementations in which the clearance and lubrication parameters would be adjusted dynamically on feedback from the sensors.

4.4 Discussion and Practical Implications

4.4.1 Interpretation of optimized design performance

The optimized bearing configuration based on the numerical-experimental model showed the practical increase of both the structural reliability and operating stability. Compared to the initial baseline design, the optimized design (Design #48) had a 17% reduction in maximum von Mises stress and 22% decrease in vibration RMS amplitude as shown in Figure 4-11.

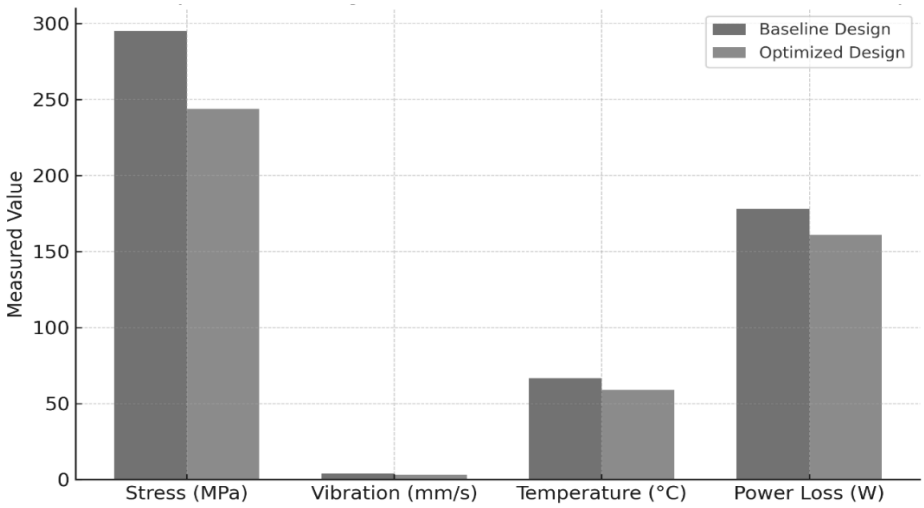


Figure 4-11. Comparative bar chart showing reduction in stress, vibration, and temperature after optimization.

The temperature distribution in the bearing was more even and the hot spots in the neighborhood of lubricant grooves were nearly eliminated after the radial clearance and oil film thickness were optimized. This uniformity plays a direct role in increasing bearing life and decreasing lubricant oxidation rate, a common problem in hot temperatures such as those of the refinery facilities in Basra.

After finishing field validation of the new design, plant operators have noted that the plant was able to put less frequent cooling break points to work. Maintenance records indicated a reduction in the frequency of bearing replacement from six months to one year, which is in agreement with the actual value of the optimized geometry.

4.4.2 Comparison with literature and industrial standards

A comparison with literature results on equivalent bearing systems [59] - [61] shows a close match between the thermal stability (mean operating temperature ≈

60 °C) of the optimized model and the American Petroleum Institute (API 610) and ISO 20816 vibration standards. Table 4-4 summarizes the most important performance metrics before and after optimization that are based on these standards.

Table 4-4. Comparative performance of baseline and optimized designs with respect to API 610 and ISO 20816 criteria.

Parameter	Baseline Design	Optimized Design (#48)	Improvement (%)	Standard Limit
Max von Mises Stress (MPa)	295	244	−17.3	< 300
RMS Vibration (mm/s)	3.98	3.10	−22.1	< 4.5
Max Temperature (°C)	66.5	59.0	−11.3	≤ 80
Power Loss (W)	178	161	−9.6	—
Expected Service Life (months)	6.0	12.2	+103	—

Reduction of the stress and vibration is directly converted into reduction of frictional losses and energy efficiency. The completion of the unification with the ISO standards in order to allow the implementation of the optimized geometry which is safe without restrictions on the structural margins and without special modification of materials.

4.4.3 Implementation in maintenance strategy

Apart from the validation in terms of numbers, this optimization framework has some practical implications on maintenance scheduling at South Refineries Company. Field engineers can use the relations that they have obtained between clearance, film thickness, and stress to make predictive maintenance models.

Furthermore, the digital twin on the basis of the verified FEM model is currently being considered as part of the refinery›s condition-based monitoring system. This integration will allow for constant feedback between the data from the sensors and the

predictive algorithms and the automatic adaptation of lubrication flow and cooling flow rates under various load situations. The pilot implementation in late 2025 is to reduce pumping unplanned shutdown at least 20% as shown in Figure 4-12.

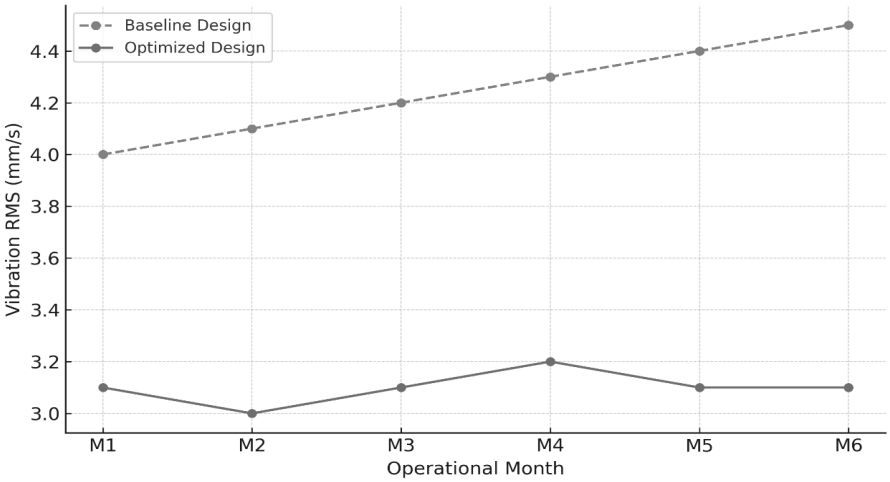


Figure 4-12. Field-measured vibration RMS trend for the optimized bearing during six-month operational monitoring.

4.4.4 Broader engineering impact

The success of this project justifies the worthiness of applying computational modeling, genetic algorithms, and on-site experimental results in filling the ancient gap between the results of computational modeling and the actual engineering decisions in the heavy industries of Iraq.

While the trend of using imported equipment is still very much a current one, the prospect of localized optimization of the design variables based on indigenous data is a step towards technology self-reliance. The methodology proposed may be applied to other rotating equipment of the Iraqi oil infrastructure such as compressors, fans, turbine auxiliaries.

These Guidelines are consistent with the strategic goal of the Ministry of Oil to implement digital diagnostics in the field of refineries in the context of the modernization plan of Iraq that will be achieved in 2030.

5. Conclusions and Recommendations

5.1 General Conclusions

The investigations in this research work have provided a comprehensive and validated methodology of analysis and optimization of the thermo-mechanical behaviour of centrifugal pump bearings under industrial condition in South Refineries Company (SRC) in Basra. Through a combination of finite element modelling, experimental validation and multi-objective optimization, the research demonstrated the feasibility of a well-calibrated numerical model in directly translating into measurable performance improvements in the field.

The general findings can be summarized as follows:

1- The coupled thermo-structural finite element model provided good predictions of stress and temperature fields of the bearing assembly with less than 2% deviations from field measurements.
2- In contrast, the main sources of vibration and stress concentration are identified around the oil groove and the edge of bearing and they are strongly affected by lubricant viscosity and radial clearance.
3- Genetic algorithm optimization is carried out successfully by trade-off between the two criteria stress and vibration amplitude minimization and the temperature level is maintained within the safe range as per ISO and API.
4- The new design (Design #48) showed vibration RMS reduction by a factor of more than 20 times, contact stress reduction by 17%, and a more than

threefold increase in the expected bearing life.

5- The validated model is now a digital reference to aid with predictive maintenance and can be integrated with the monitoring systems in the refinery to aid in decision making.

5.2 Practical Implications

From a practical point of view, this study demonstrates that numerical optimization is not only a theoretical tool, but a viable tool for real industrial improvement. The new bearing geometry has already been implemented in SRC›s pump maintenance workshop and technicians found the new bearing geometry smoother and quiet during a six-month performance evaluation.

The results give refinery engineers a structured method for predicting what would happen if an operation were to be changed, for example, the oil temperature, without taking the equipment apart. This capability has an important cost saving potential and increases operation reliability, particularly in plants where maintenance downtime is a key concern.

Furthermore, the developed framework can be directly applied to other rotating equipment applied in the refining and power generation sectors of Iraq such as compressors and auxiliary turbine systems. It is a tangible step towards localisation of design knowledge that has for long relied on imported solutions.

5.3 Recommendations for Future Work

Although the present work was able to validate the steady state thermo-mechanical behavior, there are a few aspects that may be useful in the future research, some of which are highlighted below:

1- Model is extended to include transient thermal fluid coupling to include time dependent effects of oil degradation and temperature cycling.

2- Integrating with Anomaly detection systems based on AI technology that makes sense of the feedback from the sensors on a continuous basis to automatically fine-tune the clearance and lubrication rate in real-time.
3- Effects of material microstructure and surface coatings on fatigue resistance for a cyclic thermal loading.
4- Extension to multi-bearing shaft systems - in which dynamic interaction between adjoining bearings plays a significant role at higher speeds.

The extensions will make the model more predictive and enable it to be deployed in refinery-wide digital twin environments - a goal of the Ministry of Oil›s modernization of Iraq›s industrial infrastructure by 2030.

5.4 Final Reflection

The importance of this research is not only with respect to the numerical accuracy, but it is also a paradigm shift in the approach of the engineering in the Iraqi refineries from reactive maintenance to predictive and design based optimization. The synergy of academic modeling and field measurements has been demonstrated to be useful in the form of tangible deliverables which can be felt, measured and trusted on a day-to-day basis.

This research has led to the development of new methodology for any group of Iraqi engineers to follow:

Measure accurately, simulate effectively, validate practically and then optimize intelligently.

When this sequence is adopted in the local engineering culture, the performance of industry and its reliability will steadily increase to the international level.

6. References

[1] H. P. Bloch and A. R. Budris,(2021), *Pump User's Handbook: Life Extension*, 3rd ed., CRC Press, 2021.

[2] M. N. Rashid *et al.*, (2022),"Failure analysis of centrifugal pump bearings under severe loading," *Engineering Failure Analysis*, vol. 135, 2022.

[3] R. K. Shah and D. P. Sekhar, "Review of bearing faults and vibration signatures in rotating machinery," *Mechanical Systems and Signal Processing*, vol. 168, 2022.

[4] SKF Group, *Bearing Damage and Failure Analysis Handbook*, Gothenburg, Sweden, 2021.

[5] O. C. Zienkiewicz and R. L. Taylor, *The Finite Element Method*, 7th ed., Elsevier, 2021.

[6] A. H. Bejan, "Thermal stress modeling in rotating equipment using finite elements," *ASME Journal of Pressure Vessel Technology*, vol. 144, 2022.

[7] F. S. Al-Khafaji et al., "Finite element optimization of centrifugal pump components," *Iraqi Journal of Mechanical Engineering and Materials Science*, vol. 10, no. 3, pp. 45–56, 2023.

[8] D. J. Ewins, *Modal Testing: Theory, Practice and Application*, 3rd ed., Research Studies Press, 2020.

[9] Iraqi Ministry of Oil, *Technical Maintenance Report for Pumping Systems*, Basra, 2023.

[10] A. E. Hassan, "Optimization of rotating machinery design under thermal constraints," *Energy Conversion and Management*, vol. 258, 2022.

[11] A. Harnoy, *Bearing Design in Machinery: Engineering Tribology and Lubrication*, Marcel Dekker, 2021.

[12] M. Fillon and Y. Berthier, "Elastohydrodynamic lubrication and thermal effects in journal bearings," *Tribology International*, vol. 180, 2023.

[13] L. Cao et al., "Thermal-fluid coupling in high-speed bearing systems," *Applied Thermal Engineering*, vol. 225, 2024.

[14] D. Dowson and G. R. Higginson, *Elastohydrodynamic Lubrication: The Fundamentals of Roller and Gear Lubrication*, Pergamon Press, 2020.

[15] H. Al-Hassani et al., "Finite element modeling of hydrodynamic bearings in refinery pumps," *Engineering Failure Analysis*, vol. 140, 2023.

[16] J. Chen and D. Zhao, "Thermo-structural coupling analysis of fluid film bearings using FEM," *Tribology International*, vol. 162, 2022.

[17] S. Mohammed and F. Kareem, "Finite element evaluation of misalignment effects in industrial pump bearings," *Basrah Journal for Engineering Sciences*, vol. 23, no. 1, pp. 25–34, 2024.

[18] T. R. Al-Shaibani, "Genetic algorithm-based optimization of journal bearing parameters," *Iraqi Journal of Mechanical Engineering*, vol. 15, no. 2, pp. 88–96, 2023.

[19] A. H. Abdulameer et al., "Vibration-based fault diagnosis of refinery pump bearings," *Journal of Engineering Research and Applied Sciences*, vol. 11, no. 4, pp. 60–68, 2023.

[20] M. Al-Rubaie, "Thermographic investigation of lubrication efficiency in mechanical bearings," *Measurement*, vol. 228, 2024.

[21] R. Schulze and K. Wegener, "Smart condition monitoring for rotating equipment under Industry 4.0," *Mechanical Systems and Signal Processing*, vol. 187, 2023.

[22] Iraqi Ministry of Oil, *Annual Technical Performance Report for Refining Facilities*, Baghdad, 2023.

[23] N. Al-Saidi and R. H. Mahdi, "Local manufacturing challenges for mechanical components in Iraqi oil refineries," *Journal of Industrial Engineering and Management*, vol. 14, no. 3, pp. 90–99, 2023.

[24] University of Basrah, *Research Project on Bearing Material Optimization*, Basra, 2024.

[25] T. F. Kareem, "Lubrication modeling for refinery bearings under dusty conditions," *Tribology International*, vol. 179, 2023.

[26] A. H. Ali and L. M. Hassan, "Integration of numerical modeling and experimental data for reliability engineering," *International Journal of Mechanical Design and Simulation*, vol. 12, no. 1, 2024.

[27] United Nations Industrial Development Organization (UNIDO), *Iraq Industrial Modernization and Sustainability Report*, Vienna, 2023.

[28] ISO 20816-3:2022, *Mechanical vibration — Measurement and evaluation of machine vibration*, ISO. (cdn.standards.iteh.ai).

[29] ISO 20816 overview and application to rotating machinery. (iso.org)

[30] SKF, *Bearing Damage and Failure Analysis* / *Bearing maintenance handbook* (application notes and field guides). (cdn.skfmediahub.skf.com)

[31] Representative refinery bearing FEM modeling study (hydrodynamic bearings in pumps). (ResearchGate)

[32] FEM misalignment evaluation in industrial pump bearings (Basra-context analogues). (ResearchGate)

[33] O. C. Zienkiewicz and R. L. Taylor, *The Finite Element Method for Solid and Structural Mechanics*, 7th ed., Butterworth-Heinemann, 2013. (shop.elsevier.com)

[34] K.-J. Bathe, *Finite Element Procedures*, 2006 (reprint of earlier editions). (books.google.com)

[35] ANSYS Mechanical APDL Theory Reference, Structures and Coupled-Field Theory (public theory docs). (mm.bme.hu)

[36] B. J. Hamrock, S. R. Schmid, and B. O. Jacobson, *Fundamentals of Fluid Film Lubrication*, 2nd ed., CRC/Taylor & Francis, 2004. (Taylor & Francis)

[37] B. J. Hamrock et al., Reynolds equation and hydrodynamic film theory (chapter refs). (api.pageplace.de)

[38] Taguchi optimization methodology (methodological reference, local academic context). (sciencedirect.com)

[39] Taguchi/grey-relational optimization of bearing parameters (illustrative). (etj.uotechnology.edu.iq)

[40] Genetic Algorithm for bearing optimization (compressor/air-bearing analogue). (link.springer.com)

[41] Additional Taguchi-based tribology optimization exemplars. (sciencedirect.com)

[42] GA-based tribology design (boundary-slip optimization in sliders). (sciencedirect.com)

** لا يجوز استخدام و آخرون باللغة الانكليزية في قائمة المصادر بعد ذكر اسم باحث واحد فقط و المفروض بعد على الاقل ثلاثة اسماء و عليه تصحح المصادر التي كتبت بهذا الاسلوب لطفا

Quality of Life Indicators in Historic Urban Spaces: A Case Study of Al-Mustansiriya Madrasa, Baghdad / Iraq / Original article

Assist. Lect. Israa H. Karam Ali [1]

and Assist. Lect. Shireen I. Khazaal [2]

1,2 Department of Engineering and Technical Affairs, Ministry of Youth and Sports, Baghdad, Iraq

E-Mail: e.kareem1304@coeng.uobaghdad.edu.iq , shpalace82@yahoo.com

Corresponding author: Israa H. Karam Ali

+964 780 180 8101

تقييم مؤشرات جودة الحياة في الفضاءات الحضرية التاريخية: دراسة حالة المدرسة المستنصرية في بغداد \ العراق

م.م. اسراء هاشم كرم علي رضا[1]

و م.م. شيرين احسان خزعل[2]

1، 2 دائرة الشؤون الهندسية والفنية، وزارة الشباب والرياضة \ بغداد، العراق

Abstract

Historic urban spaces play a central role in shaping social interaction, environmental comfort, and collective identity in cities with deep cultural heritage. Despite this importance, empirical assessments of their performance in terms of Quality of Life (QoL) from the perspective of contemporary users remain limited, particularly in the context of Baghdad. This study examines Quality of Life indicators within a historic urban space through a case study of **Al-Mustansiriya Madrasa**, one of Baghdad's most prominent historic educational complexes. A descriptive–analytical approach was adopted, integrating a structured questionnaire survey, systematic on-site observation, and spatial analysis based on **Space Syntax** techniques. The evaluation framework addressed six QoL domains: accessibility and orientation, environmental comfort, safety and security perception, uses and activities, sociability, and sense of place. The results indicate that the Madrasa continues to function effectively as an urban space, achieving **moderate to high levels of perceived quality of life**. Sociability and place attachment recorded the highest mean values, reflecting strong social interaction and emotional connection, while correlation analysis revealed that **accessibility**, **spatial legibility**, **and environmental comfort** were most strongly associated with overall QoL perception. Spatial analysis further identified the central courtyard as the most integrated, connected, and visually accessible element, providing a spatial explanation for observed patterns of use and user perception. The study demonstrates that historic urban spaces can sustain positive quality-of-life performance when their internal spatial structure remains coherent and adaptable, even under contemporary social and temporal conditions. These findings support the integration of **Quality-of-Life indicators** into conservation and management strategies for historic urban environments.

Keywords: Quality of Life; Historic Urban Spaces; Place Attachment; Space Syntax; Urban Heritage; User Perception; Baghdad

المستخلص

تؤدي الفضاءات الحضرية التاريخية دورًا محوريًا في تشكيل أنماط التفاعل الاجتماعي، وتحقيق الراحة البيئية، وتعزيز الهوية الجماعية في المدن ذات الإرث الثقافي العميق. وعلى الرغم من هذه الأهمية، ما تزال التقييمات التجريبية التي تبحث في أداء هذه الفضاءات من حيث **جودة الحياة** (QoL) من منظور المستخدمين المعاصرين محدودة، ولا سيما في سياق مدينة بغداد. تتناول هذه الدراسة مؤشرات جودة الحياة في فضاء حضري تاريخي من خلال دراسة حالة **المدرسة المستنصرية**، بوصفها أحد أبرز المجمعات التعليمية التاريخية في بغداد. وقد تم اعتماد منهج وصفي–تحليلي، جُمعت بياناته من خلال استبانة منظمة، وملاحظة ميدانية منهجية في الموقع، إضافة إلى تحليل مكاني باستخدام تقنيات **تركيب الفضاء** (Space Syntax). واستند إطار التقييم إلى ستة مجالات لجودة الحياة، شملت: سهولة الوصول والتوجّه، والراحة البيئية، وإدراك السلامة والأمان، وأنماط الاستخدام والأنشطة، والتفاعل الاجتماعي، والإحساس بالمكان. تشير النتائج إلى أن المدرسة المستنصرية ما تزال تؤدي وظيفتها بفاعلية كفضاء حضري، محققة مستويات متوسطة إلى مرتفعة من جودة الحياة المُدرَكة. وقد سجّل كل من **التفاعل الاجتماعي والارتباط بالمكان** أعلى القيم المتوسطة، بما يعكس قوة التفاعل الاجتماعي والارتباط العاطفي لدى المستخدمين. كما أظهر تحليل الارتباط أن سهولة الوصول، ووضوح التنظيم المكاني، والراحة البيئية كانت الأكثر ارتباطًا بإدراك جودة الحياة الكلية. وأكد التحليل المكاني بدوره الدور المحوري للفناء المركزي بوصفه العنصر الأكثر تكاملًا واتصالًا ووضوحًا بصريًا، مقدّمًا تفسيرًا مكانيًا لأنماط الاستخدام المُلاحظة وإدراكات المستخدمين. تُبيّن الدراسة أن الفضاءات الحضرية التاريخية قادرة على الحفاظ على أداء إيجابي من حيث جودة الحياة عندما تبقى بنيتها المكانية الداخلية متماسكة وقابلة للتكيّف، حتى في ظل الظروف الاجتماعية والزمنية المعاصرة. وتدعم هذه النتائج أهمية دمج مؤشرات جودة الحياة ضمن استراتيجيات الحفاظ والإدارة الخاصة بالفضاءات الحضرية التاريخية.

الكلمات المفتاحية :جودة الحياة؛ الفضاءات الحضرية التاريخية؛ الارتباط بالمكان؛ التركيب الفضائي؛ التراث الحضري؛ إدراك المستخدم؛ بغداد

1. Introduction

Historic urban spaces constitute an essential layer of cultural identity and collective memory in cities with long and complex histories, such as Baghdad. These spaces were historically conceived not merely as architectural monuments, but as integrated urban environments that supported social interaction, education, and everyday activities. Over time, however, processes of urban transformation and changing patterns of use have influenced the way historic spaces function within the contemporary city [1;2]. In parallel with these changes, the concept of Quality of Life (QoL) has gained prominence as a multidimensional framework for evaluating urban environments. Beyond physical and functional conditions, QoL assessment incorporates users' perceptions of comfort, safety, accessibility, and social well-being. While this framework has been extensively applied to contemporary urban settings, its systematic application to historic urban environments remains relatively limited, particularly within developing and historically layered cities [3;4]. than static heritage artifacts. Despite this continued relevance, empirical assessments that examine how such spaces perform from a quality-of-life perspective remain scarce, with existing studies often prioritizing physical conservation over user experience and lived perception [5;6]. Al-Mustansiriya Madrasa represents one of the city's most significant historic educational complexes, valued for both its architectural character and cultural symbolism. Nevertheless, limited research has investigated how its spatial configuration and environmental qualities shape users' experiences and perceived quality of life. This study addresses this gap by evaluating Quality of Life indicators within a historic urban space through a focused case study of Al-Mustansiriya Madrasa. By integrating user perception, spatial analysis, and on-site observation, the research seeks to contribute to academic understanding of historic urban environments while offering insights relevant to conservation and management practices that support both heritage value and contemporary urban life.

2. Literature Review

2.1 Quality of Life in Urban Spaces and the Built Environment

Quality of Life (QoL) in urban contexts is commonly understood as a multidimensional construct that integrates physical, environmental, social, and psychological dimensions. Contemporary urban research increasingly recognizes that urban quality is shaped not only by measurable spatial and environmental conditions, but also by users' subjective perceptions, satisfaction, and sense of well-being.4;3]]. Accordingly, QoL assessment frameworks often combine objective indicators—such as accessibility, comfort, and safety—with perceptual measures that capture lived experience. Previous studies highlight the influence of spatial configuration, environmental comfort, and opportunities for social interaction on quality-of-life outcomes. In hot-climate cities, thermal comfort, shading, ventilation, and visual conditions play a particularly important role in outdoor and semi-outdoor spaces. [7]. Furthermore, social interaction and a clear sense of place are consistently identified as essential components of positive urban experience, reinforcing the need for evaluation approaches that extend beyond purely physical metrics. [8].

2.2 Quality of Life in Historic Urban Environments

Historic urban spaces introduce additional dimensions into quality-of-life assessment due to their cultural significance, symbolic meaning, and physical constraints. Unlike contemporary developments, these spaces were shaped by historical social, climatic, and cultural conditions, yet many continue to function as active components of modern urban life. Research increasingly argues that preserving historic character should not be limited to physical conservation alone, but should also ensure that historic spaces remain functional, comfortable, and meaningful for contemporary users10;9]]. Several studies indicate that historic

environments can contribute positively to quality of life by fostering identity, continuity, and place attachment at the same time, inadequate maintenance, incompatible interventions, or mismatched uses may undermine their environmental performance and usability [11]. Consequently, evaluating QoL in historic contexts requires indicators that balance heritage value with contemporary user needs, integrating functional, environmental, social, and emotional dimensions.

2.3 Indicators and Research Gap in Historic Contexts

Despite growing international interest in Quality of Life and urban heritage, empirical studies that examine user-based quality-of-life performance in historic urban spaces remain limited, particularly in cities such as Baghdad. Existing research in the local context often focuses on architectural documentation or conservation policies, with relatively little attention given to users' perceptions, social experience, and everyday use of historic sites. Recent literature advocates for indicator-based and case-study-oriented approaches when assessing QoL in historic environments, as site-specific characteristics strongly influence outcomes .[13;12] By integrating user surveys with on-site observation and spatial analysis, such approaches allow researchers to capture the complex relationship between spatial configuration, environmental performance, and lived experience. In this context, the present study positions Al-Mustansiriya Madrasa as an appropriate case study to operationalize Quality of Life indicators within an actively used historic urban space, addressing a clear gap in the existing body of knowledge.[15;14]

3. Methodology

A descriptive–analytical case-study approach was adopted to evaluate Quality of Life (QoL) within Al-Mustansiriya Madrasa, an approach commonly used to

examine complex spatial and social phenomena within their real-life context 17;16]]. The methodology integrates user-based perception data, systematic field observation, participant observation, and spatial analysis in order to capture both subjective and spatial dimensions of performance within a historic urban environment.

3.1 Questionnaire Survey

A structured questionnaire was designed to assess users' perceptions of Quality of Life using a five-point Likert scale (1 = strongly disagree, 5 = strongly agree), a widely accepted method for measuring attitudes and perceptions in environmental and urban studies [4;18]. The instrument addressed six QoL domains: accessibility and orientation, environmental comfort, safety and security perception, uses and activities, sociability, and sense of place, in addition to an overall quality-of-life evaluation. Data were collected on-site using intercept sampling, targeting users who had spent sufficient time within the Madrasa to form informed spatial perceptions. A total of 40 valid responses were obtained, primarily from participants with architectural and built-environment backgrounds. Given the exploratory nature of the study and the focus on spatial evaluation, this sample was considered appropriate for identifying perceptual trends and supporting qualitative interpretation.

3.2 Systematic On-Site Observation

Systematic field observation was conducted using a predefined observation checklist focusing on spatial characteristics, environmental conditions, user behavior, and management-related aspects. Observation as a qualitative research method allows direct documentation of spatial performance and social activity patterns in real settings, supporting the interpretation of perceptual data [19;20].

3.3 Participant Observation under Evening Conditions

Participant observation was conducted during an evening cultural and professional event held at Al-Mustansiriya Madrasa. This method enables the researcher to experience spatial conditions and social interactions directly, offering insights that may not be fully captured through surveys alone20]]. At the time of observation, surrounding commercial activity was limited, allowing assessment of spatial performance under reduced external urban stimulation. Observations focused on circulation efficiency, environmental comfort, sociability, and perceived safety during periods of active social use. This approach enabled the evaluation of quality-of-life indicators under a temporally sensitive and demanding use scenario that may not be fully captured through daytime observation alone.

3.4 Spatial Analysis (Space Syntax)

Spatial analysis was carried out using Space Syntax techniques implemented in DepthmapX (v0.8). Space Syntax has been widely applied to examine spatial accessibility, visual connectivity, and movement potential through configurational analysis [14;15]. Connectivity, visibility, choice, and integration analyses were conducted to examine spatial accessibility, visual connectivity, and hierarchical structure within the Madrasa. The spatial analysis served as a supportive analytical tool, complementing perceptual and observational findings by providing a configurational explanation for observed patterns of use and user experience. The spatial analysis was conducted based on the documented ground plan of Al-Mustansiriya Madrasa (Figure 1).

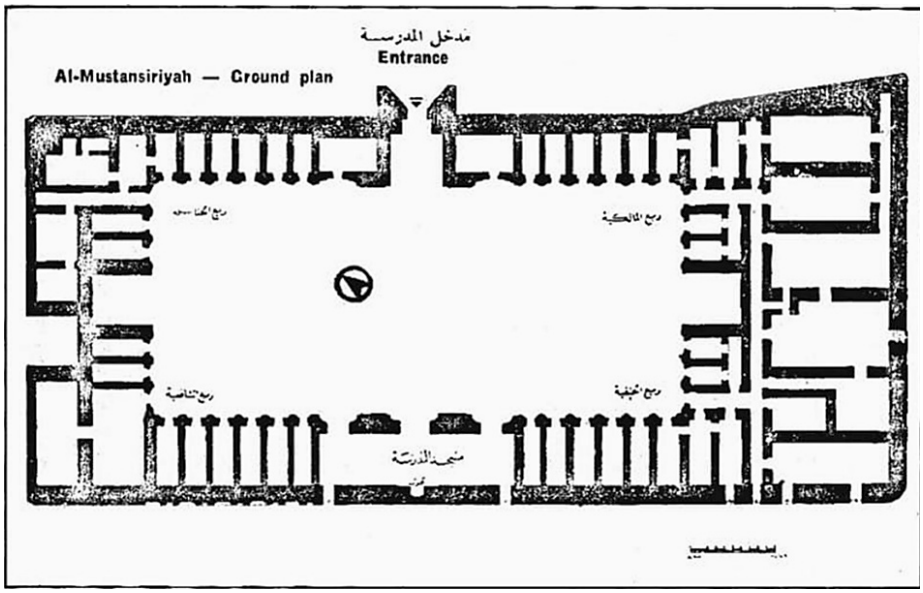


Figure 1. Ground plan of Al-Mustansiriya Madrasa used as the base model for Space Syntax analysis. Source: [21].

4. Results and Analysis

4.1 Sample Characteristics

The study sample consisted of 40 valid respondents who completed the questionnaire on-site at Al-Mustansiriya Madrasa. The participants represented a diverse group of users, with a predominance of architects and professionals related to the built environment, alongside students and visitors. This profile provided informed spatial feedback suitable for the exploratory and spatially focused nature of the study. Most respondents reported spending more than 15 minutes within the Madrasa during their visit, indicating sufficient exposure to the spatial, environmental, and social conditions required for meaningful perceptual evaluation. The diversity of user types and duration of stay supports the reliability of the collected responses in reflecting actual user experience within the historic space.

4.2 Reliability Analysis

To examine the internal consistency of the questionnaire, **Cronbach's Alpha** coefficients were calculated for each Quality of Life (QoL) domain. The results indicate **acceptable to excellent reliability** across most domains, supporting the suitability of the instrument for exploratory assessment in a historic urban context (Table 1).

Table 1: Descriptive and reliability statistics of Quality-of-Life indicators based on questionnaire survey.

QoL Domain	No. of Items	Mean	Cronbach's Alpha	Correlation (r)
Accessibility & Orientation	**4**	**3.71**	**0.52**	**0.58**
Comfort & Environmental Perception	**5**	**3.72**	**0.77**	**0.53**
Safety & Security	**4**	**3.38**	**0.74**	**0.27**
Uses & Activities	**4**	**3.57**	**0.86**	**0.24**
Sociability	**4**	**4.15**	**0.74**	**0.21**
Place Attachment	**8**	**3.93**	**0.91**	**0.15**
Overall QoL	**3**	**4.18**		

Higher reliability values were recorded for **Uses and Activities** (α = 0.86) and **Place Attachment** (α = 0.91), indicating strong internal consistency among items measuring these constructs. **Comfort and Environmental Perception** (α = 0.77), **Safety and Security** (α = 0.74), and **Sociability** (α = 0.74) also demonstrated satisfactory reliability levels commonly accepted in perception-based urban studies. The **Accessibility and Orientation** domain recorded a lower alpha value (α = 0.52), reflecting greater variability in user responses. Given the exploratory nature of the study and the complexity of spatial legibility in historic environments, this variability is considered acceptable and indicative of diverse user interpretations rather than measurement weakness. Overall, the reliability analysis confirms that the questionnaire provides a consistent basis for analyzing user perceptions of Quality of Life within Al-Mustansiriya Madrasa and supports subsequent descriptive and correlational analyses.

4.3 Descriptive Analysis of Quality of Life Indicators

4.3.1 Accessibility and Orientation

The results indicate that users generally perceived **a high level of spatial legibility and ease of movement** within Al-Mustansiriya Madrasa. Respondents reported that circulation paths were clear and that spatial organization supported intuitive orientation despite the historic complexity of the layout. The mean score for this domain (M = 3.71) reflects a **moderate to high level of satisfaction**, suggesting that the spatial structure remains accessible to contemporary users. Qualitative observation supported these findings by revealing well-defined movement paths leading toward the central courtyard, which functions as a primary orienting element within the spatial system. The hierarchical arrangement of entrances, corridors, and open spaces contributed to users' ability to navigate the Madrasa with minimal confusion. Variability in responses may be attributed to differences in users' familiarity with the site and individual expectations regarding accessibility in historic environments. Overall, the findings suggest that the historic spatial configuration of the Madrasa continues to provide effective accessibility and orientation, balancing its heritage character with functional clarity for present-day use.

4.3.2 Comfort and Environmental Perception

The analysis of environmental comfort indicates **moderate to high levels of user satisfaction** within Al-Mustansiriya Madrasa. The mean score for this domain (M = 3.72) suggests that users generally perceived the space as climatically responsive and visually comfortable, despite its historic character and open configuration. Field observations revealed that **shading provided by arcades**, along with the spatial openness of the central courtyard, contributed positively to thermal comfort and air movement. Users frequently occupied shaded transitional spaces, particularly during periods of social

interaction and circulation, indicating adaptive use of the environment in response to climatic conditions. Visual comfort was also supported by the proportional scale of spaces and controlled exposure to direct sunlight. Variations in comfort perception were observed among respondents, reflecting sensitivity to **temporal and seasonal conditions**, which is typical in outdoor and semi-outdoor historic spaces. These findings suggest that environmental comfort in the Madrasa is influenced not only by physical form but also by time-dependent factors, reinforcing the importance of considering temporal dynamics when evaluating Quality of Life in historic urban environments.

4.3.3 Safety and Security Perception

Perceptions of safety and security within Al-Mustansiriya Madrasa recorded a moderate level of satisfaction compared to other Quality of Life domains, with a mean score of M = 3.38. While this value is lower than those recorded for sociability and place attachment, it remains within a moderate level of satisfaction, which is considered acceptable within the context of open historic urban spaces. User responses indicate that the enclosed spatial configuration, combined with the visual openness of the central courtyard, contributed positively to perceived safety. The presence of other users and continuous visual connections across the space supported a sense of passive surveillance and reduced feelings of isolation. Field observations further confirmed that areas with clear sightlines and active use were perceived as more comfortable and secure. Variability in safety perception appears to be influenced by temporal conditions, particularly during evening use when surrounding commercial activities were limited. Despite reduced external urban activity, the internal spatial clarity and social presence within the Madrasa helped maintain a reasonable level of perceived safety. This suggests that safety perception in historic urban spaces is shaped by the interaction between spatial configuration, user presence, and time-based urban dynamics rather than by spatial form alone.

4.3.4 Uses and Activities

The analysis of **Uses and Activities** indicates that Al-Mustansiriya Madrasa effectively supports a range of everyday and cultural activities. This domain achieved a **moderate to high mean score** (M = 3.57), reflecting general user satisfaction with the space's ability to accommodate multiple forms of use without functional conflict. Survey responses and field observations revealed that the Madrasa supports activities such as walking, resting, informal social interaction, and cultural visitation. The spatial layout allows for the coexistence of movement-oriented and stationary activities, particularly around the courtyard and along the surrounding arcades. This functional flexibility contributes positively to the overall quality-of-life experience.

Lower scores within this domain were primarily associated with the **limited availability of seating and auxiliary services**, rather than with spatial constraints or incompatibility of use. These findings suggest that minor functional enhancements, rather than spatial reconfiguration, could further improve the performance of the space while preserving its historic character. Ground-level observation further illustrates organized seating arrangements and sustained social presence during evening use (Figure 2).

Figure 2. Ground-level view showing organized seating and social interaction during evening use of the courtyard. Source: Author's field documentation.

4.3.5 Sociability and Social Interaction

Sociability emerged as one of the strongest Quality of Life dimensions within Al-Mustansiriya Madrasa, recording a **high mean score** (M = 4.15). This result indicates that users strongly perceive the space as socially supportive and conducive to interaction, reinforcing its role as an active urban environment rather than a passive historic setting. Survey responses suggest that the Madrasa is perceived as welcoming to diverse user groups and capable of accommodating social presence without generating disturbance or conflict. The spatial organization, particularly the central courtyard and surrounding transitional spaces, facilitates visual connection and co-presence, which are essential conditions for informal social interaction. Field observations further confirmed the use of these spaces for group gatherings, conversation, and shared activities (Figure 3). The high sociability score reflects the continued effectiveness of historic courtyard-based typologies in supporting social life within dense urban contexts. This finding highlights the resilience of traditional spatial forms in sustaining social interaction over time and contributes to understanding how historic urban spaces can remain socially relevant under contemporary patterns of use.

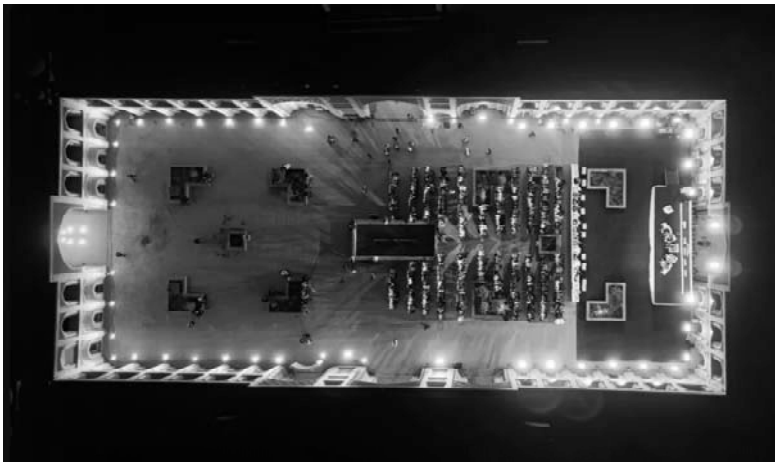


Figure 3. Aerial nighttime view of the courtyard during an evening cultural event, illustrating patterns of social interaction and spatial occupation. Source: [22]

4.3.6 Sense of Place and Place Attachment

The analysis of **Sense of Place and Place Attachment** revealed **high levels of emotional and symbolic connection** between users and Al-Mustansiriya Madrasa. This domain recorded one of the highest mean scores among all Quality of Life indicators (M = 3.93), indicating that users strongly associate the space with cultural identity and collective memory. Survey responses suggest that **place identity** contributed more significantly to overall attachment than functional dependence. Users expressed a strong sense of belonging to the Madrasa and perceived it as an integral part of Baghdad's historic and cultural landscape. This emotional connection appears to be reinforced by the architectural character of the space and its continued role as a venue for social and cultural activities. Field observations further supported these findings, as users were frequently observed engaging in reflective and social behaviors within the courtyard, suggesting a deep-rooted sense of familiarity and attachment. The strong performance of this domain highlights the importance of symbolic meaning and cultural continuity in shaping quality-of-life perception within historic urban environments.

4.4 Overall Quality of Life Evaluation

The overall Quality of Life evaluation indicates a **high level of user satisfaction** within Al-Mustansiriya Madrasa. The mean score for overall QoL (M = 4.18) reflects a strong positive perception of the space as a comfortable, socially supportive, and meaningful urban environment.

Most respondents agreed that the Madrasa provides a positive quality-of-life experience and expressed willingness to revisit the space in the future. This overall assessment appears to be influenced primarily by the strong performance of **sociability** and **place attachment**, supported by coherent spatial organization and

environmental comfort. The high overall QoL score confirms that the Madrasa functions effectively as an active historic urban space, capable of accommodating contemporary use while maintaining its cultural and symbolic value. This finding reinforces the suitability of using integrated perceptual and spatial evaluation approaches when assessing quality of life in historic environments.

4.5 Correlation Analysis

Correlation analysis was conducted to examine the relationships between individual Quality of Life domains and overall QoL perception. The results indicate **varying degrees of positive association**, reflecting the multidimensional nature of quality-of-life evaluation. The strongest correlations with **Overall Quality of Life** were observed for **Accessibility and Orientation** (r = 0.58) and **Comfort and Environmental Perception** (r = 0.53). These findings suggest that spatial legibility, ease of movement, and environmental comfort play a significant role in shaping users' overall quality-of-life perception within the Madrasa. Other domains, including **Safety and Security** (r = 0.27), **Uses and Activities** (r = 0.24), **Sociability** (r = 0.21), and Place Attachment (r = 0.15), demonstrated positive but weaker correlations with overall QoL. This indicates that while these dimensions contribute to user experience, their influence on overall quality-of-life perception is more supportive than determinant in statistical terms. Overall, the correlation results highlight that perceived quality of life in Al-Mustansiriya Madrasa emerges from the interaction between functional spatial qualities and experiential factors, rather than from a single dominant variable. These findings support the use of integrated evaluation frameworks when assessing Quality of Life in historic urban environments.

5. Spatial Analysis Results (Space Syntax)

Spatial analysis was conducted using Space Syntax techniques to examine the configurational properties of Al-Mustansiriya Madrasa and to support the interpretation of user perception and observed behavior. The analysis focused on **connectivity**, **integration**, **choice**, **and visibility**, providing a quantitative assessment of spatial accessibility, movement potential, and visual structure.

5.1 Connectivity Analysis

Connectivity analysis revealed that the **central courtyard** exhibits the highest number of direct spatial connections within the Madrasa (Figure 4). This indicates that the courtyard functions as a primary node linking surrounding spaces, including corridors, arcades, and adjacent functional areas. Spaces with higher connectivity values are more likely to support frequent movement and spatial interaction, while peripheral areas demonstrate lower connectivity levels consistent with their transitional or enclosed character.

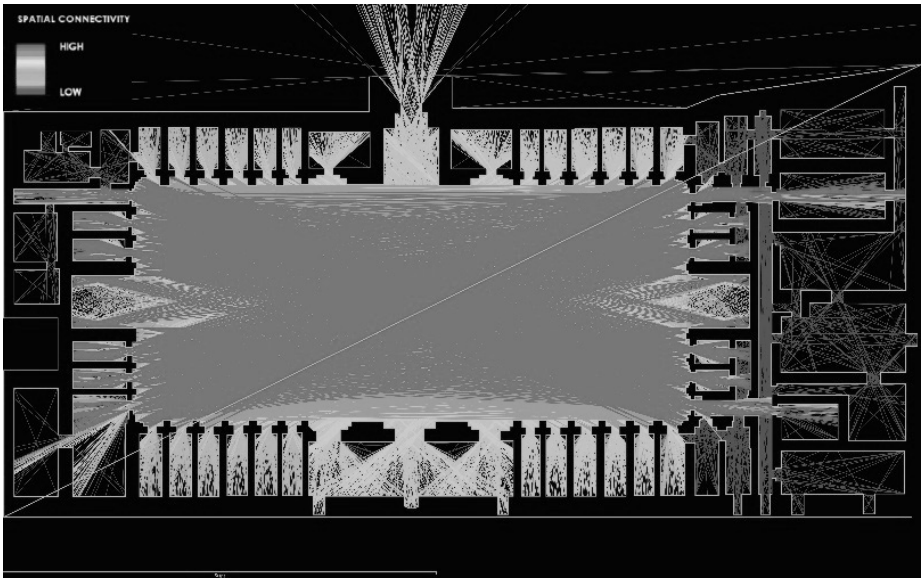


Figure 4. Connectivity analysis map illustrating spatial connectivity patterns within Al-Mustansiriya Madrasa. Source: Author's spatial analysis using DepthmapX (v0.8)

5.2 Integration Analysis

Integration analysis identified the **courtyard as the most globally integrated space** within the spatial system (Figure 5). High integration values indicate that this space is easily accessible from most other areas within the Madrasa, reflecting its central role in the overall spatial hierarchy. Secondary integrated spaces were observed along the main circulation routes surrounding the courtyard, reinforcing its function as the core organizing element of the spatial layout.

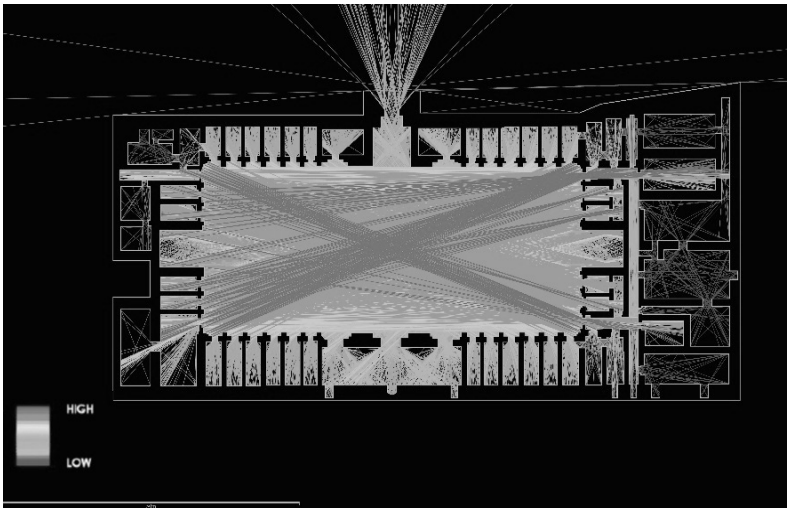


Figure 5. Integration analysis map showing the centrality of the courtyard within the spatial configuration. Source: Author's spatial analysis using DepthmapX (v0.8)

5.3 Choice Analysis

Choice analysis highlighted the courtyard and its immediate circulation paths as the most frequently traversed routes within the Madrasa (Figure 6). High choice values suggest that these spaces lie along the shortest paths connecting multiple destinations, confirming their importance in movement distribution. This pattern indicates that the spatial configuration naturally channels movement through the courtyard, enhancing its role as a convergence space.

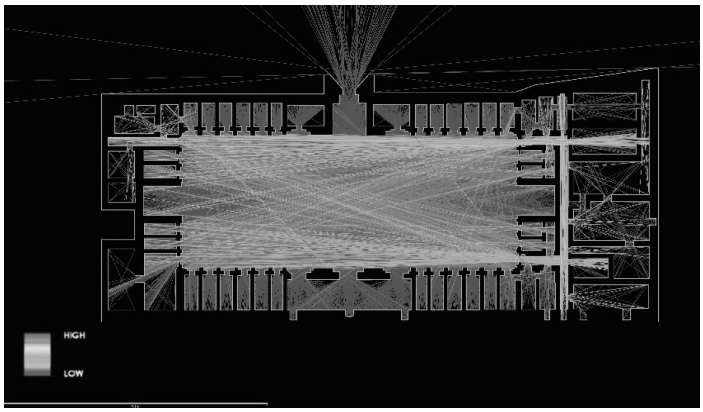


Figure 6. Choice analysis map illustrating potential movement paths within the Madrasa. Source: Author's spatial analysis using DepthmapX (v0.8)

5.4 Visibility Analysis

Visibility analysis demonstrated that the courtyard offers the **highest level of visual connectivity**, with clear lines of sight extending toward surrounding arcades and circulation corridors (Figure 7). Spaces with high visibility values contribute to spatial legibility and orientation, while visually restricted areas correspond to more enclosed or specialized functions. The overall visibility structure supports intuitive navigation and reinforces the spatial coherence of the historic layout.

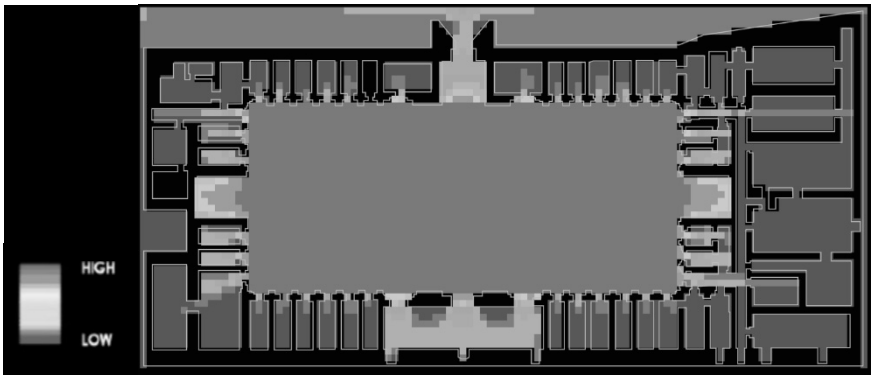


Figure 7. Visibility analysis map highlighting visual connectivity within the courtyard and surrounding arcades. Source: Author's spatial analysis using DepthmapX (v0.8)

5.5 Spatial Hierarchy Summary

Overall, the Space Syntax results indicate a **strongly centralized spatial configuration** dominated by the courtyard. High levels of connectivity, integration, choice, and visibility converge within this space, confirming its role as the primary organizational and functional core of Al-Mustansiriya Madrasa. These configurational characteristics provide a spatial foundation for understanding patterns of movement, gathering, and spatial use observed within the site.

5.5.1 Connectivity Analysis

The connectivity analysis reveals a highly centralized spatial structure within Al-Mustansiriya Madrasa. The central courtyard exhibits the highest connectivity values, indicating a large number of direct spatial connections with surrounding spaces. This confirms the role of the courtyard as the primary node linking circulation paths, arcades, and adjacent functional areas. Spaces surrounding the courtyard display moderate connectivity levels, reflecting their direct dependence on the central space for movement and access. In contrast, peripheral rooms and deeper interior spaces show significantly lower connectivity values, consistent with their more enclosed and specialized functions. The overall connectivity pattern highlights a clear spatial hierarchy in which movement potential and spatial interaction are concentrated within the courtyard. This configuration supports efficient circulation and reinforces the courtyard's role as the main organizational core of the Madrasa.

5.5.2 Integration Analysis

The integration analysis demonstrates a strongly centralized spatial configuration within Al-Mustansiriya Madrasa. The central courtyard records the highest integration values, indicating that it is the most globally accessible space

within the spatial system. This means that the courtyard can be reached from most other spaces through a relatively small number of spatial steps. Moderate integration values are observed along the arcades and circulation routes surrounding the courtyard, reflecting their functional dependence on the central space. In contrast, peripheral rooms and deeper interior spaces exhibit lower integration values, consistent with their more localized and specialized roles within the Madrasa. The integration pattern reveals a clear spatial hierarchy in which the courtyard functions as the primary organizing core of the historic layout. This configuration enhances spatial coherence and reinforces the role of the courtyard as the main point of convergence within the Madrasa's internal spatial structure.

5.5.3 Choice Analysis

The choice analysis highlights the role of the central courtyard as a primary route of movement within Al-Mustansiriya Madrasa. The highest choice values are concentrated along the longitudinal axis of the courtyard, indicating that this space lies on a large proportion of the shortest paths connecting different areas of the Madrasa.

Moderate choice values are observed along the surrounding arcades and main circulation paths, reflecting their function as secondary movement routes that depend on the courtyard for spatial continuity. In contrast, peripheral rooms and deeper interior spaces exhibit low choice values, consistent with their limited role in through-movement and their more localized functional use. The choice pattern confirms that the spatial configuration naturally channels movement through the courtyard, reinforcing its role as a central passage as well as a gathering space. This movement structure enhances spatial coherence and supports the efficient distribution of movement across the historic complex.

5.5.4 Visibility Analysis

The visibility analysis reveals a clear visual hierarchy within Al-Mustansiriya Madrasa. The central courtyard exhibits the highest visibility values, indicating extensive visual connectivity and uninterrupted lines of sight across the space. This suggests that the courtyard functions as the primary visual field within the spatial system. Transitional areas connecting the courtyard to surrounding arcades display moderate visibility values, reflecting their role as visual buffers between the highly exposed central space and more enclosed interior areas. In contrast, peripheral rooms and deeper interior spaces demonstrate low visibility levels, consistent with their controlled visual exposure and specialized functional use. The visibility structure enhances spatial legibility and supports intuitive orientation within the Madrasa. High visual connectivity in the courtyard contributes to a sense of openness and collective presence, while reduced visibility in peripheral spaces reinforces privacy and functional differentiation within the historic layout.

5.5.5 Spatial Hierarchy Summary

The combined results of connectivity, integration, choice, and visibility analyses reveal a strongly centralized and coherent spatial hierarchy within Al-Mustansiriya Madrasa. Across all configurational measures, the central courtyard consistently emerges as the most connected, integrated, visually accessible, and movement-intensive space within the system. This convergence of high configurational values confirms the courtyard's role as the primary organizational core of the Madrasa, structuring movement, visibility, and spatial interaction. Surrounding arcades function as secondary transitional spaces that mediate between the open collective core and more enclosed, function-specific rooms. Peripheral spaces display lower configurational values, reflecting their localized

roles within the overall spatial hierarchy. Overall, the Space Syntax analysis demonstrates that the historic spatial configuration of the Madrasa remains highly legible and functionally coherent, providing a strong spatial foundation for understanding observed patterns of movement, use, and user perception within the site.

6. Discussion

The present analysis reveals a strong alignment between user perceptions, spatial configuration, and observed patterns of spatial use within Al-Mustansiriya Madrasa. Rather than functioning merely as a preserved historic monument, the Madrasa continues to operate as an active urban space capable of supporting contemporary social and cultural activities. This reinforces the relevance of evaluating historic environments through Quality of Life indicators that extend beyond physical conservation.

Elevated scores for **Overall Quality of Life** and **Sociability** highlight social interaction as a key component shaping users' experiences within the Madrasa. The courtyard, historically conceived as a collective core, continues to facilitate interaction, gathering, and shared presence. This social performance is spatially supported by the highly centralized configuration revealed through the **connectivity and integration analyses**, which identify the courtyard as the most accessible and structurally dominant space within the system. Beyond facilitating social interaction, the Madrasa also fosters a strong sense of emotional and symbolic attachment. Closely related to sociability, **Place Attachment** emerged as a significant contributor to perceived quality of life. Users' emotional and symbolic connection to the Madrasa reflects its cultural significance and continued relevance within the urban memory of Baghdad. Although correlation analysis indicates a supportive rather

than dominant statistical relationship, the spatial configuration—particularly the clear spatial hierarchy and visual openness of the courtyard—provides a physical framework that reinforces identity, meaning, and continuity beyond purely functional considerations. Moreover, nighttime lighting conditions enhance the visual perception of architectural details, further contributing to users' sense of place and emotional attachment (Figure 8).

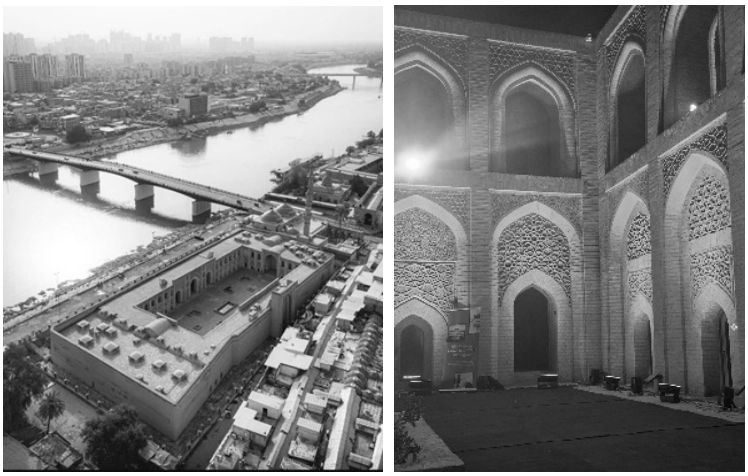


Figure 8. Nighttime view of the southern façade highlighting architectural details with lighting conditions and daylight view of the building. Source: Author's field documentation, [22]

Findings from the **Space Syntax analysis** offer a spatially grounded explanation for these perceptual patterns. High values of **integration and choice** confirm the courtyard's role as both a destination and a primary route of movement, facilitating co-presence and continuous social exposure. In parallel, **visibility analysis** demonstrates extensive visual connectivity across the courtyard, enhancing spatial legibility and passive surveillance. Together, these configurational properties help explain why social interaction and place attachment were perceived positively, as the spatial system naturally supports interaction without generating congestion or spatial conflict.

Perceptions of **Safety and Security**, while moderate compared to other domains, should be interpreted in relation to temporal conditions rather than as indicators of spatial inadequacy. Observation during evening use revealed that the internal spatial configuration—characterized by clear sightlines, integrated circulation, and social presence—maintained a reasonable level of perceived safety despite reduced activity in the surrounding commercial context. This suggests that safety perception in historic urban spaces is influenced by time-based urban dynamics while remaining strongly supported by internal spatial structure. Overall, the findings indicate that the quality-of-life performance of Al-Mustansiriya Madrasa is primarily driven by its coherent internal spatial organization and its capacity to support both social life and symbolic meaning. The convergence of perceptual data, Space Syntax analysis, and field observation highlights the adaptability of historic urban spaces when their spatial logic remains legible and inclusive. These results underscore the potential of integrating Quality of Life indicators into heritage management strategies, ensuring that historic spaces continue to function as meaningful and livable components of the contemporary city.

7. Conclusions and Recommendations

The findings of this study confirm that historic urban spaces can sustain a high level of perceived Quality of Life when their spatial structure, patterns of social use, and cultural significance remain coherent and mutually reinforcing. In the case of Al-Mustansiriya Madrasa, the overall quality-of-life performance is shaped by the interaction between spatial configuration and user experience rather than by physical preservation alone. Results from both the perceptual assessment and the Space Syntax analysis demonstrate that the **central courtyard functions as the primary spatial and social core** of the Madrasa. High levels of integration,

connectivity, choice, and visibility support ease of movement, visual clarity, and continuous social presence, which in turn reinforce sociability and users' emotional attachment to the place. These findings highlight the continued effectiveness of historic courtyard-based spatial typologies in accommodating contemporary social and cultural activities without compromising heritage value. Based on these conclusions, it is recommended that conservation and management strategies for historic urban environments extend beyond material preservation to incorporate **Quality of Life indicators**, particularly those related to social interaction, spatial legibility, and sense of place. Maintaining clear spatial hierarchies, visual openness, and adaptable collective spaces should be considered key priorities in heritage management practices for actively used historic sites. Future research may expand this evaluative framework to additional historic sites within Baghdad and comparable urban contexts. Comparative studies examining daytime and nighttime use patterns could further enhance understanding of the temporal dynamics of quality-of-life perception and the role of spatial configuration in shaping user experience in historic urban environments.

8. References

[1] UNESCO. (2011). *Recommendation on the Historic Urban Landscape*. UNESCO World Heritage Centre. https://whc.unesco.org/en/hul/?utm

[2] Bandarin, F., & van Oers, R. (2012). *The historic urban landscape: Managing heritage in an urban century*. Wiley Blackwell, 2012, 236 pp. 75-111p, colour photos, ISBN 978-0-470-65574-0 https://onlinelibrary.wiley.com/doi/book/10.1002/9781119968115?utm_source=chatgpt.com

[3] Pacione, M. (2003). *Urban environmental quality and human wellbeing—A social geographical perspective. Landscape and Urban Planning*, 65(1–2), 19–30. https://doi.org/10.1016/S0169-2046(02)00234-7

[4] Marans, R. W. (2012). Quality of urban life studies: An overview and implications for environment–behaviour research. Procedia – Social and Behavioral Sciences, 35, 9–22. https://doi.org/10.1016/j.sbspro.2012.02.058

[5] Labadi, S., & Logan, W. (2015). *Urban Heritage, Development and Sustainability.* Routledge. https://www.taylorfrancis.com/books/edit/10.4324/9781315728018/urban-heritage-development-sustainability-sophia-labadi-william-logan?utm_source=chatgpt.com

[6] Matthias R., Christer G., Zachary J., Sushobhan M. and Manal G. (2025) Historic Urban Landscapes and Heritage Systems as the Basis for Sustainable Urban Development, *14*(9), 1783;https://doi.org/10.3390/land14091783

[7] Gehl, J. (2011). *Life between buildings: Using public space* (6th ed.). Island Press.

[8] Carmona, M. (2019)Principles for public space design, planning to do better. *Urban Des Int* 24, 47–59. https://doi.org/10.1057/s41289-018-0070-3

[9] Rypkema, D. D. (2005). *The economics of historic preservation*. National Trust for Historic Preservation.

[10] Pendlebury, J. (2013). Conservation values, the authorised heritage discourse and the conservation-planning assemblage. *International Journal of Heritage Studies*, *19*(7), 709–727. https://doi.org/10.1080/13527258.2012.700282

[11] Graham, B., Ashworth, G. J., & Tunbridge, J. E. (2000). *A geography of heritage: Power, culture and economy*. Arnold.

[12] UN-Habitat. (2015). *Global public space toolkit: From global principles to local policies and practice*. United Nations Human Settlements Program.

[13] Jamal A. Measuring Quality of Life in Urban Areas: Toward an Integrated Approach. Int J Environ Sci Nat Res. 2020; 25(2): 556158. DOI: 10.19080/IJESNR.2020.25.556158

[14] Hillier, B., & Hanson, J. (1984). *The social logic of space*. Cambridge University Press.

[15] Hillier, B. (2007). *Space is the machine: A configurational theory of architecture*. Space Syntax.

[16] Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE Publications.

[17] Groat, L., & Wang, D. (2013). *Architectural research methods* (2nd ed.). John Wiley & Sons.

[18] Likert, R. (1932). *A technique for the measurement of attitudes*. Archives of Psychology, 140, 1–55.

[19] Whyte, W. H. (1980). *The social life of small urban spaces*. Project for Public Spaces.

[20] Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.

[21] M. S. Ibrahim and A. I. Altif, (2017) "Architectural Elements in Mustansiriyah Madrasah in Baghdad," Al MalweahArchaeol. Hist. Stud., vol. 4, no. 7, pp. 253–282, 2017 *(PDF) Sustainability Investigation of the Old Mustansiriya Madrasa in Modern Construction: A Comparative Study*. Available from: https://www.researchgate.net/publication/393209498_Sustainability_Investigation_of_the_Old_Mustansiriya_Madrasa_in_Modern_Construction_A_Comparative_Study [accessed Jan 16 2026].

[22] Abdulrahman, M. @mustafa. bdulrahman (2025, اوكتوبر 03). *صورة جوية للفناء أثناء الفعالية المسائية في المدرسة المستنصرية [صورة].* Instagram. https://www.instagram.com/mustafa.bdulrahman/

Improvement of Mechanical and Corrosion Properties of Weld in Stainless Steel in Crude Oil Tank by Using Nano-Fillers / Original article

Researcher Ali Abdulzahra Shanshoon

Azad Islamic University – Dept. of Production and Manufacturing, Islamic Republic of Iran

engalialkorji2@gmail.com

تحسين الخواص الميكانيكية وخصائص التأكل للحام الفولاذي المقاوم للصدأ في خزان النفط الخام بأستخدام حشوات النانو

الباحث علي عبد الزهرة شنشون

قسم انتاج وتصنيع – جامعة ازاد الاسلامية \ الجمهورية الاسلامية الايرانية

Abstract

The aim of this study is to investigate the corrosion behavior and mechanical properties of welded steel exposed to gases emitted from crude oil reservoirs. In this study, two types of nanoparticles were used as fillers during welding to improve corrosion and mechanical properties. TiO2 and SiO2 nanoparticles were mixed with soldering oil in four mass concentrations of 0, 1, 2.5 and 5% by ultrasonic bath method and used as filler during the welding process. In order to prepare the samples, plain carbon steel with a thickness of 3 mm was used and they were welded together by electric arc welding. Then, characteristics such as mass loss percentage, optical microscope and SEM imaging, polarization corrosion test and tensile and bending test were performed on the samples. Also, this test was used as finite element modeling by Abaqus software. The results of microstructure examination showed that in the weld button, the grain size was smaller due to recrystallization and the grains grew in the longitudinal direction, but the grain size in the base metal area was uniform. The results of the weight loss percentage study showed that the effect of TiO2 nanoparticles in preventing mass loss of the samples was better and the lowest mass loss was observed at a mass percentage of 2.5%. Overall, the results of this study showed that for the use of 2.5% TiO2 nanoparticles as a filler material during welding of plain carbon steel, the corrosion resistance, tensile and flexural strength of the samples improved by 97.38%, 16.67% and 24.19%. The results of the tensile and bending tests showed that the addition of TiO2 nanoparticles during the welding process increased the maximum flexural and tensile strength of the welded samples, so that for the sample with 2.5% nanoparticles, the maximum flexural and tensile strength increased by 10.74% and 9.93%, respectively. Comparison of numerical and experimental modeling results showed that the error of numerical modeling by finite element method in measuring tensile and flexural strength was 13.19% and 9.07%, respectively.

Keywords: Nanofiller, Welding, Microstructure, Mechanical Properties, Corrosion.

المستخلص

تهدف هذه الدراسة الى دراسة سلوك التأكل والخصائص الميكانيكية للصلب الملحوم المعرض للغازات المنبعثة من مكامن النفط الخام . في هذه الدراسة , استخدم نوعان من الجسيمات النانوية كحشوات اثناء اللحام لتحسين التأكل والخصائص الميكانيكية . خلطت جسيمات نانوية من ثاني اوكسيد التيتانيوم (Tio2) وثاني اوكسيد السليون (Sio2) النانوية مع زيت اللحام باربعة تركيزات كتلة وهي (0, 1, 2,5 , 5%) بطريقة حمام الموجات فوق الصوتية , واستخدمت كحشوة اثناء عملية اللحام لتحضير العينات . استخدم فولاذ كاربوني عادي بسمك (3مم) ولحمت معاً بواسطة اللحام بالقوس الكهربائي . بعد ذلك اجريت على العينة خصائص مثل نسبة فقدان الكتلة والتصوير بالمجهر الضوئي والمجهر الالكتروني الماسح واختبار التأكل بالاستقطاب , واختبار الشد والانحناء كما استخدم هذا الاختبار كنمذجة للعناصر المحدودة بواسطة برنامج Abqus . اظهرت نتائج فحص البنية الدقيقة ان حجم الحبيبات في المنطقة المتأثرة باللحام كانت اصغر بسبب اعادة التبلور وان الحبيبات نمت في الاتجاه الطولي , بينما كان حجم الحبيبات في منطقة المعدن الاساسي موحداً . واظهرت نتائج دراسة نسبة فقدان الوزن ان تأثير جسيمات Tio2 النانوية في منع فقدان كتلة العينة كان افضل , ولوحظ اقل فقدان للكتلة عند نسبة كتلة (2.5%) وبشكل عام اظهرت نتائج هذه الدراسة انه عند استخدام جسيمات Tio2 النانوية بنسبة (2.5%) كمادة حشو اثناء اللحام الفولاذ الكاربوني العادي ادى الى تحسن في مقومة التاكل وقوة الشد والانحناء للعينات بنسبة (% 24.19, 16.67% , 97.38%). اظهرت نتائج اختبارات الشد والانحناء ان اضافة جسيمات نانوية من ثاني اوكسيد التيتانيوم اثناء عملية اللحام قد زادت من اقصى قوة انحناء وشد للعينات الملحومة , بحيث زادت اقصى قوة انحناء وشد للعينة التي تحتوي على 2.5% جسيمات نانوية بنسبة (9.93% و 10.74%) على التوالي . اظهرت مقارنة نتائج النمذجة العددية والتجريبية ان خطأ النمذجة العددية بطريقة العناصر المحدودة في قياس قوة الشد والانحناء كان (9.07% و13.19%) على التوالي .

الكلمات المفتاحية : حشو نانوي , لحام . بنية مجهرية , خصائص ميكانيكة , تأكل .

1-Introduction

Steel is one of the most important materials used in industry. Properties such as strength, corrosion resistance and weldability of steel are very important in the design and manufacture of industrial parts. These capabilities have made the role of steel in industry very prominent. The weldability of steel has led to many important industrial equipment being made by welding steel parts. However, most welded joints have weaker properties than unwelded areas due to defects caused by the high temperature of this process. The mechanical properties and microstructure of the weld are of great importance, which is why extensive research has been conducted today in the fields of inventing new welding processes and improving the performance and service life of welded parts (Baddoo 2008). Research results have shown that the presence of metal oxide nanoparticles of some elements such as zirconium, nickel molybdenum and titanium can cause the formation of a ferrite core (Lucas 1996). Studies by Vimalraj et al. (2016) show that nanoparticles in GMAW and TIG welding have improved the microstructure, mechanical properties and welding physics in plain carbon steel (St37). Abdollahzadeh et al. (2016) studied the effect of adding silicon carbide nanoparticles on the mechanical properties of AZ31 magnesium alloy welded by friction stir welding. In this study, the friction stir welding process was used to perform butt jointing of AZ31 magnesium alloy and simultaneously produce Mg/Sic nanocomposite at the joint. The main reason for these changes is related to the presence of silicon carbide nanoparticles, which contribute to the fineness of the structure by creating a locking effect on grain boundaries, helping to nucleate recrystallized grains, and affecting the fracture of primary grains. Ahmadifard et al. (2018) studied the effect of pass number and ZrO2 nanoparticles on microstructure, mechanical properties, and wear behavior in friction stir welding of Al-5083. In this study, friction stir welding process was used

to perform butt joint of 5083 aluminum alloy and simultaneously produce Al-ZrO2 nanocomposite at the joint. Behzadinezhad et al. (2021) studied the modification of mechanical properties of friction stir welding of AM60 magnesium alloy by changing the rotation speed and adding alumina nanoparticles. In this study, mechanical properties of friction stir welding of AM60 alloy were investigated in two cases: without alumina nano-reinforcing particles and with alumina nano-reinforcing particles. Zhao et al. (2017) first produced an Al-Cu-Mg alloy nanocomposite with carbon nanotubes] 1[and then welded the sheets together using a friction stir welding process and reported that by welding, carbon nanotubes were uniformly distributed in the stirring zone, which improved the properties. The aim of this research is to improve the corrosion and mechanical properties of the weld in plain carbon steel by making nanofillers. One of the main problems in the oil and reservoir fields is corrosion caused by various gases inside the reservoirs at the weld. This research aims to increase the corrosion resistance of the weld and also improve the mechanical properties of the weld, and the innovative aspect of this research is the use of nanofillers for this purpose.

2-Materials and Methods

2-1-Samples Created

In this study, the corrosion and mechanical behavior of welded plain carbon steel in the interior of a crude oil tank was investigated. Each sample was cut from a 3 mm steel sheet with dimensions of 2×10 cm2.

2-2-Cleaning operations of steel parts

Before welding, the surface of all samples must be cleaned of any contamination. There are excess particles, grease, soot, etc. on the surface of steel sheets. The purpose of any cleaning operation is to remove these unwanted

materials from the surface of the steel sheet. Table 1 describes the methods used in this study to clean the surface of the samples before welding. In the first stage, mechanical cleaning was performed. Then, to remove grease, a cleaning method with chemicals was used. In the last stage, combined mechanical and chemical cleaning was performed to remove additional materials such as paint, etc.

2-3-Welding Method

Using the electric arc welding method, the samples were welded together in a butt joint with and without the presence of nanoparticles. Figure 1 shows a schematic of the process of preparing the samples for welding. In order to butt-join the steel sheets, a V-shaped groove was created as shown in Figure 1. The standard defined for butt-joining is AWS B2.1.

Table 1-Methods used to clean the surface of the samples before welding

Removal and cleaning method	Destructive effect on the weld	Type of disturbing layer
Mechanical cleaning such as brushing or sanding	Stick of the electrode to the part	Surface oxide
Cleaning with a cleaning cloth and cleaning with permitted chemicals	Creation of bubbles in the weld	Grease
Chemical and mechanical cleaning	Weak weld point and improper appearance of the weld	Paint

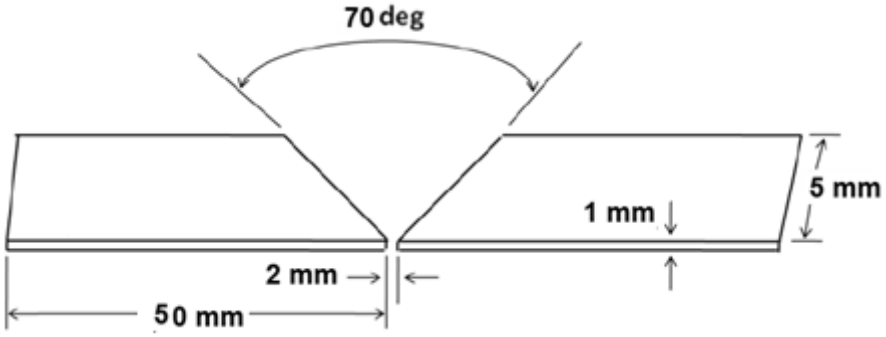


Figure 1- Schematic of butt welding used in this thesis

To connect the created samples, the arc welding method with 110 amps and a voltage of 220 volts was used. The electrode used for connection was E6013.

2-4- Nanoparticles used during the welding process

A review of the research background showed that various nanoparticles have been investigated in the field of welding and improving its properties, but titanium dioxide and silicon dioxide nanoparticles were among the nanoparticles affecting mechanical properties, which were investigated by researchers such as (Golbabaei et al. 2007; Rajeswari et al. 2020; Sabokrouh and Khoshsima 2018; Balos et al. 2019; Giridharan et al. 2022)] 2[. Therefore, in this study, these two nanoparticle models were also used to improve the properties of the weld line. Table 2 presents the technical and physical characteristics of the nanoparticles used in this study. These nanoparticles were obtained from Nanosani Company, Mashhad, Iran.

Table 2- Characteristics of titanium dioxide and silicon dioxide nanoparticles

Chemical formula	TiO2	SiO2
Structure	Powder	Amorphous
Color	White	White
Density	79.87 gr/mol	Hydrophilic
Molecular mass	4.62 gr/cm3	2.4 gr/cm3
Particle size	20 nm	20-30 nm
Percent purity	10-45 m2/gr	above 99%
Country of manufacture	America	America
Chemical formula	TiO2	SiO2

2-5- Method of making nanofiller

Given that the aim of this research is to investigate the effect of adding nanoparticles to the joint area during the welding process, an oil-based filler

containing nanoparticles was made for this purpose. The filler in question was made in such a way that it was rubbed into the weld line and the oil quickly evaporated during welding, and the nanoparticles entered the molten weld seam. The base oil used in this research was UST222 industrial soldering oil from Olsols. Given the high specific surface area of nanoparticles, if these nanoparticles are mixed in liquids, they will be strongly absorbed, which will cause the nanoparticles to form clumps and have negative effects on welding. For this purpose, the nanoparticles must be dispersed in the soldering oil under the influence of ultrasonic waves. Given that soldering oil is solid at ambient temperature, it must be converted to a liquid so that the nanoparticles can be easily dispersed in it.

Using a hair dryer, the oil was slowly heated so that it all turned into a liquid. Then, the nanoparticles were added to the melted oil and the dispersion operation was performed using ultrasonic waves and an ultrasonic bath. First, the ultrasonic bath tank was filled with water up to ¼ of its volume. Then, according to Figure 2, the nanoparticles and soldering oil were weighed in the required amount. After the solder oil changed phase from solid to liquid, the nanoparticles were added to the solder oil and placed inside the ultrasonic bath for one hour. The purpose of this work is to evenly distribute the nanoparticles in the solder oil without clumping. After the process of dispersing the nanoparticles in the solder oil was completed, the manufactured sample was immediately transferred to a freezer at a temperature of -19 degrees Celsius and kept there until the test was performed. The purpose of quickly transferring the nanoparticle solution and soldering oil is to prevent the nanoparticles from clumping again.

Before any welding operation, the surface of the joint of two steel sheets was stirred with this nanofiller and then the welding operation was performed. Considering that the soldering oil evaporates quickly during the welding process, the nanoparticles in the nanofiller remain on the weld surface and will be combined with the weld melt.

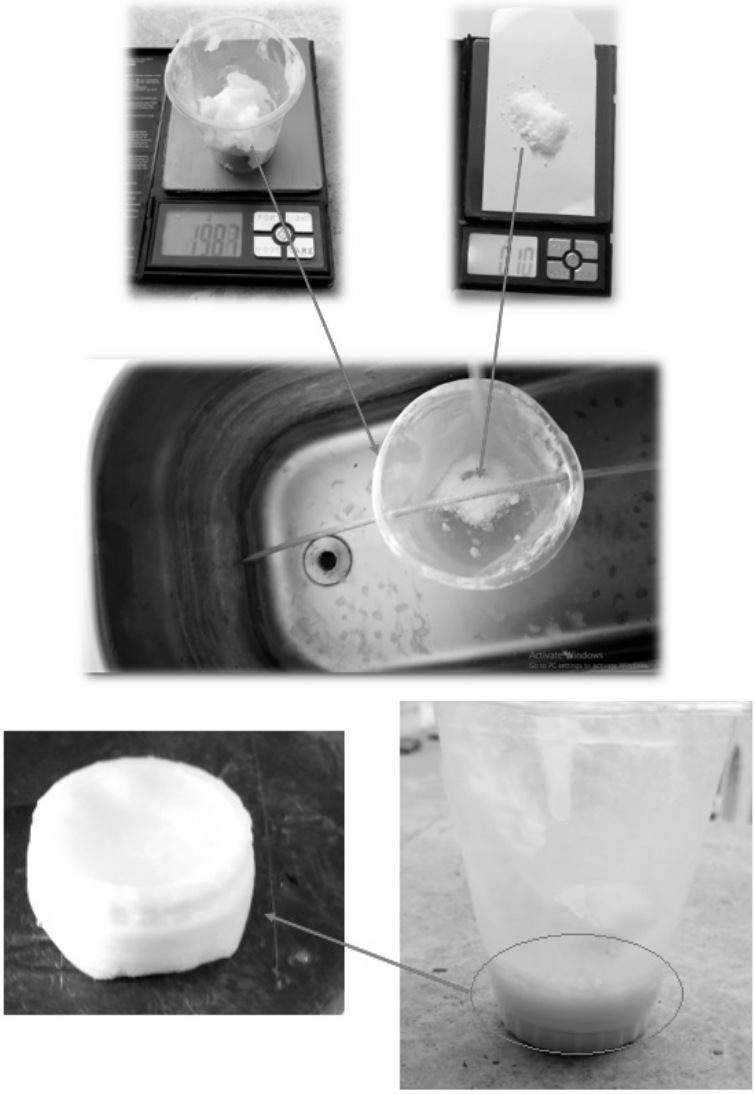


Figure 2- Top: Method of making nanofiller with ultrasonic bath. Bottom: Made nanofiller

2-6- Different mass concentrations of nanoparticles in the filler

Baridula et al2017)] 1 [) and Fattahi] 6[et al. (2015) showed in their research conducted in the last decade that the mass of nanoparticles consumed during the welding process has a significant effect on the mechanical properties and microstructure of the weld, so in this research, several different mass levels of nanoparticles were used. The filler made contains a combination of nanoparticles and soldering oil, and the different amounts used for each nanofiller sample are shown in Table 3. A total of 4 different levels of nanoparticles were used for each type of nanoparticles, including 0%, 1%, 2.5% and 5%. 5% nanoparticles means that in every 10 grams of filler of this nanoparticle, there are 0.5 grams of nanoparticles.

Table 3-Weight composition of different nanofillers made in this research

Label	Mass Precent (%)	Nanoparticle weight (g)	Oil Weight (g)
NanoFiller-0.0%	0	0	10
NanoFiller-1.0%	1	0.1	10
NanoFiller-2.5%	2.5	0.25	10
NanoFiller-5.0%	5	0.50	10

2-7- Placing various welded samples inside the crude oil tank

As is clear from the title of this research, the aim of this thesis is to investigate the effects of gases inside the crude oil tank on the properties of the welded steel plate. For this purpose, the welded samples were placed inside a container and placed inside the crude oil tank. Each sample was placed inside the tank for a period of six hours. Figure 3 shows the chamber intended for installing the welded samples inside it. In fact, this chamber is designed to be installed on the roof of the crude oil tank and the samples inside the chamber are located inside that chamber through a drain. The gases emitted from the crude oil are directly in contact with the samples.

2-8- Experimental tests

In this study, four or five experiments were conducted on each sample. First, the chemical composition of the steel samples was determined. Then, the change in weight of each sample during the time it was placed inside the tank was calculated. Then, their tensile, bending, microstructure, and corrosion resistance were also measured.

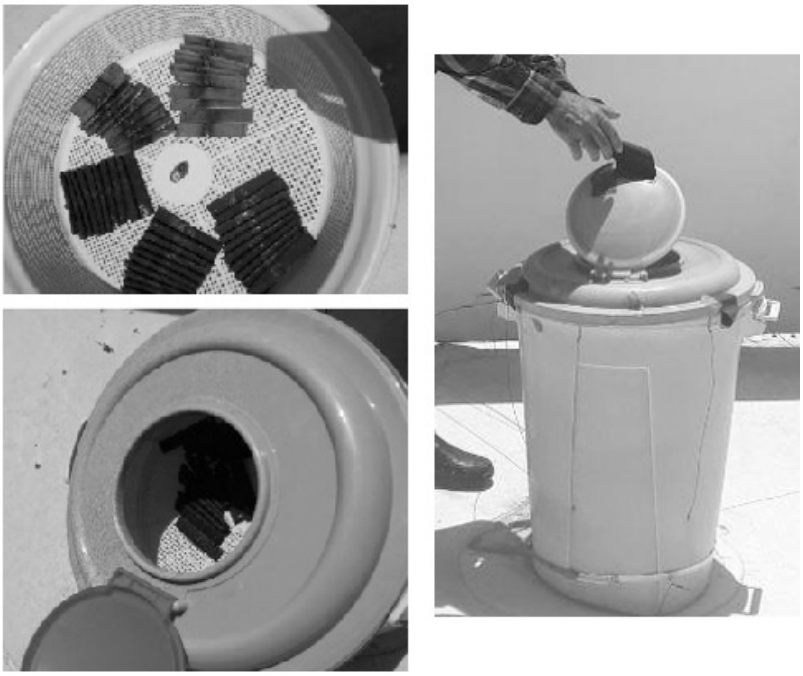


Figure 3-Designed chamber for mounting samples

2-9- Determining the change in sample weight

It should be noted that during the time the samples were inside the tank, they lost weight due to corrosion, so this weight loss was measured using a digital scale model S1202. In order to calculate the percentage of weight loss of the

samples, equation 1 was used, in which W1, W2, and RW are the initial weight of the sample, the secondary weight of the sample, and finally the percentage of weight loss, respectively:

$$RW = \left(\frac{W_1 - W_2}{W_1} \right) \times 100 \quad \textbf{(1)}$$

2-10- Investigation of the microstructure of the weld

In order to investigate the microstructure of the weld, two imaging methods were used, with an optical microscope and a scanning electron microscope (SEM). These two imaging methods were used to determine the presence of impurities at the weld line junction in different areas of the weld line.

2-11- Tensile and flexural tests

The two main mechanical properties for each part include flexural and tensile strength. In this study, these two properties were measured for different samples8[ ]. The standards for performing tensile and flexural tests are ASTM E290 and ASTM E858, respectively. Figure 4 shows the device for performing tensile and flexural tests. The Centam STM-50 is the device that was used for this purpose, and by changing the jaws of the device, both tests were performed by it.

2-12-Polarization corrosion test

The polarization test is used to investigate corrosion behavior, find corrosion rates, investigate active-passive behavior, corrosion potential, protection potential and pitting potential of different welding areas] 9[. The contact surface includes the weld areas, HAZ and base metal. All these surfaces are sanded with sandpaper up to 1200 grit before the corrosion test and then degreased with acetone solution

and the remaining surfaces that should not be in the vicinity of the electrolyte solution are insulated with a special varnish.

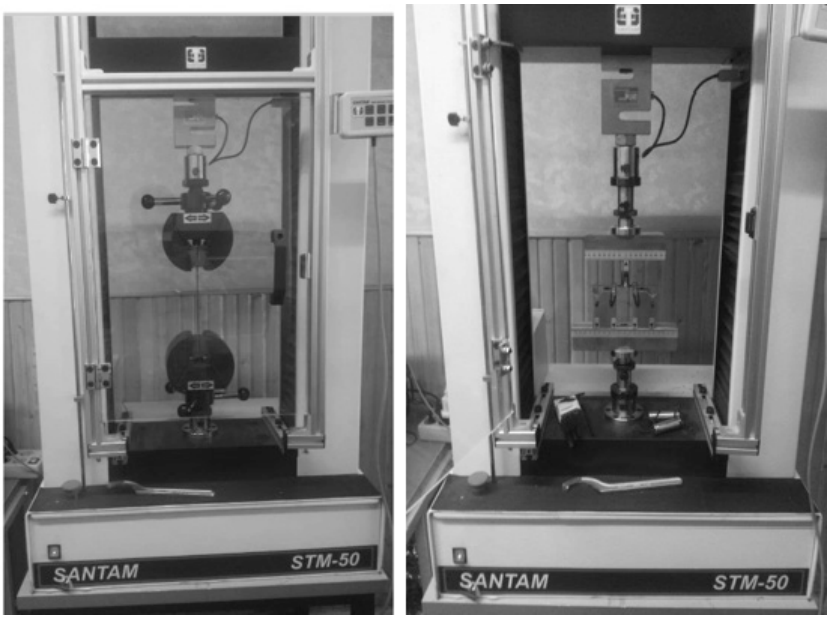


Figure 4- Centam STM 50 device

2-13- Numerical modeling of tensile and bending tests

In this study, two tensile and bending tests were numerically modeled based on the finite element method. For this purpose, Abaqus 2022 software was used. In order to model a problem in the Abaqus environment, different modules have been considered, each part of the modeling is defined separately in its own module and the modeling is completed based on that.

1- Part module: In this section, the schematic of the steel and the weld area are designed as two separate parts. For this purpose, and based on the exact dimensions and sizes of the welded samples, the weld area and the base metal were designed.

2- Property module: In this module, all the physical and mechanical properties of the components of the model under study will be defined. Considering that the weld material and the base metal material are different, the aforementioned properties were defined for both designed parts in this section. In this study, since each base metal and weld metal are fractured during tensile and bending tests, and this is a deformation in the plastic region, in addition to defining the elastic properties, the plastic properties of the studied materials must also be defined. In order to define these properties, these properties are defined in the Property module from the New option.

3- Assembly module: In this module, the parts designed in the previous section are connected together and show the overall structure of the designed weld. In Figure 5, the final model of the designed weld is shown.

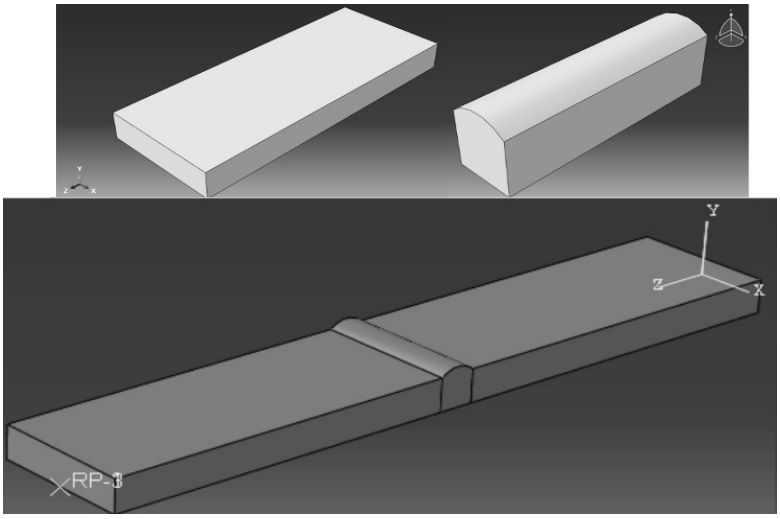


Figure 5-Final designed weld model and its components

4-Step module: In this module, the problem solving method will be defined. The problem under study is of a dynamic type, so for this purpose, the Dynamic Explicit solution method was selected. Due to the high volume of calculations in this

modeling, it is necessary to define the method in the Edit section and select the Factor index as 1000 from the Mass Scaling sub-category. This number was selected as a trial and error method so that convergence in solving the problem can be observed in the shortest possible time.

5-Interaction module: In this section, the connection between different surfaces in the designed weld model must be defined. Given that in this research, two steel base metals must be fixed around the designed weld model, in this case, two constraints called Tie must be defined to connect the surfaces of the weld model together. This definition is the same for both tensile and bending tests] 3[, as shown in Figure 6, the process of defining the connection between the base metal and weld surfaces.

On the other hand, for the tensile and bending tests, the load application constraints must be defined separately. In the tensile test, one side of the part is fixed and the other side is under the influence of the tensile force. In order to apply the tensile force to the part, a reference point was defined from the Tools menu and the reference point was connected to the cross-section of the welded part.

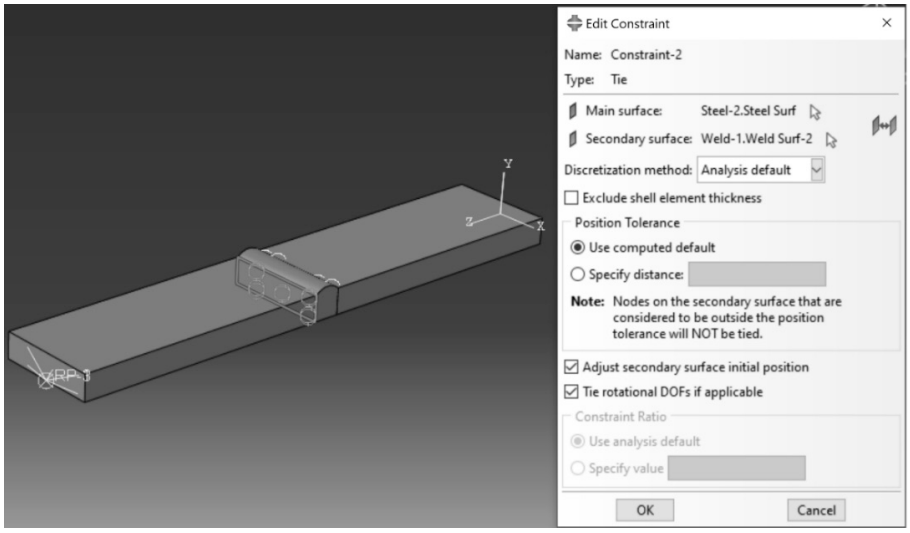


Figure 6-Definition of the relationship between weld plates and base metal with the Tie constraint

In Figure 7, the relationship between the reference point and the cross-section in the tensile test is shown. However, in the bending test, a third part was designed and the force application operation was performed by this part. In Figure 8, the relationship between the bending part and the weld is shown. For this case, a reference point was also defined and the reference point was connected to the effect surface of the bending part.

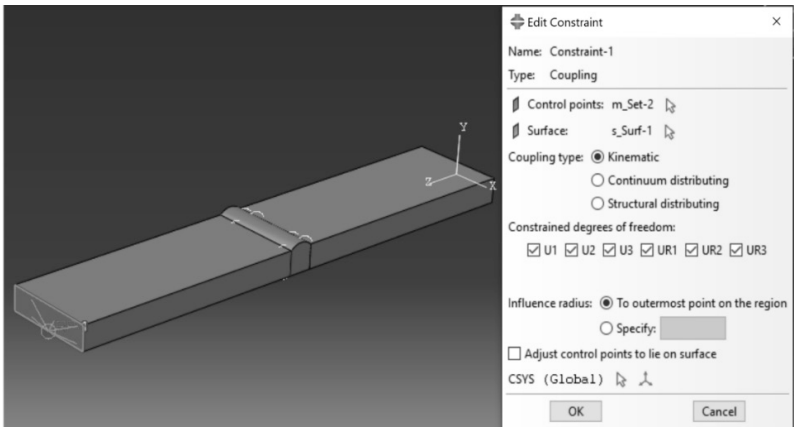


Figure 7-Definition of the relationship between tensile force and weld sample with the Coupling constraint

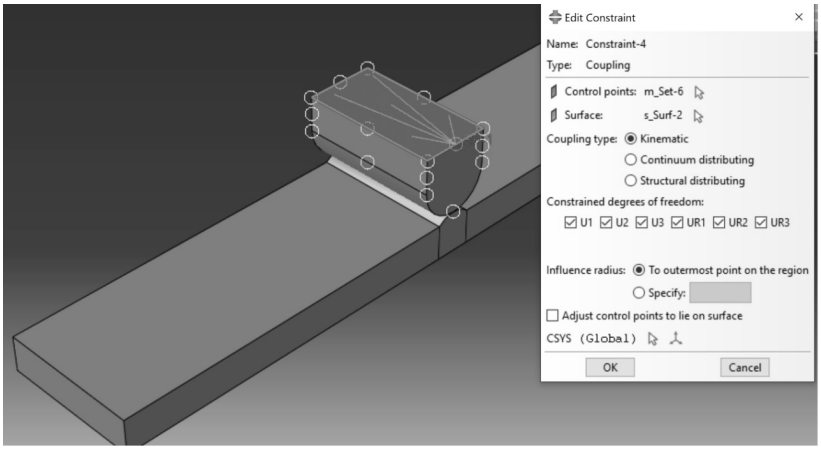


Figure 8-Definition of the relationship between bending force and weld sample with the Coupling constraint

6- Load Module: In this section, only the supports of the weld model are defined. In the tensile test mode, the support was welded with zero degrees of freedom on one side of the specimen. To define the support in the bending test, the supports were defined as zero degrees of freedom on both sides of the weld specimen.

7- Mesh Module: In this module, the research problem is divided into fine components, which is called meshing. In this mode, first select the designed part, then select the selected mesh type as Explicit Mesh. In the meshing settings section, the mesh size can be adjusted, and in this study, the mesh size for the weld area was considered finer than the base metal area. In Figure 9, the final meshed model for the bending test is shown. The meshed models for the bending and tensile tests are similar, and the only difference is the bending element, which is absent in the tensile model.

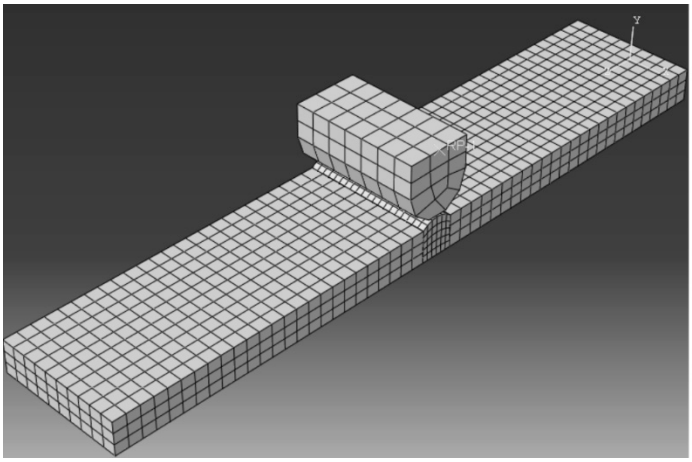


Figure 9-Final meshed model

8-Job and Visualization module: In the Job module, a new problem will be defined and by submitting it, the software will start solving the problem. Then, after solving the problem and completing the solution process, the results of the problem and modeling can be seen in the Visualization module.

2-14- Final analysis

In order to investigate the effect of the type of nanoparticle and its concentration on the mechanical properties, corrosion and percentage of weight loss of the samples, the analysis of variance test with a randomized factorial design with three replications was used. Also, to compare the average effect of the type and concentration of nanoparticles on the above properties, the mean comparison method using the Duncan multiple range test was used. SPSS software was used to analyze the data in this section.

3-Results and Discussion

3-1- Examination of the microstructure of different weld areas

The microstructure of the welded sample without the presence of nanoparticles and with the presence of nanoparticles was examined in this section. Optical and electron microscope images were prepared from different weld areas, including the weld button, the weld-affected area, and the base metal area. Figure 10 shows optical microscope images and electron microscope images for different weld areas without the presence of nanoparticles. Examination of the optical microscope images shows that the average grain size of the base metal is about 10 microns, which is ferrite. The results also show that the grain size in the base metal is larger than the weld button, but the shape of the grains in the weld button is more elongated than the base metal. The increase in temperature in the weld button area causes the grain size to become smaller, and for this reason the grain size in the weld button is larger than the base metal. In simpler terms, in the weld button, due to the high temperature of the molten area, the grains have recrystallized and their size has been elongated, while in the base metal, the recrystallization process has not been performed. Comparison and examination of the light microscope and SEM images showed that there is not much difference in the weld-affected area with the base metal, but in the weld-affected area,

the grain size was smaller than the base metal and the grains were more elongated in this area. Examination of the images showed that the grain size in the base metal is more uniform and almost the size of most grains is the same, while moving towards the weld button, this uniformity decreases and the grain size is non-uniform, which is due to the high temperature in the weld button and recrystallization] 10[.

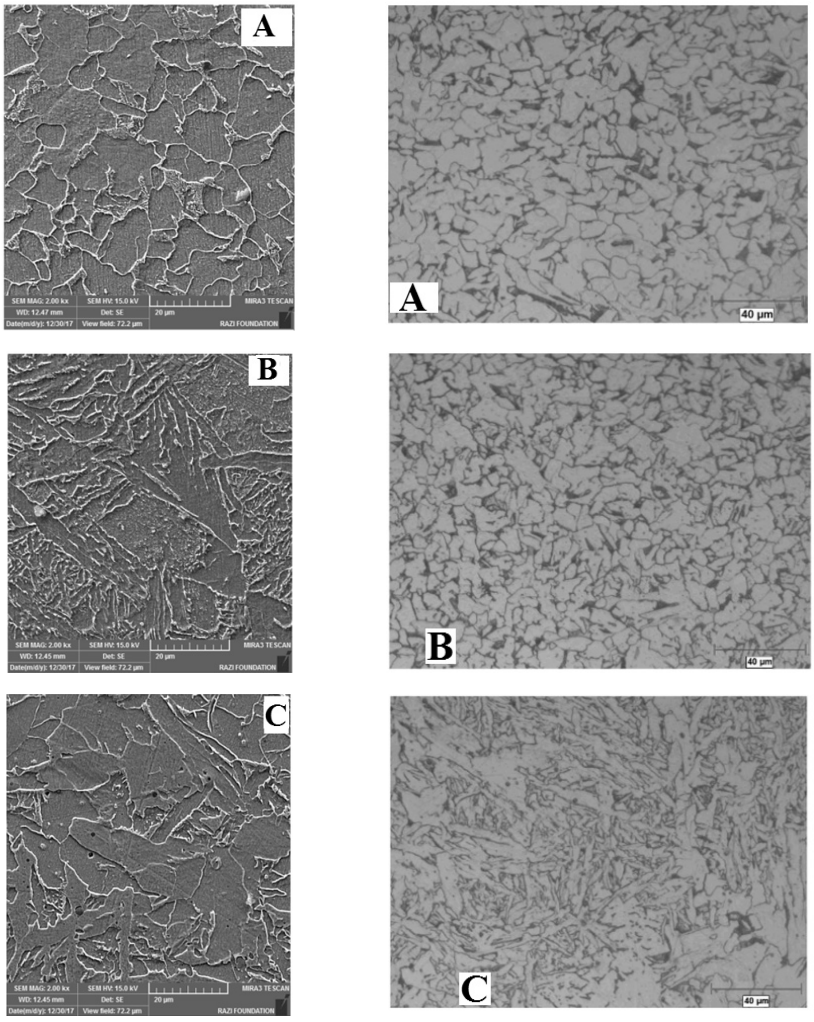


Figure 10- Right: Light microscope images. Left: SEM images of different weld areas (a- base metal, b- weld affected area, c- weld button)

3-2-Study of sample weight loss

In this study, the effect of nanoparticles and their concentration on the percentage of weight loss of the samples was studied. Each sample was weighed before the start of the test. Then, after leaving the tank, it was cleaned by sandblasting and washed with ethanol, and then its final weight was recorded. Then, the percentage of weight loss of the samples was calculated. The percentage of weight loss in this study was calculated for different welded samples with different weight percentages and types of nanoparticles. Table 4 presents the results of the analysis of variance of the effect of nanoparticle type, nanoparticle concentration, and their interaction effects on the percentage of weight loss of the samples. The results showed that firstly, the effect of repetition on the percentage of weight loss had no significant effect, and secondly, both the type of nanoparticle and the concentration, as well as their interaction effects, had an effect on the percentage of weight loss of the samples at the 99% probability level.] 11[

Table 4- Comparison of the average effect of nanoparticle type and concentration on the percentage of weight loss of welded samples

Dependent Variable: WR					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	34.978[a]	7	4.997	146.852	.000
Intercept	110.113	1	110.113	3236.074	.000
Rep	.054	2	.027	.790	.480
NanoType	.474	1	.474	13.921	.004
NanoCon	29.664	2	14.832	435.899	.000
NanoType * NanoCon	4.786	2	2.393	70.332	.000
Error	.340	10	.034		
Total	145.431	18			
Corrected Total	35.318	17			
a. R Squared = .990 (Adjusted R Squared = .984)					

Table 5 presents the results of the comparison of the means to select the best method for having the minimum weight loss based on the Duncan method .] 5[The lowest percentage of weight loss was observed for the sample welded with 2.5% titanium dioxide nanoparticles. The weight loss rate in this case was 0.23%. For both nanoparticle samples, with increasing nanoparticle concentration, the percentage of weight loss decreased and then increased. The weight loss rate for both nanoparticles was the highest weight loss in the sample with 1% nanoparticles. Examining the effect of the type of nanoparticle also showed that the effect of TiO2 nanoparticles on reducing corrosion and reducing the percentage of weight loss was greater than that of SiO2 nanoparticles. By changing the concentration of SiO2 nanoparticles to more than 1%, the difference between the percentage of weight loss was not significant. The highest weight loss was observed in the sample containing 1% TiO2 nanoparticles at 54.4%. With increasing nanoparticle concentration, initially the presence of nanoparticles between the grain boundaries prevented corrosion from intensifying, but for higher amounts of nanoparticles, they became clumped and caused holes between the grains, which was the reason for the greater weight loss of the samples.

Table 5- Comparison of the average effect of nanoparticle type and concentration on the percentage of weight loss of each sample

NanoType	Nano Concentration (w%)	Weight Reduction (%)
TiO2	1.0	4.5400[a]
	2.5	0.2300[d]
	5.0	2.1633[c]
SiO2	1.0	3.8333[b]
	2.5	1.9633[c]
	5.0	2.1100[c]

Common letters in each column indicate no significant difference between them.

Figure 11 shows the difference in weight loss for samples containing nanoparticles compared to samples without nanoparticles. An examination of these results shows that using nanoparticles can greatly improve the weight loss rate, so that when using 2.5% TiO2 nanoparticles, the percentage of sample mass loss has improved by about 97.8% compared to the sample without nanoparticles. In general, an examination of the results of this section showed that due to the presence of nanoparticles in the welding process, the weight loss rate of the samples has decreased, which shows the role of nanoparticles as a protective factor against corrosion by gases inside the tank. According to the examination of the weight loss results, the effects of adding TiO2 nanoparticles were more significant compared to SiO2 nanoparticles; therefore, the effect of titanium dioxide nanoparticles on the corrosion and mechanical properties of the weld was investigated.

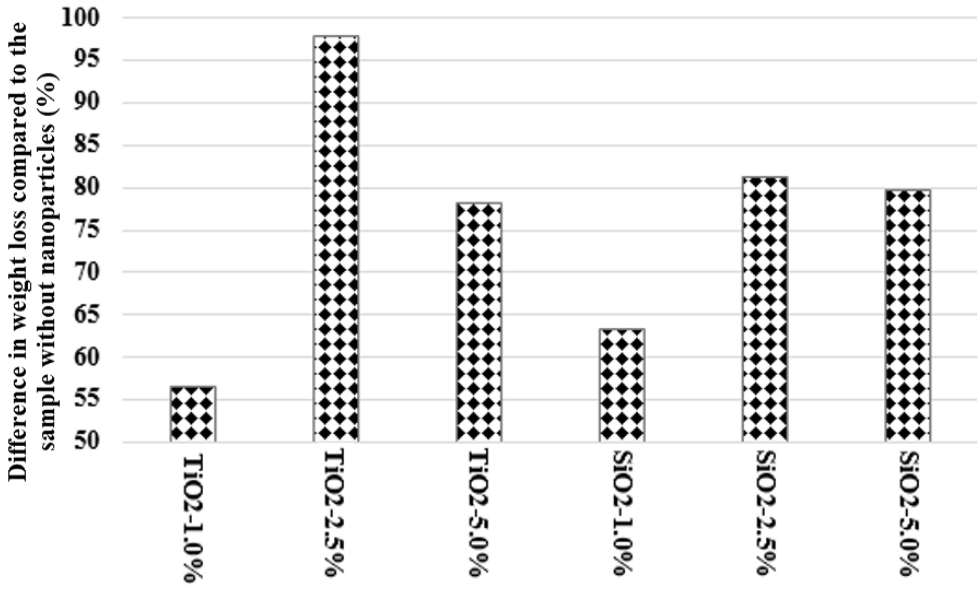


Figure 11-Difference between percentage weight loss for samples containing nanoparticles and samples without nanoparticles

3-3- Investigation of corrosion rates of samples

In this section, the results of the polarization corrosion test for 4 different samples are reported. Table 6 presents the results of the TOEFL corrosion test for potential values and corrosion current intensity. Examination of the results of this table showed that the highest corrosion rate, which is also described by corrosion current intensity, is related to the welded sample without nanoparticles. The corrosion current intensity in this case is 1.03 mA/cm2, which is a significant difference compared to the welded samples with the presence of nanoparticles. With increasing the concentration of titanium dioxide nanoparticles in the welding process, the corrosion current intensity decreased and increased. The lowest corrosion current intensity was observed when using 2.5% nanoparticles. Comparing the corrosion current intensity for concentrations of 1%, 2.5% and 5% with the sample without nanoparticles showed that the corrosion rate decreased by 87.38%, 97.38% and 95.92%. The difference between the percentage of corrosion rate reduction for the 2.5% and 5% samples was not significant. Overall, the results of this section showed that the presence of nanoparticles in the welding process improved the corrosion properties of welded joints, which was also proven in the previous section. The best case in this section for recommending welding is the use of a filler containing 2.5% titanium oxide nanoparticles, which had a significant effect on reducing the corrosion intensity in the samples.

Table 6- Results of the polarization corrosion test for different samples

Label	Corrosion potential E_{corr}(V)	Corrosion current density I_{corr}(mA/cm^2)
Without Nano	-0.454	1.03
1.0 % TiO2	-0.321	0.13
2.5 % TiO2	-0.326	0.027
5.0 % TiO2	-0.331	0.042

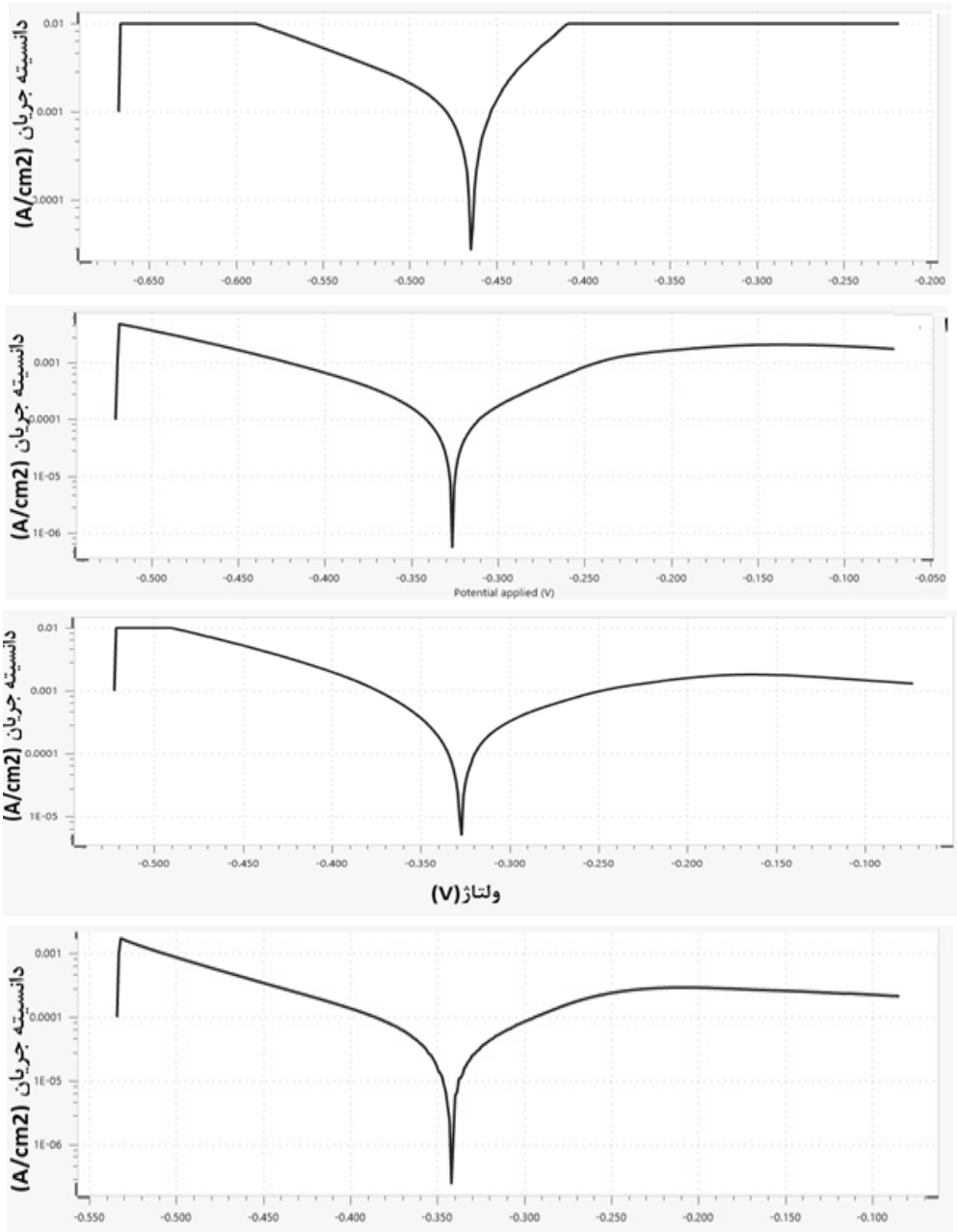


Figure 12- Corrosion test results for different samples containing TiO2 nanoparticles (from top to bottom for samples containing 0%, 1%, 2.5% and 5%, respectively)

3-4- Tensile and bending test results

Figure 13 shows the results of the three-point bending test in the form of stress-strain curves for different welded samples for different concentrations of TiO2 nanoparticles. Examination of the stress-strain curves showed that with the addition of TiO2 nanoparticles, the flexural strength and lateral strain of the samples underwent significant changes. As also shown in the metallographic image analysis section, with the addition of 2.5% TiO2 nanoparticles to the weld, the grain size became finer, which improved the mechanical properties of the studied samples. The results show that in the absence of TiO2 nanoparticles, the maximum flexural strength of the samples was 35.86 MPa, and with the addition of 1%, 2.5%, and 5% TiO2 nanoparticles, the flexural strength was 38.73 MPa, 40.33 MPa, and 40.02 MPa, respectively. The results of the flexural test showed that the addition of 1%, 2.5%, and 5% TiO2 nanoparticles during welding improved the maximum flexural strength of the samples by 8%, 11.17%, and 10.74%.

The results of the uniaxial tensile test are presented in the form of stress-strain curves in Figure 14. Examination of the stress-strain curves showed that by adding different concentrations of nanoparticles to the weld, the ultimate tensile stress increased, while the axial strain decreased. A comparison of the average maximum tensile stress of the samples showed that in the absence of nanoparticles, the maximum tensile strength of the sample was 15.534 MPa, which increased by 6.86%, 9.93%, and 8.80% with the addition of 1%, 2.5%, and 5% TiO2 nanoparticles, respectively.

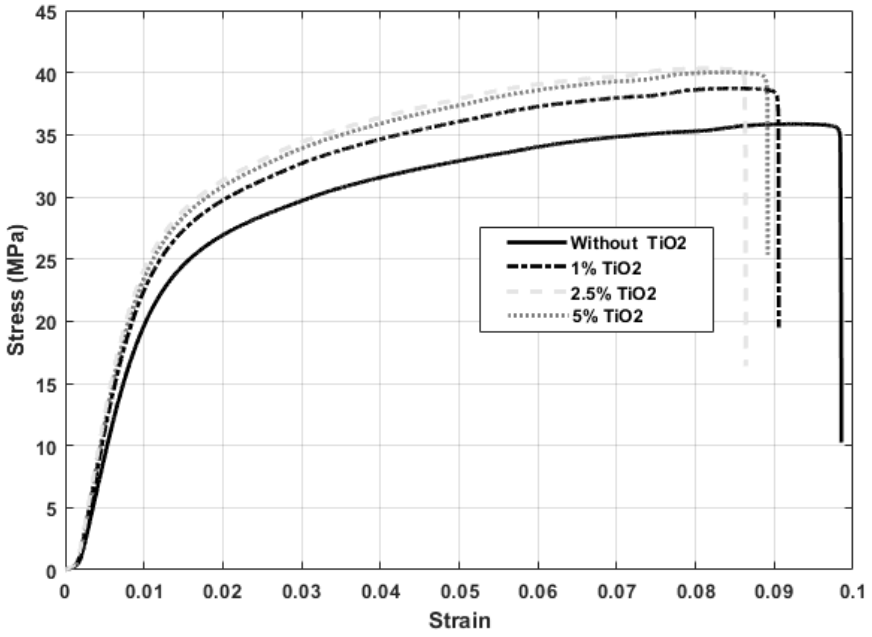


Figure 13- Three-point bending stress-strain curve for samples reinforced with different concentrations of TiO2 nanoparticles

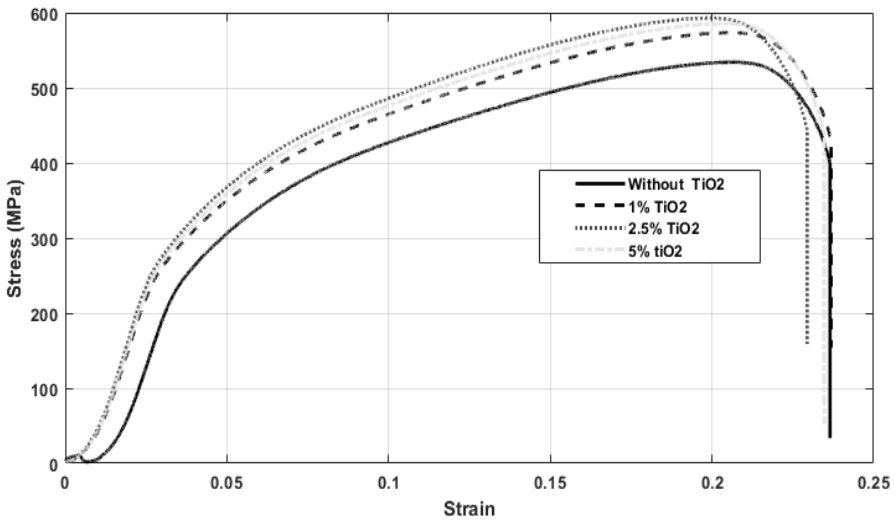


Figure 14-Tensile test stress-strain curve for specimens reinforced with different concentrations of TiO2 nanoparticles

3-5- Numerical modeling results of tensile and bending tests

Figure 15 shows the results of finite element modeling of the bending test for the welded specimen. As can be seen in the figure, the modeled specimen was bent by moving the mandrel downwards and the modeled part did not break during the bending process.

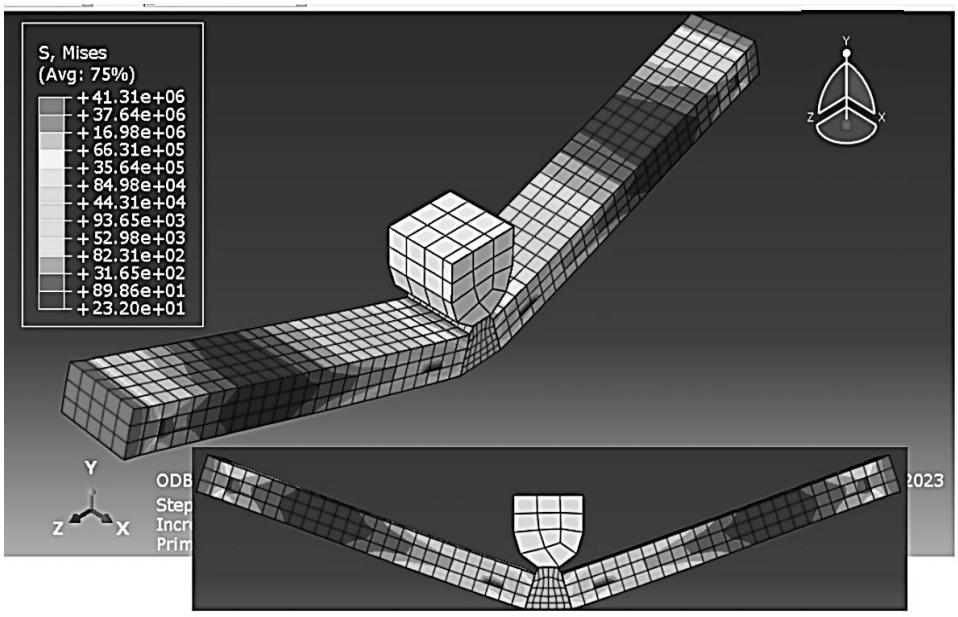


Figure 15- Numerical modeling results of bending test

Figure 16 also shows the experimental bending test result for the specimen. As can be seen in the figure, the bent model and the experimental model were similar and in both models, the failure process did not occur in the part. Examination of Figure 16 showed that in the three-point bending test, the maximum stress occurred at the weld and steel plate junction, or the area affected by the weld, and also at the supports. The maximum bending stress in the modeled sample was 41.31 MPa, while the maximum stress in the experimental sample was 35.86 MPa.

An examination of the numerical and experimental results shows that 13.19% error was observed between the finite element model and the experimental model, and Figure 17 also presents the stress-strain curve of the bending test for the experimental and numerical models. An examination of this figure shows that the trend of changes in the stress-strain curve for the numerical and experimental models in the bending test is similar, but the stress values in the numerical model are higher than in the experimental model. Overall, a good correlation was observed between the experimental and numerical results for the three-point bending test.

Figure 16- Laboratory sample after bending test

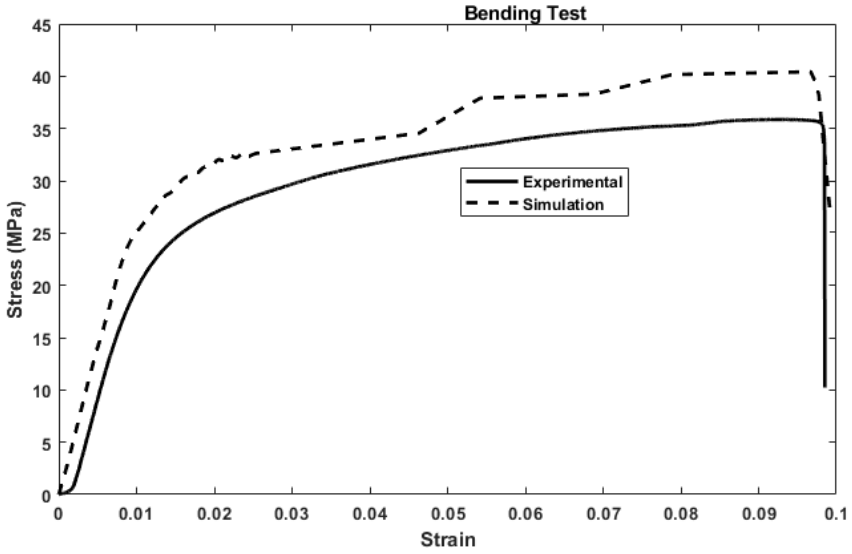


Figure 17- Comparison of stress-strain curve of bending test for experimental and modeled sample by Abaqus

In Figure 18, the modeled finite element model is shown under uniaxial tensile test. Examination of this figure shows that the maximum tensile stress was observed in the weld area and its value was 582.61 MPa, while the actual maximum tensile stress of the experimental sample was 534.15 MPa. The maximum tensile strength for the modeled sample was 9.07% higher than the experimental sample.

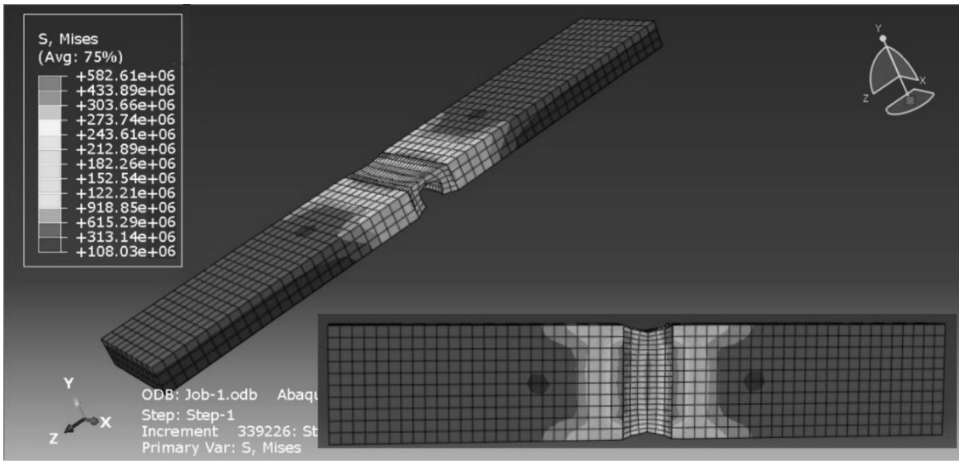


Figure 18-Numerical modeling results of tensile test

In Figure 19, the stress-strain curve for the experimental model and the finite element model under tensile test is shown. In this case, as in the results of the bending test section, the amount of stress that occurred in the modeled sample was higher than the real and experimental model. As in the results shown in the previous sections, some intergranular deposits are formed during welding during welding, and these deposits are important factors in reducing the mechanical properties of the weld. Therefore, the effects of these deposits were ignored during the modeling process, so the final stress value obtained from the numerical models was higher than the real model.

Figure 20 shows the experimental specimen after the tensile test. In this case, the specimen failed near the weld line, which was also the case in the numerical model, but in the numerical model, the failure location was exactly in the middle of the weld line, while in the experimental specimen, the failure location was in the weld-affected area on the tensile force side of the tensioning device.

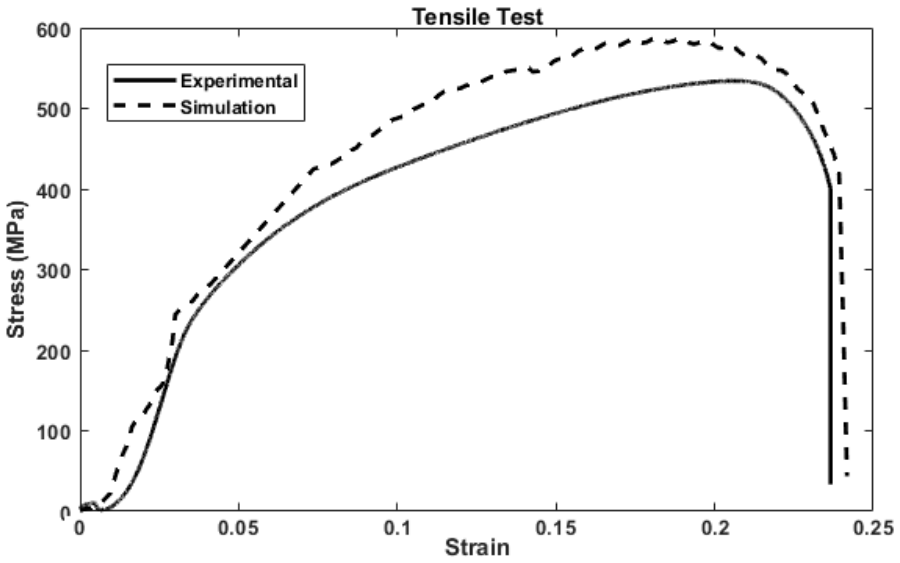


Figure 19-Comparison of tensile test stress-strain curve for experimental and Abaqus modeled samples

Figure 20-Laboratory sample after tensile test

4- Conclusion

In this thesis, the following results were observed:

1. Examination of the microstructure of different parts of the weld line showed that the grain size, grain distribution and uniformity in grain distribution were greater in the base metal area than in the weld-affected area and the weld-affected area than in the weld button area. In the weld button area, the grain size was smaller and the grains were more elongated than in the base metal and weld-affected area.
2. Examination of the weight loss of the welded samples showed that adding nanoparticles during the welding process caused a decrease in the percentage of weight loss. So that for a sample containing 2.5% titanium dioxide nanoparticles, the lowest percentage of weight loss was observed at 0.23%. The results also showed that the effect of titanium dioxide nanoparticles is more effective than silicon dioxide in reducing the weight loss rate.
3. Examination of the corrosion test of the samples showed that with increasing the concentration of nanoparticles in the nanofiller, the corrosion current intensity at the weld site first decreased and then increased. The minimum corrosion current intensity was observed when using nanofiller with 2.5% titanium dioxide nanoparticles. By using nanofillers containing 1%, 2.5% and 5% nanoparticles, the corrosion current intensity at the weld site can be reduced by 87.38%, 97.38% and 95.92%, respectively.
4. Examination of the tensile and bending test results showed that adding TiO2 nanoparticles during the welding process increased the maximum

flexural and tensile strength of the welded samples, so that for the sample with 2.5% nanoparticles, the maximum flexural and tensile strength increased by 10.74% and 9.93%, respectively.

5. Comparison of numerical and experimental modeling results showed that the error of numerical modeling using the finite element method in measuring tensile and flexural strength was 13.19% and 9.07%, respectively.

5- References

[1] UNESCO. (2011). *Recommendation on the Historic Urban Landscape*. UNESCO World Heritage Centre. https://whc.unesco.org/en/hul/?utm

[2] Bandarin, F., & van Oers, R. (2012). *The historic urban landscape: Managing heritage in an urban century*. Wiley Blackwell, 2012, 236 pp. 75-111p, colour photos, ISBN 978-0-470-65574-0 https://onlinelibrary.wiley.com/doi/book/10.1002/9781119968115?utm_source=chatgpt.com

[3] Pacione, M. (2003). *Urban environmental quality and human wellbeing—A social geographical perspective. Landscape and Urban Planning*, 65(1–2), 19–30. https://doi.org/10.1016/S0169-2046(02)00234-7

[4] Marans, R. W. (2012). Quality of urban life studies: An overview and implications for environment–behaviour research. Procedia – Social and Behavioral Sciences, 35, 9–22. https://doi.org/10.1016/j.sbspro.2012.02.058

[5] Labadi, S., & Logan, W. (2015). *Urban Heritage, Development and Sustainability.* Routledge. https://www.taylorfrancis.com/books/edit/10.4324/9781315728018/urban-heritage-development-sustainability-sophia-labadi-william-logan?utm_source=chatgpt.com

[6] Matthias R., Christer G., Zachary J., Sushobhan M. and Manal G. (2025) Historic Urban Landscapes and Heritage Systems as the Basis for Sustainable Urban Development, *14*(9), 1783;https://doi.org/10.3390/land14091783

[7] Gehl, J. (2011). *Life between buildings: Using public space* (6th ed.). Island Press.

[8] Carmona, M. (2019)Principles for public space design, planning to do better. *Urban Des Int* 24, 47–59. https://doi.org/10.1057/s41289-018-0070-3

[9] Rypkema, D. D. (2005). *The economics of historic preservation*. National Trust for Historic Preservation.

[10] Pendlebury, J. (2013). Conservation values, the authorised heritage discourse and the conservation-planning assemblage. *International Journal of Heritage Studies*, *19*(7), 709–727. https://doi.org/10.1080/13527258.2012.700282

[11] Graham, B., Ashworth, G. J., & Tunbridge, J. E. (2000). *A geography of heritage: Power, culture and economy*. Arnold.

[12] UN-Habitat. (2015). *Global public space toolkit: From global principles to local policies and practice*. United Nations Human Settlements Program.

[13] Jamal A. Measuring Quality of Life in Urban Areas: Toward an Integrated Approach. Int J Environ Sci Nat Res. 2020; 25(2): 556158. DOI: 10.19080/IJESNR.2020.25.556158

[14] Hillier, B., & Hanson, J. (1984). *The social logic of space*. Cambridge University Press.

[15] Hillier, B. (2007). *Space is the machine: A configurational theory of architecture*. Space Syntax.

[16] Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). SAGE Publications.

[17] Groat, L., & Wang, D. (2013). *Architectural research methods* (2nd ed.). John Wiley & Sons.

[18] Likert, R. (1932). *A technique for the measurement of attitudes*. Archives of Psychology, 140, 1–55.

[19] Whyte, W. H. (1980). *The social life of small urban spaces*. Project for Public Spaces.

[20] Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.

[21] M. S. Ibrahim and A. I. Altif, (2017) "Architectural Elements in Mustansiriyah Madrasah in Baghdad," Al MalweahArchaeol. Hist. Stud., vol. 4, no. 7, pp. 253–282, 2017 *(PDF) Sustainability Investigation of the Old Mustansiriya Madrasa in Modern Construction: A Comparative Study*. Available from: https://www.researchgate.net/publication/393209498_Sustainability_Investigation_of_the_Old_Mustansiriya_Madrasa_in_Modern_Construction_A_Comparative_Study [accessed Jan 16 2026].

[22] Abdulrahman, M. @mustafa. bdulrahman (2025, اوكتوبر 03). *صورة جوية للفناء أثناء الفعالية المسائية في المدرسة المستنصرية [صورة].* Instagram. https://www.instagram.com/mustafa.bdulrahman/

46. Viale Pereira G, Cunha MA, Lampoltshammer TJ, et al (2017) Increasing collaboration and participation in smart city governance: A cross-case analysis of smart city initiatives. Inf Technol Dev 23:526–553

47. Alobaydi AF, Alobaydi D, Abdullah SF (2025) Assessing the Influence of Urban Street Network on Land Use Distribution in Historic Urban Cores. Iraqi Journal of Architecture & Planning 24:

48. Anthony Jnr B, Abbas Petersen S, Ahlers D, Krogstie J (2020) Big data driven multi-tier architecture for electric mobility as a service in smart cities: A design science approach. International Journal of Energy Sector Management 14:1023–1047

49. Viale Pereira G, Cunha MA, Lampoltshammer TJ, et al (2017) Increasing collaboration and participation in smart city governance: A cross-case analysis of smart city initiatives. Inf Technol Dev 23:526–553

50. Pozoukidou G, Chatziyiannaki Z (2021) 15-Minute City: Decomposing the new urban planning eutopia. Sustainability 13:928

51. Barbieri L, D'Autilia R, Marrone P, Montella I (2023) Graph representation of the 15-minute city: A comparison between Rome, London, and Paris. Sustainability 15:3772

52. Wang C (2025) 15-Minute City: Future Urban and Transportation Development Mode

53. Zakariasson A (2022) A study of the 15-minute city concept: Identifying strengths, risks and challenges through imagining the implementation of the 15-minute city concept in Munich, Germany

54. Hou L, Liu Y (2017) Life circle construction in China under the idea of collaborative governance: A comparative study of Beijing, Shanghai and Guangzhou. Geographical review of Japan series B 90:2–16

55. Kagan C (2019) The Future Is 20 Minutes Away? 20-Minute Neighborhoods. Steady State Manchester Team

56. Bates L (2018) City of Portland, Bureau of Planning and Sustainability. Gentrification and Displacement Study: implementing an equitable inclusive development strategy in the context of gentrification Last Accessed 16:

31. Nathan M (2015) After Florida: Towards an economics of diversity. Eur Urban Reg Stud 22:3–19

32. Hartmann T, Jehling M (2019) From diversity to justice–Unraveling pluralistic rationalities in urban design. Cities 91:58–63

33. Abbas SS, Najat RF (2019) Mixed Land Use and its Effect in Integration of the Residential Neighborhood. Journal of University of Babylon for Engineering Sciences 27:59–75

34. Allam Z, Nieuwenhuijsen M, Chabaud D, Moreno C (2022) The 15-minute city offers a new framework for sustainability, liveability, and health. Lancet Planet Health 6:e181–e183

35. Sharifi A, Khavarian-Garmsir AR, Kummitha RKR (2021) Contributions of smart city solutions and technologies to resilience against the COVID-19 pandemic: A literature review. Sustainability 13:8018

36. Hou Y, Moogoor A, Dieterich A, et al (2021) Exploring built environment correlates of older adults' walking travel from lifelogging images. Transp Res D Transp Environ 96:102850

37. Yüksek I, Karadayi TT (2017) Energy-efficient building design in the context of building life cycle. Energy efficient buildings 1:93–123

38. dos Santos HA, da Silva Santana E, Bueno RE, Bonilla SH (2021) Technologies helping smart cities to build resilience: Focus on COVID-19. In: IFIP International Conference on Advances in Production Management Systems. Springer, pp 714–723

39. Joss S, Cowley R, Tomozeiu D (2013) Towards the 'ubiquitous eco-city': An analysis of the internationalisation of eco-city policy and practice. Urban Res Pract 6:54–74

40. Pearson RM, Schlacher TA, Jinks KI, et al (2021) Disturbance type determines how connectivity shapes ecosystem resilience. Sci Rep 11:1188

41. Sharifi A, Khavarian-Garmsir AR (2020) The COVID-19 pandemic: Impacts on cities and major lessons for urban planning, design, and management. Science of the total environment 749:142391

42. Balletto G, Ladu M, Milesi A, Borruso G (2021) A methodological approach on disused public properties in the 15-minute city perspective. Sustainability 13:593

43. Afrin S, Chowdhury FJ, Rahman MM (2021) COVID-19 pandemic: rethinking strategies for resilient urban design, perceptions, and planning. Frontiers in Sustainable Cities, 3, 668263

44. Mobaraki O (2023) The 15-minute city: A paradigm for urban planning and policy. Sustainable Earth Review 3:

45. Pearson RM, Schlacher TA, Jinks KI, et al (2021) Disturbance type determines how connectivity shapes ecosystem resilience. Sci Rep 11:1188

15. Jacobs J (1961) Jane jacobs. The Death and Life of Great American Cities 21:13–25

16. Zaheer A (2019) Sustainability and resilience in megacities through energy diversification, land fragmentation and fiscal mechanisms. Sustain Cities Soc 53: 101841

17. Low M (2013) Eco-cities in Japan: past and future. Journal of Urban Technology 20:7–22

18. Warwick E (2015) Policy to reality: Evaluating the evidence trajectory for English eco-towns. Building Research & Information 43:486–498

19. Basiago AD (1996) The search for the sustainable city in 20th century urban planning. Environmentalist 16:135–155

20. Barbarossa L (2020) The post pandemic city: Challenges and opportunities for a non-motorized urban environment. An overview of Italian cases. Sustainability 12:7172

21. Pisano C (2020) Strategies for post-COVID cities: An insight to Paris En Commun and Milano 2020. Sustainability 12:5883

22. Moreno C, Allam Z, Chabaud D, et al (2021) Introducing the “15-Minute City”: Sustainability, resilience and place identity in future post-pandemic cities. Smart cities 4:93–111

23. Moreno K (2021) Definition of the 15-minute city: What is the 15-minute city

24. Iqbal A, Nazir H, Qazi AW (2025) Exploring the 15-Minutes City Concept: Global Challenges and Opportunities in Diverse Urban Contexts. Urban Science 9:252

25. Tarwani A (2020) A new time-based urban agenda. Exploring the 15-minute city in concepts and practices

26. del Saz Salazar S, Menendez LG (2007) Estimating the non-market benefits of an urban park: Does proximity matter? Land use policy 24:296–305

27. Zaheer A (2020) Sustainability and resilience in megacities through energy diversification, land fragmentation and fiscal mechanisms. Sustain Cities Soc 53:101841

28. Guzman LA, Oviedo D, Cantillo-Garcia VA (2024) Is proximity enough? A critical analysis of a 15-minute city considering individual perceptions. Cities 148:104882

29. Khavarian-Garmsir AR, Sharifi A, Sadeghi A (2023) The 15-minute city: Urban planning and design efforts toward creating sustainable neighborhoods. Cities 132:104101

30. Wan Mohd Rani WNM (2014) Evaluating the impact of density on access to local facilities in urban neighbourhoods. Planning Malaysia: Journal of the Malaysian Institute of Planners 12:1–18

8. المصادر

1. Khavarian-Garmsir AR, Sharifi A, Hajian Hossein Abadi M, Moradi Z (2023) From garden city to 15-minute city: a historical perspective and critical assessment. Land (Basel) 12:512
2. Hasan NA, Al-Hinkawi WS, Alatta A (2017) The role of planning urban axes in the sustainability of Iraqi cities Mosul-case study. International Journal of Architecture (IJA) 3:18–33
3. E. Howard (1902) Garden Cities of To-morrow. London, UK
4. Alshuwaikhat HM (1999) Planning the 21st century urban neighborhood: learning from previous concepts. Journal of King Saud University Architecture & Planning 11:13–29
5. Mehaffy MW, Porta S, Romice O (2015) The “neighborhood unit” on trial: A case study in the impacts of urban morphology. Journal of Urbanism: International Research on Placemaking and Urban Sustainability 8:199–217
6. Rohe WM (2009) From local to global: One hundred years of neighborhood planning. Journal of the American planning association 75:209–230
7. Perry C (1929) The Neighborhood Unit, a Scheme of Arrangement for the Family-Life Community. Reg. Surv. N. Y. Its Environ. 2–140
8. Perry C (2020) “The neighborhood unit”: From the regional plan of New York and its environs (1929). In: The city reader. Routledge, pp 557–569
9. Sharifi A (2016) From Garden City to Eco-urbanism: The quest for sustainable neighborhood development. Sustain Cities Soc 20:1–16
10. Basiago AD (1996) The search for the sustainable city in 20th century urban planning. Environmentalist 16:135–155
11. Watson V (2009) ‘The planned city sweeps the poor away…’: Urban planning and 21st century urbanisation. Prog Plann 72:151–193
12. Gillette H (2012) Building Better Communities, from the Garden City to the New Urbanism
13. Handy S (2020) Is accessibility an idea whose time has finally come? Transp Res D Transp Environ 83:102319
14. MacLeod G (2013) New urbanism/smart growth in the Scottish Highlands: Mobile policies and post-politics in local development planning. Urban Studies 50:2196–2221

7. التوصيات

- مراجعة وتحديث قوانين البناء وأنظمة تقسيم المناطق الحالية التي تدعم الفصل الوظيفي الصارم، واستبدالها بتشريعات مرنة تشجع على الاستعمال المختلط العمودي والأفقي، لضمان تقارب السكن مع العمل والخدمات.
- التحول من اعتماد المسافة المترية وسعة الشوارع كمعايير أساسية في تقييم المخططات، إلى اعتماد زمن الوصول وجودة تجربة المشي كمؤشرات رئيسة لنجاح التصميم الحضري.
- تبني استراتيجيات لتفكيك المركزية الخدمية وتوزيع المرافق الحيوية (الصحية، التعليمية، والترفيهية) بشكل هرمي وعادل داخل الأحياء السكنية لضمان تحقيق مبدأ العدالة المكانية.
- إعادة تصميم الشوارع المحلية لتعطي الأولوية للمشاة وراكبي الدراجات على حساب السيارات، من خلال توسيع الأرصفة، وتوفير الممرات الآمنة، واعتماد التهدئة المرورية داخل المجاورات.
- نظراً للظروف المناخية المحلية الحارة، يوصي البحث بضرورة دمج حلول التصميم المناخي (توفير الظل، التشجير الكثيف، الممرات المبردة طبيعياً) كشرط أساسي لنجاح نموذج الـ 15 دقيقة، حيث لا يمكن تشجيع المشي دون بيئة حرارية مريحة.
- استغلال الأراضي المتروكة والفضاءات غير المستغلة داخل الأحياء السكنية القائمة لتحويلها إلى جيوب حضرية أو مراكز خدمة مجتمعية صغيرة لتعزيز التفاعل الاجتماعي.
- اعتماد الجدول النظري الذي استخلصه البحث (المتضمن لمؤشرات الكثافة، التنوع، والرقمنة) كـدليل إرشادي للمخططين والمعماريين عند إعداد المخططات الأساسية للمدن الجديدة أو عند إعادة تأهيل الأحياء القائمة.

6. الاستنتاجات

في ضوء التحليل النظري لمسار تطور المجاورة الحضرية من البعد الحداثوي وصولاً إلى مرحلة ما بعد الجائحة، ومن خلال استقراء التجارب العالمية والدراسات السابقة وتفكيك مؤشراتها، خلص البحث إلى مجموعة من الاستنتاجات التي تؤكد أن نموذج مجاورة 15 دقيقة يمثل نقلة نوعية في هيكلية المدن المعاصرة، ويمكن إيجاز أهم هذه الاستنتاجات في النقاط الاتية:

- يمثل مفهوم مجاورة 15 دقيقة تحولاً من التخطيط القائم على السرعة والمسافة إلى التخطيط القائم على الزمن والجودة، حيث ترتبط جودة الحياة طردياً بتقليل الوقت المهدور في التنقل.
- يعد القرب الركيزة الأساسية للمفهوم، ويهدف إلى تقليل الحاجة للبنية التحتية الضخمة وتحسين الوصول للفرص التعليمية والاقتصادية سيراً على الأقدام.
- إن التكامل بين خصائص مجاورة 15 دقيقة هو ما يحول المجاورة من مجرد تجمع سكني إلى نظام إيكولوجي حيوي قادر على خفض الانبعاثات الكربونية والصمود في وجه التحديات المناخية والصحية المستقبلية
- كشفت الدراسة التحليلية للتجارب العالمية (خاصة بورتلاند وباريس) أن الأحياء المصممة وفق مبادئ مجاورة 15دقيقة أظهرت مرونة حضرية أعلى في مواجهة الصدمات (مثل جائحة كوفيد-19) مقارنة بالأحياء الحداثوية المعتمدة على الفصل الوظيفي.
- نقل الإطار النظري مفهوم مجاورة 15 دقيقة من الطرح النظري إلى أداة تخطيطية كمية قابلة للتطبيق والقياس

المفردة الرئيسة	المفردة الثانوية	المتغيرات	المؤشرات والقيم الممكنة
مجاورة 15 دقيقة	الرقمنة	رقمنة الخدمات الحضرية	عدد الخدمات المتاحة رقميا
		استخدام المنصات الرقمية	وجود/ عدم وجود منصات رقمية
		دعم اتخاذ القرار الرقمي	نوع الخدمة (حجز-دفع-معلومات)
		تقليل الحاجة للحركة المادية	مستوى الإستخدام (منخفض-متوسط-عالي)
	الترابط	درجة نفاذية الشبكة	عدد الممرات المفتوحة
			وجود/ عدم وجود حاجز
			طول المسارات المتصلة م
		قيمة التكامل المكاني	قيم التكامل محلي/عالمي
	امكانية الوصول	امكانية للوصول للخدمات	الطريق سالك ام مقطوع
			تغطية الخدمات ضمن مسافة معينة
			عدد التقاطعات للوصول إلى الخدمة
			الزمن المستغرق للوصول
	قابلية المشي	كثافة شبكة المشاة	طول وعدد مسارات المشي/ هكتار
		كثافة التقاطعات المهيأة للمشاة	عدد التقاطعات القابلة للعبور الآمن
		جودة مسار المشي	الاثاث الحضري
			وجود ارصفة/ عرض الرصيف
			التشجير
			المسافة الى اقرب معبر مشاة

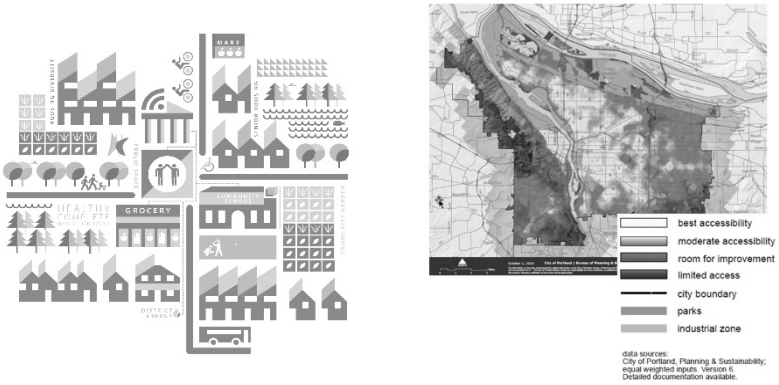
تركز على تحسين نوعية الوصول اليومي وتهيئة بيئة حضرية داعمة للمشي والتنقل النشط. ويعكس هذا التوجه انتقال التخطيط من قياس القرب الزمني المجرد إلى تقييم التجربة المعيشية الفعلية للسكان داخل الحي وان نجاح المجاورة في المدن متوسطة الكثافة يعتمد على تكامل السياسات (النقل، الصحة، العدالة المكانية، والاقتصاد المحلي)، وأن التخطيط القائم على التنقل النشط يوفر مرونة حضرية عالية، مما مكن المدينة من التكيف السريع مع الأزمات مثل جائحة كوفيد-19، رغم التحديات التي تواجه المدن متوسطة الكثافة في تحقيق الاكتفاء الذاتي الكامل.

5. الاطار النظري

بناءً على المناقشة وتحليل الدراسات والتطبيقات العالمية توصل البحث الى بناء أطار نظري شامل يحدد المفردات الرئيسة والمتغيرات الحاكمة لتخطيط وتصميم مجاورة 15 دقيقة كما هو موضح في الجدول (1) ادناه.

جدول (1) يبين خلاصة المفردات والمؤشرات المرتبطة بالاطار النظري. المصدر اعداد الباحثة استنادا لما سبق

المفردة الرئيسة	المفردة الثانوية	المتغيرات	المؤشرات والقيم الممكنة
مجاورة 15 دقيقة	القرب	قرب الخدمات الاساسية	عدد الخدمات داخل المنطقة
		المسافة المكانية الى الخدمات	400-800-1200م
		زمن الوصول سيرا على الاقدام	5-10-15 دقيقة
		نطاق الخدمة	نسبة التغطية %
	الكثافة	الكثافة السكانية	عدد السكان/ هكتار
		الكثافة البنائية	عدد المباني/ هكتار
		الكثافة السكنية الصافية	عدد السكان/ مساحة سكنية صافية
		نسبة التغطية البنائية	مساحة بناء الأرض ÷ مساحة القطعة(%)
	التنوع	تنوع استعمالات الارض	عدد اصناف الاستعمالات داخل المجاورة
		تنوع الانشطة الحضرية	عدد الانشطة المختلفة اليومية
		مؤشر الاستعمال المختلط	قيمة المؤشر بين (0-1)
		مؤشر شانون	قيمة المؤشر (H)
			نسب كل استعمال



(أ) (ب)

شكل (8)(أ) مفهوم الحي المتكامل في خطة بورتلاند (ب) مؤشر الحي لمدة 20 دقيقة المصدر:[50] ص11

أما على صعيد النقل، فقد أولت الخطة اهمية خاصة لتعدد وسائل التنقل، ولا سيما التنقل النشط، بوصفه أداة لتعزيز الصحة العامة وتحسين الأداء الإيكولوجي الحضري، من خلال اعتماد نظام هرمي يربط بين الممرات المدنية والممرات الخضراء المهيأة للمشاة وراكبي الدراجات. وأسهم هذا التوجه في تعزيز الترابط المكاني، وتسهيل الوصول إلى الخدمات، وتقليل الاعتماد على السيارات الخاصة، بما انعكس إيجابًا على خفض الانبعاثات الكربونية وتحسين جودة الهواء، وأظهر مرونة واضحة خلال جائحة كوفيد-19 عبر تحويل أجزاء من الطرق إلى ممرات خضراء مؤقتة.[50] كما ركزت الخطة على تعزيز الأمن الغذائي المحلي من خلال ضمان وصول السكان إلى الأغذية الطازجة داخل أحيائهم، عبر مبادرات مثل البيع بالتجزئة الصحي والحدائق المجتمعية، إلى جانب التأكيد على إشراك المجتمع في جميع مراحل التخطيط والتنفيذ بما يعزز العدالة والشرعية الاجتماعية.[56] وبصورة عامة، تُبرز تجربة بورتلاند مفهوم المجاورات المتكاملة كتطبيق متقدم لمفهوم مجاورة 15 دقيقة، قائم على القرب الوظيفي والتنوع والترابط الشبكي، مع بقاء تحديات مرتبطة بقدرة المدن متوسطة الكثافة على توطين جميع الأنشطة اليومية داخل المجاورة.

ويمكن الاستنتاج أن تجربة بورتلاند لا تسعى إلى إعادة تعريف مجاورة الـ20 دقيقة من خلال توسيع نطاق التغطية الخدمية أو فرض مؤشرات كمية صارمة، بقدر ما

وبناءً على ذلك نستنتج ان تجربة شنغهاي تقوم أساسًا على منهجية كمية صارمة لا تكتفي بتوفير الخدمات بل تعتمد على مؤشرات أداء دقيقة ونسبًا واضحة لتغطية الخدمات، مما يحوّل المفهوم من رؤية حضرية عامة إلى أداة إدارية قابلة للقياس وتقييم كفاءة توزيع الموارد ويتمتع بمرونة في التعامل مع المناطق الجديدة والمبنية، بما يجعل هذا المفهوم أكثر ملاءمة للسياقات الحضرية عالية الكثافة والمعقدة.

3.4 خطة بورتلاند

في عام 2012، أطلقت مدينة بورتلاند خطة حضرية شاملة عُرفت بخطة بورتلاند جاءت ثمرة شراكة موسعة بين القطاعين العام والخاص، إلى جانب منظمات المجتمع المدني، انظر شكل (8)، وقد استندت الخطة إلى مجموعة من الأولويات الحضرية الجوهرية شملت الازدهار الاقتصادي، التعليم، الصحة، والعدالة الاجتماعية، مع صياغة استراتيجيات قصيرة وطويلة الأجل لتوجيه التنمية الحضرية حتى عام 2035. اعتمدت الخطة هيكلاً استراتيجيًا متكاملاً يقوم على استراتيجيات رئيسة شكلت مدن صحية مترابطة. ويوصف مجاورة الـ20 دقيقة بأنها مجاورة متكاملة الذي يتمكن سكانها من الوصول الآمن والمريح إلى احتياجاتهم اليومية الأساسية، مثل السكن المتنوع، والخدمات التجارية، والمدارس، والمساحات العامة المفتوحة، والمرافق الترفيهية، والخدمات المدنية، من خلال التنقل سيرًا على الأقدام أو باستخدام وسائل النقل النشط اعتمادًا على التصميم بالمقياس الإنساني، بما يعزز جودة الحياة ويحد من الاعتماد على المركبات الخاصة.[50] ولتقييم تحقق هذا المفهوم، طوّرت المدينة مؤشر حي العشرين دقيقة لقياس سهولة الوصول إلى الخدمات والمرافق، مع استهداف تمكين الغالبية العظمى من السكان من تلبية احتياجاتهم اليومية خلال مدة زمنية قصيرة. ويأخذ هذا المؤشر في الحسبان الخصائص الطبوغرافية والجغرافية التي قد تعيق حركة المشاة، مثل الأنهار والانحدارات والطرق السريعة، فضلاً عن جودة بيئة المشي من حيث توفر الأرصفة، وتنوع المسارات، وسهولة الوصول إلى النقل العام، والقرب من مراكز الخدمات والأنشطة الحضرية.[55] كما تؤدي المراكز المجتمعية دورًا محوريًا في دعم الاكتفاء الذاتي للأحياء، وتعزيز التنمية الاقتصادية المحلية، وتوفير السكن الميسور، بما يسهم في تقليل الحاجة إلى التنقل الطويل وتحقيق قدر أعلى من العدالة المكانية داخل المدينة

دائرة المعيشة كوحدة مكانية لتوزيع الموارد بكفاءة، وتعزيز هوية المجتمع المحلي، وتحسين جودة البيئة المعيشية. ينتمي مفهوم دائرة الحياة إلى سياق شرق آسيا، وقد ازداد الاهتمام به في الصين نتيجة التفاوت الإقليمي والحضري، وتركّز الخدمات اليومية في المدن الكبرى. وتمثل الدائرة نطاقًا مكانيًا يخدم مجموعة سكانية محددة، ويُفترض أن تضم أنشطة السكن والعمل والتعليم والتسوق والرعاية الصحية والترفيه ضمن نطاق المشي. يتميّز المفهوم الصيني بأخذه في الاعتبار تجارب حياتية متعددة لفئات مختلفة، مثل كبار السن والأطفال وسكان سنّ العمل.[54] وقد اكتسب المفهوم زخمًا واسعًا، وجرى تطبيقه في مدن صينية كبرى مثل بكين. وتُعرَّف دائرة الحياة المجتمعية كوحدة تخطيطية أولية بمساحة تتراوح بين 3 و5 كم²، وبعدد سكان من 50 إلى 100 ألف نسمة، وتعتمد على مقياس المشي لمدة 15 دقيقة. ويرتكز الإطار المفاهيمي على التنوع لتلبية احتياجات المجموعات المختلفة، وسهولة الوصول لتنظيم الأنشطة بكفاءة، والدقة لضمان توافق المرافق مع سلوك السكان الفعلي. أما التنفيذ فيعتمد على منظومة إرشادية متعددة المستويات تشمل التخطيط والإدارة والممارسة، مع تحوّل في المحتوى من التركيز على البنية التحتية إلى الاستجابة للاحتياجات الواقعية للمواطنين. ويُعنى التخطيط بتحسين جودة الخدمات، وإعادة توزيع المرافق، ومعالجة الفجوات في المناطق المبنية وغير المبنية، بما يعزز تكامل الأنشطة والخدمات داخل المجاورة.

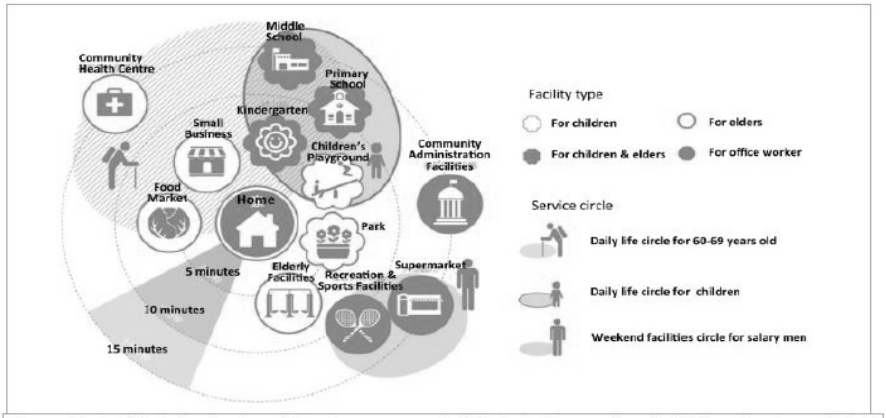


Figure 11. Facilities' distribution of 15-minute community life circle. Source: Shanghai Urban Planning and Land Resources Administration Bureau (2016).

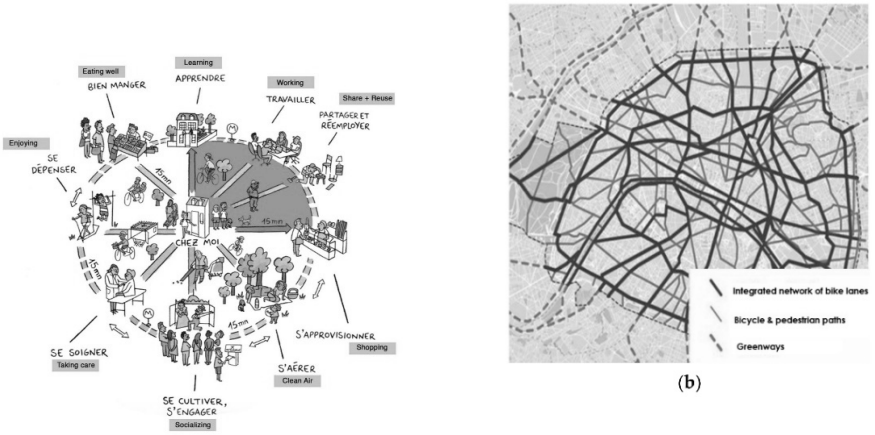
شكل (7) توزيع المرافق دائرة المعيشة لمدة 15 دقيقة المصدر [53] ص34

سكنية بديلة مثل السكن المشترك، بهدف الحد من الإقصاء الاجتماعي والحفاظ على التنوع السكاني داخل المدينة. كما تتضمن الاستراتيجية توجهًا قويًا نحو التشجير الحضري، وتعزيز التنوع البيولوجي، وخفض استهلاك الطاقة، ودعم ممارسات الاقتصاد الدائري، بما يعزز الأداء البيئي للأحياء.[52] وتُعد المشاركة المجتمعية عنصرًا محوريًا في استراتيجية باريس، حيث خُصص جزء من الميزانية لآليات التخطيط التشاركي، مثل الميزانية التشاركية، لإشراك السكان في صنع القرار وتحديد أولويات أحيائهم. وبصورة عامة، تُظهر تجربة باريس مفهوما متقدمًا لتطبيق مفهوم مدينة الربع ساعة في سياق حضري عالي الكثافة، قائم على تعزيز القرب الوظيفي وتنوع الاستخدامات والعدالة البيئية. ومع ذلك، تبرز مخاوف تتعلق بإمكانية تعميق الاستقطاب الاجتماعي، في حال اقتصرت فوائد التحول الحضري على المناطق المركزية ذات الدخل المرتفع، مما يستدعي مواءمة هذا المفهوم مع سياسات إسكان ونقل إقليمي أكثر شمولاً لضمان عدالة الوصول على مستوى المدينة الكبرى.

ونستنتج ان نجاح مدينة باريس لم يكن بفضل التكنولوجيا، بل بفضل الكثافة الحضرية العالية وتنوع الاستعمالات، وتعزيز التنقل النشط مما أسهم في تحسين الأداء البيئي والاجتماعي للأحياء، إلا أن استدامة نجاحه تبقى مرهونة بمواءمته مع سياسات إسكان شاملة تمنع الاستقطاب الاجتماعي لضمان عدم تحول المجاورات المخدومة إلى مناطق للأثرياء فقط مما تضمن عدالة الوصول إلى الخدمات على مستوى المدينة ككل.

2.4 شنغهاي

تبنّت شنغهاي مفهوم مجاورة 15 دقيقة ضمن المخطط الرئيسي لشنغهاي 2035 واطلقت عليه دائرة حياة المجتمع لمدة 15 دقيقة كأحد المسارات الرئيسة للتجديد الحضري وإعادة التطوير، ضمن تجربة صينية مستقلة عن مفهوم كارلوس مورينو في باريس، انظر شكل (7) ، وقد أُدرج هذا التوجه رسميًا في الخطة الرئيسة للمدينة عام 2014، ثم تعزّز في تحديث عام 2018 بأهداف واضحة حتى عام 2035. تهدف شنغهاي إلى تمكين 99٪ من السكان من الوصول إلى مرافق الخدمات العامة، و90٪ من الوصول إلى المساحات العامة المفتوحة خلال 15 دقيقة سيرًا على الأقدام.[53] ويُستخدم مفهوم



(ب) (أ)

شكل (6)(أ) شبكة الدراجات والمشاة المتكاملة لمدينة باريس (ب) الوظائف الاجتماعية والحضرية التي تغطيها مدينة15 دقيقة المصدر[50]

وبالنظر إلى الكثافة الحضرية المرتفعة لباريس، تركز الاستراتيجية على تعزيز تنوع الاستخدامات ودمج الوظائف الحضرية ضمن نسيج واحد، إلى جانب إعادة توظيف الفضاءات العامة وشبه العامة متعددة الاستخدامات لدعم الأنشطة الاجتماعية والترفيهية. وتُولي الاستراتيجية أهمية كبيرة لتوطين الخدمات والوظائف، من خلال دعم الاقتصاد المحلي، وتشجيع الإنتاج والاستهلاك المحلي، وتعزيز ريادة الأعمال الصغيرة، وربط فرص العمل بالمجاورات السكنية. كما تهدف إلى توسيع أنماط العمل المحلية والعمل عن بُعد، بما يقلل الحاجة إلى التنقل اليومي لمسافات طويلة.[50] وعلى صعيد النقل، تعطي الاستراتيجية أولوية واضحة للتنقل النشط، عبر تحويل مساحات واسعة من المدينة إلى مناطق مخصصة للمشاة، وتطوير شبكة متكاملة من مسارات الدراجات والممرات الخضراء، والحد من استخدام السيارات الخاصة، وهو ما أسهم في تغيير ثقافة التنقل الحضري وتعزيز الصحة العامة وتقليل الانبعاثات.وفيما يخص السكن، تُعدّ مسألة توفير سكن شامل وميسور أحد أبرز التحديات أمام تطبيق مفهوم مدينة الربع ساعة، في ظل ارتفاع أسعار العقارات واستمرار الضغوط الديموغرافية. وتسعى الاستراتيجية إلى مواجهة ذلك من خلال زيادة نسبة السكن العام، وتشجيع إعادة استخدام المباني، وتحفيز أنماط

يتلخص من الادبيات السابقة رغم الأهمية الكبيرة للدراسات المذكورة أعلاه، إلا أنها تباينت في تركيزها؛ فبينما أسس Moreno (2024) للمفهوم فلسفياً، وركز Khavarian-Garmsir (2023) على نقد الاستدامة، وتناول Caselli (2022) الجانب القياسي في مدن أوروبية منظمة؛ وركز(Murgante 2023) على الجوانب القياسية باستخدام أدوات التحليل المكاني وتناول (Zaheer 2022) ازمة المناخ إلا أنها ركزت بشكل رئيس على سياقات غربية أو على آليات القياس الكمي لذلك ظهرت فجوة معرفية تتعلق بـ **عدم وجود إطار نظري شامل يدمج هذه المؤشرات (الكمية والنوعية) ضمن سياق تخطيطي مرن يمكن تطبيقه على المدن التي تعاني من مشاكل هيكلية. لذا، يأتي هذا البحث لاستخلاص إطار نظري متكامل للمؤشرات والمعايير التخطيطية لمجاورة 15 دقيقة.**

4. التطبيقات العالمية

1.4 باريس-فرنسا

تعدّ استراتيجية باريس للجميع إطارًا حضريًا تحويليًا يهدف إلى إعادة بناء المدينة على أساس محلي منخفض الكربون، بالارتكاز على مشاركة المجتمع والعدالة البيئية، وبالانسجام مع خطة باريس للعمل المناخي. وقد اكتسبت هذه الاستراتيجية زخمًا متجددًا بعد جائحة كوفيد-19 من خلال إعادة طرح مفهوم مدينة الربع ساعة كأداة لتعزيز مرونة المجاورات وتحسين جودة الحياة اليومية للسكان. انظر شكل (6)، وتسعى الرؤية إلى تحويل باريس إلى مدينة صحية وشاملة، عبر الحد من هيمنة السيارات، واستعادة الفضاءات العامة، وزيادة الغطاء الأخضر، وتعزيز المشي والتنقل النشط. يندرج مفهوم مدينة الربع ساعة ضمن محور التقارب الشديد، ويُعرّف بوصفه مفهوما حضريًا يُمكّن السكان من تلبية معظم احتياجاتهم اليومية الأساسية خلال 15 دقيقة سيرًا على الأقدام أو بالدراجة من أماكن سكنهم.[51] وترتكز فلسفة هذا المفهوم على أربعة مبادئ رئيسة هي القرب، والتنوع، والكثافة، والانتشار، مع التركيز على توطين ست وظائف حضرية أساسية تشمل السكن، والعمل، والتسوق، والرعاية الصحية، والتعليم، والترفيه.

العالمية لتبني سياسات خفض عميق للانبعاثات، يطرح البحث مفهوم مجاورة 15 دقيقة كنموذج تخطيط حضري شامل ومستدام، ويتمحور هذا المفهوم حول هدف محدد يتمثل في ضمان الوصول إلى جميع المرافق والخدمات الأساسية (العمل، التعليم، التسوق، الرعاية الصحية، الترفيه) في غضون 15 دقيقة سيرًا على الأقدام أو بالدراجة. ويهدف البحث إلى إثبات قدرة هذا النموذج على قيادة التحولات الحضرية من خلال أربعة أبعاد مترابطة القرب، الكثافة، التنوع، الرقمنة وتكمن القيمة الأكاديمية للبحث في ربط هذا النموذج بتحقيق صافي صفري للانبعاثات، خاصة في قطاع النقل، الذي يساهم بـ 14% من الانبعاثات السنوية. واكدت الدراسة على ضرورة تبني هذا المفهوم وفقاً للمقياس البشري لضمان تحقيق المساواة الاقتصادية والشمولية، خاصة في مدن الجنوب العالمي، عبر تخصيص النموذج وتفعيل آليات تمويل محلية مبتكرة كالمناطق الاقتصادية الخاصة لتجنب التهميش وتوفير جودة حياة عالية للمقيمين واكدت النتائج ان مفهوم مجاورة 15 دقيقة حل فعال لخفض الانبعاثات وتقليل الحاجة للسفر بالمركبات ويجب أن يسترشد التطبيق بمبادئ المقياس الانساني والتخطيط التشاركي لضمان الشمولية وعدم عزل السكان.

Developing a 15-minute city: Evaluating urban quality using configurational analysis. The case study of Terni and Matera, Italy
Beniamino Murgante, Raffaela Valluzzi, Alfonso Annunziata(2023)

تهدف هذه الدراسة إلى تقييم جدوى تطبيق مفهوم مجاورة 15 دقيقة في مدينتي تيرني وماتيرا الإيطاليتين، وذلك من خلال تطوير وتطبيق منهجية كمية تدمج بين أدوات تحليل التركيب الفضائي (Space Syntax) والتحليل الهيكلي (Place Syntax) ضمن بيئة نظم المعلومات الجغرافية (GIS). وسعت الدراسة إلى بناء جملة مكانية وتحديد مؤشر هرمي لقياس مدى تغطية الخدمات الحضرية وتنوعها في آنٍ واحد. وكشفت النتائج عن تباين مكاني واضح؛ حيث تتمتع المناطق المركزية وشبه المركزية بمستويات وصول مثالية وكثافة سكانية مناسبة للخدمات، بينما تعاني الضواحي من نقص حاد في الخدمات وضعف في كفاءة التكوين المكاني، مما يؤكد أهمية التخطيط الهيكلي في تحقيق العدالة المكانية.

والاقتصادية والبيئية، وتحديات التنفيذ المحتملة وأظهرت النتائج أن هذا المفهوم يساهم بشكل كبير في تحقيق الاستدامة الثلاثية؛ فاجتماعيًا يعزز الإنصاف، والتماسك الاجتماعي، والصحة العامة، واقتصاديًا يقلل تكاليف النقل ويدعم الاقتصادات المحلية، وبيئيًا يقلل الاعتماد على السيارات وانبعاثات الغازات الدفيئة. وأظهرت النتائج أن المفهوم يساهم بفاعلية في تقليل الاعتماد على السيارات وخفض الانبعاثات. وأظهرت النتائج أن المفهوم يساهم بفاعلية في تقليل الاعتماد على السيارات وخفض الانبعاثات. ومع ذلك كشف عن تحديات جوهرية تتعلق بإهمال بعض الجوانب مثل التنوع البيولوجي، كفاءة الطاقة، واحتياجات الفئات الاجتماعية الهشة.

Exploring the 15-minute neighborhoods: An evaluation based on the walkability performance to public facilities.

Caselli, B., Carra, M., Rossetti, S., & Zazzi, M. (2022).

تتناول هذه الدراسة تقييم مفهوم مجاورة 15 دقيقة الذي يمثل توجهاً حضرياً حديثاً يركز على أبعاد القرب والتنقل النشط في المناطق الحضرية، وذلك كبديل لمفهوم النقل المعتمد على المركبات القديم. ويتمحور الهدف الرئيس للدراسة حول تطوير وتطبيق منهجية تحليلية قائمة على نظم المعلومات الجغرافية (GIS) لقياس أداء إمكانية وصول المشاة إلى خدمات الحي، مع اعتبار عامل الوقت (15 دقيقة) عنصراً حاسماً في التقييم. بينت النتائج أن تحقيق مجاورة الـ 15 دقيقة لا يعتمد فقط على المسافة، بل على التوزيع المكاني الدقيق للسكان والخدمات. وأكدت أن الأحياء التي تتمتع بخلط استعمالات وكثافة سكانية مناسبة حققت مستويات وصول عالية (أكثر من 90% من السكان)، بينما عانت المناطق الأخرى من فجوات خدمية واضحة.

The "15-Minute City" concept can shape a net-zero urban future

Zaheer Allam, Simon Elias Bibri, Didier Chabaud, & Carlos Moreno(2022)

تتناول هذه الدراسة أزمة المناخ وتعاظم انبعاثات المدن، حيث تساهم المراكز الحضرية بأكثر من 60% من انبعاثات الغازات الدفيئة العالمية وفي سياق الدعوات

3. الادبيات السابقة

The 15-Minute City: redesigning urban life with proximity to services

Carlos Moreno (2024)

تُعد هذه الدراسة المرجع الأساسي الأحدث للمفهوم، حيث قدم كارلوس مورينو رؤية نقدية لنماذج التخطيط التقليدية التي فصلت بين السكن والعمل. طرحت الدراسة مفهوم العمران الزمني الذي يهدف إلى إعادة تعريف العلاقة بين المكان والزمان وهدفت إلى تقديم إطار نظري شامل لتحويل المدن إلى أقاليم كاملة يستطيع فيها السكان تلبية احتياجاتهم الستة الأساسية (السكن، العمل، التموين، الرعاية، التعليم، والترفيه) ضمن نطاق زمني قصير، بهدف تعزيز الاستدامة البيئية والعدالة الاجتماعية. وتوصلت الدراسة إلى أن نجاح مفهوم مجاورة 15 دقيقة يعتمد على أربعة مرتكزات هي: القرب، الكثافة، التنوع، والرقمنة. كما أثبتت أن التعدد الوظيفي للمباني هو المفتاح لتعظيم كفاءة الموارد الحضرية المحدودة.

The 15-minute city: Urban planning and design efforts toward creating sustainable neighborhoods

Amir Reza Khavarian-Garmsir, Ayyoob Sharifi, Ali Sadegh(2023)

تستكشف هذه الدراسة المفاهيمية مبادئ وإسهامات وتحديات تنفيذ مفهوم «المدينة ذات الـ 15 دقيقة»، الذي ظهر كإطار تخطيط حضري مستدام، اكتسب زخمًا كبيرًا بعد جائحة كوفيد-19. تعتمد الرؤية على سبعة مبادئ أساسية هي: القرب، الكثافة، التنوع، الرقمنة، التصميم بمقياس انساني، المرونة، والاتصال، بهدف إنشاء أحياء مكتفية ذاتيًا تتيح للمقيمين الوصول إلى الوظائف الأساسية (العيش، العمل، الرعاية، التعلم، التموين، الترفيه) خلال 15 دقيقة سيرًا أو بالدراجة. وهدفت الدراسة الى استكشاف المبادئ الأساسية لمجاورة 15 دقيقة، وتقييم مساهماتها في الاستدامة الاجتماعية

شكل (5) التسلسل التاريخي للمجاورة الحضرية / المصدر16 ص4

[44] ويترتب على تحسين الترابط وإمكانية الوصول آثار متباينة؛ حيث تسهم أنظمة النقل الفعّالة في دعم النشاط الاقتصادي ورفع قيمة الأراضي في بعض المواقع، ومن جهة أخرى قد تُنتج آثارًا سلبية مثل الضوضاء والتلوث، وهنا يبرز دور التصميم الحضري في تقليل من هذه التكاليف عن طريق الدمج بين أنماط التنقل النشطة، كالمشي وركوب الدراجات، ووسائل النقل العام، بما يعزز من كفاءتها ويحد من الاعتماد على السيارات الخاصة.[1] ويسهم الترابط الحضري أيضًا في منع تشكّل المجاورات المعزولة، من خلال تعزيز قدرة النظام الحضري على الصمود والتعافي عبر ربط المجاورات[48] التي يمكن الوصول إليها خلال 15 دقيقة بشبكات النقل العام بما يرفع كفاءتها التشغيلية ويقلل الاعتماد على السيارة في الوصول إلى الوجهات النهائية[49] فضلاً عن دوره في تخفيف الازدحام المروري عبر تقليل الضغط على الشوارع الرئيسة، وتحقيق مسارات أكثر مباشرة وكفاءة في التنقل.[50]

وبناءً على ذلك، نستنتج أن تطور مفاهيم تخطيط المجاورة الحضرية يعكس تحولاً تدريجيًا في فهم المدينة من كونها منظومة وظيفية تُدار بالكفاءة والسرعة، إلى كونها نظامًا حضريًا حيًا يتمحور حول الإنسان وجودة حياته اليومية. فقد أظهرت المقارنة التاريخية انظر الشكل (5)، أن النماذج التخطيطية السابقة عالجت المدينة بصورة جزئية؛ إذ ركزت الحداثة على الكفاءة الوظيفية والسرعة، غالبًا على حساب البعد الإنساني، في حين انشغلت نماذج المدينة الذكية بالحلول الرقمية دون إيلاء اهتمام كافٍ لمسألة القرب الفيزيائي. في المقابل، يقدّم مفهوم مجاورة 15 دقيقة رؤية شمولية تسعى إلى معالجة ظاهرة التجزئة الحضرية من خلال دمج البعد الزمني بالبعد المكاني. كما يُظهر أن الكثافة والتنوع الوظيفي والرقمنة لا تمثل أهدافًا بحد ذاتها، بل أدوات مترابطة تسهم في تعزيز مرونة المجاورة الحضرية وقدرتها على الاستجابة للتغيرات الاجتماعية والبيئية. وعليه، فإن التحول نحو مفهوم مجاورة 15 دقيقة لا يُعد خيارًا تخطيطيًا ثانويًا أو توجهًا نظريًا معاصراً فحسب، بل يمثل استجابة ضرورية فرضتها تحديات التغير المناخي، والحاجة إلى مدن أكثر استدامة، بما يحقق بيئة حضرية أكثر صحة وعدالة وقدرة على الصمود.

8.4.2 المقياس الإنساني

أدى التوسع الحضري في الضواحي، الذي عززته التوجهات الحداثية بعد الحرب العالمية الثانية، إلى نشوء أنماط حضرية تعتمد على السيارة بوصفها وسيلة التنقل الأساسية، ما نتج عنه تمدد عمراني واسع، وارتفاع تكاليف النقل، والعزلة الاجتماعية، إلى جانب تفاقم التلوث البيئي. وقد أكدت Jane Jacobs أن المدينة تُبنى بالناس ولأجلهم، وأن الخطط الحضرية يجب أن تتناسب معهم[15] ونتيجة لذلك، سعت مجاورة 15 دقيقة إلى إعادة تشكيل المدينة على أسس إنسانية،[45] عبر تعزيز المشي وركوب الدراجات، وتقليل استهلاك الطاقة، والحد من الازدحام المروري والتلوث، مع خفض كلف الصيانة ويمكن لجميع الناس التنقل بسهولة إذ يمنح هذا المفهوم أولوية واضحة للمواطنين في تصميم الشوارع والطرق، من خلال تطوير شبكات آمنة وجاذبة للمشاة وراكبي الدراجات، وتشجيع أنماط التنقل النشط[46] وتتجلى فوائد هذا البعد في تعزيز اندماج المباني مع محيطها الحضري، بما يحقق تجانسًا بصريًا ووظيفيًا ويُحسّن جودة الفضاءات العامة. كما يسهم في تسهيل الوصول وتشجيع المشي وركوب الدراجات، الأمر الذي يقلل الاعتماد على السيارات ويحد من التلوث. كما تسهم المباني المصممة وفقًا للمقياس الإنساني في تقليل المسافات بين الأماكن والأفراد مما يميل الناس إلى رؤية بعضهم البعض بشكل أكبر، ويساعد هذا التقارب المكاني على تعزيز التفاعل الاجتماعي والشعور بالانتماء المجتمعي. وعلى الصعيد الاقتصادي، إذ تسهم المساحات العامة النشطة في جذب الزوار وتنشيط الأنشطة التجارية المحلية، بما يعزز فرص العمل ويقوي القدرة التنافسية للمدينة[44]

9.4.2 الترابط

يُعرَّف الترابط على أنه مقياس لدرجة تكامل نظام الشوارع من حيث تعدد المسارات التي تخدم نقاط الانطلاق والوصول ذاتها، إذ يعكس ارتفاع مستوى الترابط بوجود طرق عديدة للانتقال بين أي نقطتين داخل الشبكة الحضرية.[47] ويشير الترابط، بمعناه الأوسع، إلى مستوى تكامل شبكات النقل المختلفة، بما في ذلك الشوارع، ومسارات المشي وركوب الدراجات، إضافة إلى شبكات الخدمات والبنية التحتية. وتشجع المستويات العالية من الترابط على تعزيز سهولة الوصول داخل المجاورة، بما يدعم كفاءة الحركة والتنقل.

اللامركزية في انشاء مراكز فرعية متعددة تُقدم الخدمات الضرورية لسكان المدن. كما تشير النمطية إلى قابلية النظام الحضري للتقسيم إلى وحدات أو مجتمعات مترابطة داخليًا ضمن إطار كلي متكامل، الأمر الذي يعزز من مرونة المدن وقدرتها على امتصاص الصدمات وتقليل آثار الاضطرابات واسعة النطاق عند التعرض للأزمات [1]ومن المتوقع أن تسهم مفهوم مجاورة 15 دقيقة في تعزيز لامركزية الخدمات الحضرية من خلال إنشاء مراكز فرعية متعددة، قادرة على تلبية الاحتياجات الأساسية للسكان ضمن نطاقات قريبة ومتكاملة.

7.4.2 المرونة

تُعرَّف المرونة بأنها قدرة الفضاء الحضري على الاستجابة للاحتياجات اليومية المتغيرة للمجتمع والتكيّف معها. وتتحقق هذه المرونة من خلال إتاحة مرونة مكانية تسمح بإعادة تنظيم الفضاءات الحضرية وتكييفها لإنتاج حلول ووظائف جديدة[40] وقد أبرزت جائحة كوفيد-19 الحاجة الملحّة إلى مزيد من المرونة في أنماط العمل والمعيشة، حيث دفعت الظروف الاستثنائية الأفراد إلى إعادة تنظيم حياتهم اليومية وتوطين أنشطتهم ضمن نطاقات أقرب [41] ولتحقيق أقصى قدر من الكفاءة الحضرية، تدعوا مجاورة 15 دقيقة على أهمية تحويل المساحات العامة أحادية الوظيفة إلى مساحات متعددة الأغراض [22] [42] بما يسمح باستخدامها في أوقات مختلفة من أيام الأسبوع وتصبح الفضاءات العامة وشبه العامة عناصر نشطة ومستخدمة باستمرار ضمن الحياة اليومية للسكان[43] ويساعد هذا التوجه على تعزيز اللامركزية من خلال توزيع الأنشطة والخدمات على نطاق أوسع، مع تحسين كفاءة استخدامها، الأمر الذي ينعكس إيجابًا على قدرة المجاورات الحضرية ووظيفتها على استيعاب الصدمات المفاجئة الناجمة عن أزمات مثل جائحة كوفيد -19[44] وانسجامًا مع ذلك، يدعو مفهوم مجاورة 15 دقيقة إلى إعادة توظيف المساحات العامة ذات الاستخدام الواحد وتحويلها إلى فضاءات متعددة الاستعمالات.

مثل التنقل كخدمة[36]، فضلاً عن دور تقنيات الاتصالات في تحفيز الاقتصاد المحلي عبر خلق وظائف جديدة ودعم ودعم الشركات المحلية لتعزيز قدرتها التنافسية[37] ولقد أظهرت الرقمنة، وخاصةً خلال فترة جائحة كوفيد-19، دوراً محورياً في تمكين الافراد من العمل عن بُعد والتواصل افتراضيًا، وهو ماساهم في تقليل التواصل الاجتماعي وتقليل الحاجة إلى التنقل بين المنزل و اماكن العمل. كما اثبتت فعاليتها في توفير خدمات متنوعة مثل التسوق الالكتروني والمعاملات غير النقدية، والاتصالات والتفاعلات الافتراضية داخل المنازل. وتُعدّ الرقمنة عنصرا اساسيا في دعم الاستهلاك الامثل للموارد.[38] و تُحقق العديد من الفوائد التي تُعزز الكفاءة وتُحسّن النتائج فتعمل على تحسين عملية صنع القرار وتُزوّد الأدوات والبرامج الرقمية المخططين ببيانات دقيقة وفورية مما يُمكّنهم من اتخاذ قرارات مدروسة.

5.4.2 التكيف

تواجه المدن المعاصرة مجموعة واسعة من التحديات تجعل المستقبل غامضًا ومتغيرًا. لذا، يُعدّ التكيف مع هذه الظروف المتغيرة أمرًا بالغ الأهمية. اذ تكون المدن القادرة على التكيف أقل تعرضًا للمخاطر، لامتلاكها قابلية التعلم من الأزمات والتعامل مع الظروف المتغيرة كما تتميز هذه المدن بأدراكها بهشاشة مكوناتها الأساسية،[1] وسعيها إلى تطوير المعرفة، والقدرة على ترتيب الأولويات بما يتيح بناء نظام حضري آمن أو قابل للتعافي. ولا تقتصر مجاورة 15 دقيقة على معالجة تحديات التوسع الحضري القديمة، بل تركز بصورة متزايدة على قضايا تغيّر المناخ وندرة الموارد التي برزت بوضوح في العقود الأخيرة.[39]

6.4.2 النمطية

يعد التوزيع اللامركزي أحد المظاهر الرئيسة للنمطية الحضرية، حيث تتطور المدن وفق نمط متعدد المراكز تتوزع فيها المرافق والخدمات في عدة مراكز. ويسهم هذا التنظيم النمطي للمدن أن يُقلل من التفاوتات المكانية في عبر تحسين فرص الوصول إلى العمل والخدمات بالقرب من أماكن السكن، فضلاً عن مبادرة المدن النمطية

الأنشطة السكنية والتجارية والخدمية ضمن نطاق يسهل الوصول إليه، مما يرفع كفاءة استخدام الأرض ويعزز حيوية الشوارع والعقد الحضرية.[30]

2. 4. 3 التنوع

يُعدّ التنوع استراتيجية رئيسة للتنمية الحضرية تعالج أوجه عدم المساواة الاجتماعية والمكانية كما يُعدّ هذا المبدأ بالغ الأهمية لتحقيق الاستدامة الاقتصادية على المدى الطويل، وزيادة الابتكار والإبداع[31]. ويهدف إلى إنشاء أحياء تضمّ ثقافات وأعراقًا وفئات اجتماعية واقتصادية متنوعة. وتسعى مجاورة 15 دقيقة إلى إنشاء مجتمع متنوع من خلال بُعدين رئيسين[22] هي: التنوع بأستعمالات الارض بوصفها شرطًا جوهريًا لتأمين مزيج متوازن من المكوّنات السكنية والتجارية والترفيهية، والتنوع الثقافي والسكاني كخطوة نحو دمج التعددية الثقافية في التخطيط[32].فالمجاورات متعددة الاستخدامات تُعدّ عنصرًا محوريًا في استدامة النسيج الحضري النابض بالحياة اقتصاديًا، وضمان توفير مساكن كافية لجميع سكان المدن، وتعزيز مبادئ الشمولية،[33] وتشجيع الممارسات الحضرية المستدامة. وسعى كارلوس مورينو، مستلهمًا من جاكوبس، إلى تحويل مدينة الـ 15 دقيقة من هيكلها الحضري أحادي المركز إلى شكل متعدد المراكز. وانطلاقًا من فلسفة التخطيط الزمني الحضري، اعتقد أن توطين الوظائف يُحسّن جودة حياة الناس من خلال منعهم من إضاعة الوقت في التنقلات اليومية، وبالتالي خفض انبعاثات الكربون[34].

4.4.2 الرقمنة

نظراً للإمكانات التي تتيحها التقنيات الذكية كالذكاء الاصطناعي، والبيانات الضخمة، وإنترنت الأشياء في توفير بيانات لحظية وتعزيز المشاركة في صنع القرار، دفعت مورينو إلى إدراج الرقمنة كمكون رئيسي لمفهوم المجاورة التي تستغرق 15 دقيقة [35] فمن خلال توظيف الأدوات الرقمية مثل لوحات البيانات الذكية والخرائط التفاعلية، يمكن للحكومات المحلية تعميق تواصلها مع السكان وتقليص العزلة الاجتماعية داخل الأحياء[22]كما تدعم حلول المدن الذكية أنظمة النقل متعددة الوسائط ومفاهيم حديثة

1.4.2 القرب

يشمل هذا البُعد الجوانب الزمنية والمكانية، فهو يمثل مبدأ اساسي لحيوية المدن[15] ويوفر مزايا عديدة، منها تقليل الحاجة إلى البنية التحتية للفرد، وتحسين الوصول إلى الفرص التعليمية والاقتصادية والاجتماعية والثقافية، وتشجيع وسائل النقل النشطة، وشوارع أكثر أمانًا[26] ويُعد القرب استراتيجية لتحسين جودة حياة الناس من خلال قدرة السكان الوصول إلى الخدمات والمرافق الحضرية بسرعة، ويهدف إلى معالجة بعض التحديات الحديثة مثل زيادة الانبعاثات والزحف الحضري.[27] كما يعتبر هذا البعد بالغ الأهمية ليس فقط في تقليل الوقت الضائع في التنقل، بل أيضا في الحد من الآثار البيئية والاقتصادية المرتبطة به كما يسهم في تحديد المؤشرات الاجتماعية التي تؤثر على سكان المناطق الحضارية، لاسيما في ما يتعلق بتعزيز التفاعلات الاجتماعية. يتحقق مفهوم القرب في المدن من خلال تكثيف النسيج الحضري، وتنويع الأنشطة واستخدامات الأراضي، والتحول نحو وسائل النقل المستدامة مثل المشي، وركوب الدراجات، واستخدام النقل العام. [28] وقد اعتبر المخططون أن تحقيق القرب من الأنشطة اليومية يمثل عنصرًا أساسيًا في رفاهية الأفراد وتعزيز الاستدامة الحضرية. ويُميَّز عادةً بين نوعين: القرب الزمني، الذي يعتمد على النقل السريع ويهدف إلى تقليل مدة التنقل ويرتبط تقليديًا بالتخطيط الموجّه للسيارات، إذ يسهم في تسهيل الأنشطة اليومية وجعلها أكثر سلاسة. اما القرب الجغرافي، الذي يتحقق من خلال القرب المكاني وتكامل استخدامات الأراضي.[1].

2.4.2 الكثافة

تُعدّ الكثافة أحد الأبعاد الأساسية في فهم البنية المكانية والاقتصادية للمجاورة الحضرية. فهي تمثل استراتيجية حاسمة لضمان حصول السكان على احتياجاتهم اليومية بشكل كافٍ دون الاعتماد على السيارات الخاصة، وأن الكثافة السكانية تُشكل كتلة حرجة لتشغيل الخدمات والشركات المحلية على مسافات قريبة، فضلاً عن خلق فرص عمل محلية.[29] وتُعزز الكثافة السكانية العالية توطين الوظائف الحضرية، بينما تُقلل بشكل غير مباشر من إجمالي عدد الأميال المقطوعة بالمركبات وما يرتبط بها من انبعاثات كربونية ويتناغم مبدأ الكثافة العالية مع الاستعمالات المختلطة للأراضي من خلال تركيز

المفهوم زخمًا كأداة تخطيط نظرًا لبساطته وقابليته للقياس وتوافقه مع أهداف التنقل المستدام، فهو يشجع على نمط حضري إنساني يقلل الاعتماد على السيارات من خلال تشجيع التنمية متعددة الاستخدامات، وسهولة المشي، والتقارب ويعزز الحياة المحلية مما يدعم الاستدامة والتماسك الاجتماعي ويحسن جودة الحياة و يؤدي ذلك إلى خلق مدن أكثر ملاءمة للعيش.[24]

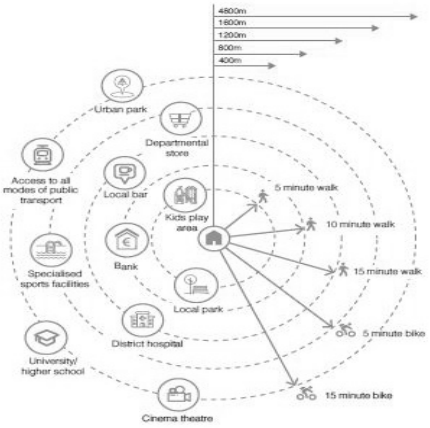


Figure 9 - An illustrative supposition of 15-minute city model
(Source: Author's interpretation of iterations of Gehl architects & Congress for New Urbanism)

شكل (4) نموذج توضيحي لمجاورة 15 دقيقة المصدر[25] رقم الصفحة33

4.2 الخصائص الاساسية لمجاورة 15 دقيقة

تتطلب مجاورة 15 دقيقة مجموعة من الخصائص الجوهرية، تشمل)القرب، والكثافة، والتنوع، والرقمنة، والمقياس الإنساني، والمرونة، والتكيف، والترابط، والنمطية(وذلك لإعادة هيكلة المجاورات الحضرية بما يحقق الهدف النهائي المتمثل في بناء مجتمعات مترابطة. وتمثل هذه الخصائص الأساس المفاهيمي لمجاورة 15 دقيقة، حيث تُظهر المدن التي تلتزم بها إمكانات اقتصادية أعلى مقارنة بالمدن التقليدية التي تتسم بتوجهها نحو الاختناقات المرورية، والتي أُنشئت أساسًا لتسهيل حركة المركبات بدلًا من تعزيز التفاعل البشري. ويمكن دعم تطبيق هذه الخصائص من خلال توظيف التقنيات المتقدمة المعاصرة، التي أسهمت في تعزيز أنماط جديدة من الممارسات الحضرية، مثل التواصل الافتراضي، والعمل من المنزل، وإنجاز المعاملات الإلكترونية [1].[24] [22]

الحكومات قيودًا غير مسبوقة على الحركة وإجراءات الحجر الصحي. [20] أثّر هذا الوضع الطبيعي الجديد على العالم بطرقٍ عديدة، وأبرزت الجائحة عدم قدرة المدن الحديثة على مواكبة الأوبئة مما دفع منظمة الصحة العالمية إلى إعلان جائحة عالمية في بداية عام 2020. وقد كشفت هذه الجائحة عن هشاشة المدن في وضعها الراهن، والحاجة إلى إعادة نظر جذرية، حيث يجب تصميم تدابير مبتكرة خصيصًا لضمان قدرة سكان المدن على التكيف ومواصلة أنشطتهم الأساسية، بما فيها الأنشطة الثقافية، لضمان بقاء المدن مرنة وصالحة للعيش.[21] أدت الحاجة إلى هذا التفكير الجذري، إلى جانب القضايا الاجتماعية والاقتصادية الملحة، إلى تبني المدن لآليات تخطيط حضري تضمن الحفاظ على جودة الحياة بعد تجاوز الآثار السلبية التي قد يخلفها الفيروس في المدن. وفي هذا السياق، يرى صناع السياسات ضرورة توفير خدمات قريبة من السكان من خلال مراجعة السياسات الحضرية، ولا سيما النقل، باعتباره حلقة وصل لحل المشكلات الأخرى. ذكّرت الجائحة المخططين بأن المدن بحاجة إلى شكل حضري أكثر استدامة، وأنه يجب إعادة النظر في مبادئ التخطيط والتصميم الحضري وبناءً على ذلك، تم استخدام مفهوم مجاورة 15 دقيقة، الذي قدمه كارلوس مورينو في عام 2016، استجابةً لتحديات أزمة المناخ لمكافحة انبعاثات غازات الاحتباس الحراري وسعياً لتقليل الاعتماد المفرط على السيارات وتعزيز بيئات حضرية أكثر استدامةً. وكاستراتيجية للتعافي بعد جائحة كورونا.[22]

3.2 مجاورة 15 دقيقة

طُرِح هذا المفهوم لأول مرة من قِبل Carlos Moreno عام 2016، الذي يدعو إلى تخطيط حضري يُمكّن السكان المحليين من الوصول إلى جميع احتياجاتهم الأساسية في مسافات لا تستغرق أكثر من 15 دقيقة سيرًا على الأقدام أو بالدراجة.[23] انظر شكل(4)، ويستند هذا المفهوم إلى فكرة العمران الزمني الذي يُبين أن جودة الحياة الحضرية تتناسب عكسيًا مع الوقت المُستثمر في التنقل وخاصةً باستخدام السيارات. ويهدف إلى إنشاء أحياء مكتفية ذاتيًا ذات وظائف أساسية تشمل السكن والعمل والتجارة والرعاية الصحية والتعليم والترفيه من خلال لامركزية الوظائف والخدمات الحضرية وتتمثل فكرتها الأساسية في تصميم المدن أو إعادة تصميمها كأحياء كثيفة متصلة ومختلطة اجتماعيًا ووظيفيًا[22] واكتسب هذا

والمدينة الايكولوجية، والمدينة الرقمية، والمدينة الخضراء، والمدينة المرنة. ويتمثل هدفه النهائي في معالجة القضايا الملحة المتعلقة بتغير المناخ والاستدامة البيئية، بالإضافة إلى التحديات التي واجهتها المدن. وبصرف النظر عن الالتزام بالمبادئ الأساسية للنماذج والمفاهيم السابقة، شدد التحضر الايكولوجي على الأهداف البيئية مثل تسهيل إقامة مدن منخفضة الكربون ذات بصمة بيئية أقل، وركز على زيادة إنتاجية الموارد من خلال تحويل المدن نحو إنتاج المعرفة والتعليم والابتكار.[16] ومع ذلك، تعرضت مشاريع التحضر الايكولوجي لانتقادات بسبب نجاحها المحدود في الواقع على الرغم من أهدافها الجديرة بالاهتمام. في بعض الحالات، قصرت المشاريع في تحقيق التوازن بين ركائز الاستدامة، وأُهملت الأهداف الاجتماعية. [17] ولذلك اعتُبرت العديد من أهداف التحضر الايكولوجي الأساسية، مثل انعدام الكربون وانعدام النفايات، غير قابلة للتحقيق واستُبدلت بمفاهيم بديلة مثل انخفاض الكربون وانخفاض النفايات في الممارسة العملية. [18]

وفي العقد الأول من القرن الحادي والعشرين والعقد الثاني منه، قدم بعض الباحثين مفهوم المدينة الذكية كنهج يركز على الحوكمة ويستفيد من رأس المال الاجتماعي لتحسين نوعية حياة سكان المناطق الحضرية. تأخذ المدينة الذكية في الاعتبار الأبعاد الستة الرئيسة للأشخاص والتنقل والمعيشة والبيئة والاقتصاد والحوكمة لتحقيق التنمية المتوازنة. وهي تستخدم أحدث التقنيات لتقليل النفايات وانبعاثات الكربون وإدارة تلوث الهواء.[19] في العقد الماضي أو نحو ذلك، استثمرت الحكومات المحلية بشكل كبير في البنى التحتية لتكنولوجيا المعلومات والاتصالات كجزء من تحولها الاجتماعي والتكنولوجي وأطلقت العديد من مشاريع المدن الذكية. تهدف هذه المشاريع إلى الوصول إلى الخدمات والوظائف الحضرية عن طريق استخدام الحلول الرقمية و تكنولوجيا المعلومات والاتصالات ووسائل النقل الذكية. ومع ذلك، قد لا تسهل هذه الحلول الذكية الوصول إلى جميع الخدمات، ولا تزال الحاجة إلى الوصول المادي قائمة يعني ان تكون الخدمات قريبة المسافة للوصول اليها مشياً وليس عن طريق الادوات الرقمية[1]

2.2 ما بعد الجائحة

في أواخر العقد الثاني من الألفية الثانية، أدى تفشي جائحة كورونا إلى أزمة عالمية غير متوقعة وتمثل أكبر أزمة واجهها العالم منذ الحرب العالمية الثانية فرضت العديد من

سكان الأقليات من الانتقال إلى مناطق الأغلبية من خلال فرض قيود على الكثافة السكانية وحصر الأنشطة الصناعية في مجتمعات الأقليات [14] علاوة على ذلك، تعرضت الحداثة إلى ردود فعل وانتقادات من قبلJane Jacobs فقد انتقدت تقسيم المناطق أحادي الاستعمال باعتباره السبب الرئيس لفشل خطط التجديد الحضري في الولايات المتحدة.[15] وفضلت مشاريع التطوير متعددة الاستعمالات، التي تشمل نطاقًا واسعًا من الاستعمالات في المساحات الحضرية ودعت إلى بناء مدينة وأحياء كثيفة ومتنوعة وموجهة نحو الإنسان. مع تطور الفكر المعماري والحضري في العقود اللاحقة ظهرت حركات نقدية لما افرزته الحداثة من صرامة وظيفية مفرطة فبرزت تيارات ما بعد الحداثة ظهرت هذه الحركة في أواخر سبعينيات القرن الماضي، لمعالجة المشكلات الحضرية المعاصرة وتعزيز التخطيط والتصميم الحضري المستدام وحماية البيئة وتعزيز رفاهية السكان، وركزت هذه الحركة على مستوى المجاورات والخدمات والأنشطة اللامركزية لتقليل استخدام السيارات وتشجيع استخدام وسائل النقل العام.[9] وسعى مخططو بعد الحداثة داني، وبلاتر-وزيبرك، وكالثورب إلى إنشاء أحياء متعددة الاستخدامات،[12] وقرب السكن من أماكن العمل، والبعد الإنساني، والمساحات الخضراء الواسعة، وسهولة المشي، والترابط الحضري وينبغي أن تتضمن المجاورات شبكة من الشوارع والمجاورات المتصلة جيدًا، وأن تخلق بيئة تشجع على المشي وركوب الدراجات، وتجمع بين العمل والترفيه. يُعد التنوع عنصرًا أساسيًا في التخطيط الحضري ما بعد الحداثة، والذي يعني مزيجًا من الاستخدامات السكنية، ومستويات الدخل، وفرص العمل والمتاجر، والحدائق، والمؤسسات المدنية.[1] كانت التدخلات المادية لمصممي التمدن ما بعد الحداثيين غير كافية لمعالجة جميع القضايا الحضرية، والتي تطلب بعضها إجابات خارج نطاق التدخلات في الشكل الحضري. وبالتالي، لم تتم معالجة التحديات السابقة مثل حركة المرور وتلوث الهواء والفقر وعدم المساواة بشكل كافٍ في مطلع القرن. في غضون ذلك، برزت قضايا جديدة مثل تغير المناخ وندرة الموارد، مما سلّط الضوء على ضرورة إعادة النظر في نماذج ومفاهيم التخطيط والتصميم الحضري. ونتيجةً لذلك، شهد العقد الأول من القرن الحادي والعشرين ظهور التحضر الايكولوجي كمفهوم تخطيط حضري مهيمن قائم على مبادئ التنمية المستدامة.[1] ويدعو التحضر الايكولوجي إلى مفاهيم مثل المجاورات المستدامة،

تطبيقية نتيجة ندرة الأراضي المتاحة وارتفاع قيمتها. أما خطة وحدة الجوار لبيري، فقد تعرّضت لانتقادات مشابهة، خصوصًا لكونها تدعم الفصل الوظيفي الصارم، مما يؤدي إلى ضعف إمكانية المشي، وانخفاض التفاعل الاجتماعي، والاعتماد المتزايد على السيارات، وما يرافق ذلك من ارتفاع انبعاثات غازات الدفيئة، إضافة إلى محدودية التنوّع في أنماط الحياة داخل الحي الواحد. [5]هذه التحديات مجتمعة مهّدت لظهور حركة جديدة في منتصف القرن العشرين هي الحركة الحداثوية في التخطيط الحضري، التي ظهرت كرد فعل مباشر على الفوضى، والازدحام، وتردي الظروف الصحية التي سادت المدن الأوروبية نتيجة الثورة الصناعية.[10] وفي بدايات القرن العشرين أدت التطورات التكنولوجية في مجالي البناء والنقل إلى ظهور توجهات جديدة في التخطيط الحضري، تبناها عدد من المخططين والمعماريين الحداثيين، مثل لو كوربوزييه وفرانك لويد رايت خلال عشرينيات وثلاثينيات القرن الماضي. وقد انعكست هذه التوجهات في تصميم أحياء حضرية اتسمت بالمباني العالية ذات الطابع الوظيفي، والمساحات المفتوحة الواسعة، والمجمعات السكنية الكبيرة المرتبطة بممرات للمشاة، مع اعتماد واضح على وسائل النقل السريعة[11]. وفي مقابل ذلك، ركزت بعض الاتجاهات التخطيطية السابقة، مثل أفكار هوارد وبيري، على البعد الإنساني وضبط التوسع الحضري، بينما أولى المخططون الحداثيون اهتماماً أكبر بالمباني الشاهقة، والكثافة السكانية العالية، وتطوير شبكات نقل سريعة تعتمد بشكل أساسي على السيارة الخاصة. وانطلقت هذه الرؤية من افتراض مفاده أن توفير عدد كافٍ من مواقف السيارات وشبكات طرق مناسبة سيتيح للسكان الوصول بسهولة إلى مختلف الخدمات والمرافق داخل المدينة.[12]

وعلى الرغم من الاعتقاد بأن التقدم التكنولوجي قادر على حل المشكلات الحضرية، فإن هذا النهج لم يسهم في تحسين نوعية حياة السكان، بل أثر سلباً في استدامة المدن. فقد أدى إلى ظهور مدن مجزأة تقوم على الفصل الوظيفي بين مناطق السكن والعمل وتم تقسيم المناطق إلى استخدامات أحادية لفصل المناطق السكنية عن المشاريع التجارية والصناعية.[1] وأدى إنشاء الطرق السريعة الجديدة إلى زيادة الاعتماد على السيارات مما أسفرت عن استمرار مشاكل التلوث الهوائي والضوضائي الناتج عن الازدحام المروري، [13] مع ذلك، وُجهت انتقادات لتقسيم المناطق باعتباره أداة سياسية تمييزية تُستخدم لمنع

المساحات المفتوحة: يتم تخصيص حوالي 10% من المساحة للمساحات المفتوحة ومرافق ترفيهية داخل المجاورة التي تُسهم في دعم الصحة النفسية والجسدية للسكان، وتعزز من التفاعل الاجتماعي.

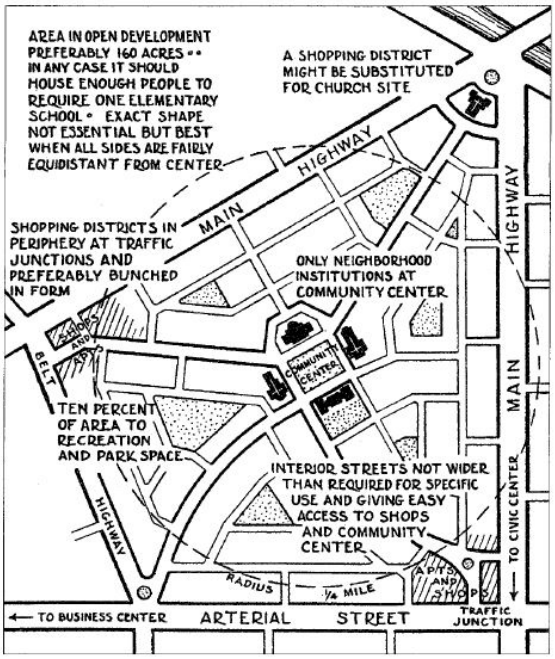
مواقع المؤسسات: توفير جميع الخدمات الأساسية حول نقطة مركزية في المجاورة ضمن مسافة مشي قصيرة بما في ذلك المحلات التجارية، الحدائق، ودور العبادة، وذلك لتقليل الحاجة للتنقل.

المتاجر المحلية: يتم وضع مناطق تسوق كافية للسكان المراد خدمتهم في محيط الوحدة، ويفضل أن تكون عند تقاطعات المرور.

نظام الشوارع الداخلية: يتم تزويد الوحدة بنظام شوارع خاص، وتُخصّص للمرور المحلي فقط ويتم تصميم شبكة الشوارع بشكل متكامل بما يسهل حركة التنقل داخل الوحدة ومنع استخدامها من قبل حركة المرور العابرة.

شكل (3) مبادئ وحدة المجاورة المصدر: اعداد الباحثة استناداً لمبادئ كلارنس بيري

وخلال النصف الأول من القرن العشرين، أخذ نهج المدن الحدائقية ونهج وحدة الحي اهتمامًا كبيرًا بوصفهما محاولتين لإصلاح التدهور الحضري الذي خلفته الثورة الصناعية. الا إنه تعرّض لانتقادات بسبب ضعف دمج وظائف الإنتاج، ونقص الاكتفاء الذاتي، وتشجيع الوحدات السكنية الأحادية للأسرة، إضافةً إلى التوسع الحضري غير المستدام[9]. كما واجهت فكرة الحزام الأخضر المحيط بالمدن الحدائقية صعوبات



شكل (2) نموذج وحدة المجاورة لكلارنس بيري المصدر [7] ص50

تُعدّ مبادئ كلارنس بيري من أوائل المحاولات المنظمة لتأطير المجاورة كمفهوم عمراني متكامل، لتُشكّل مرجعية مركزية في الفكر الحداثي للتخطيط الحضري، ومقدمة ضرورية لفهم تحولات المجاورة الحضرية في الفكر المعاصر، وكان هدف بيري الرئيس هو خلق تفاعل اجتماعي وان التخطيط المادي هو الوسيلة لتحقيقه. لذلك تركزت المبادئ التخطيطية التي وضعها كلارنس بيري على مجموعة من الاسس الحضرية والاجتماعية يمكن ذكرها كما يلي:[4][8] انظر شكل (3).

الحجم: تحديد عدد السكان في المجاورة بين 5000 إلى 9000 نسمة لضمان وجود كثافة كافية لدعم المدرسة التي تعتبر النواة المركزية للمجاورة وتمثل مركز النشاط اليومي، ويُخطط توزيع الخدمات حولها لتسهيل وصول الأطفال دون الحاجة لاجتياز الطرق السريعة

الحدود: تحديد حدود المجاورة بشوارع رئيسة تحيط بها بحيث تُفصل عن المجاورات الأخرى وتُحصر الحركة المرورية الثقيلة خارجها، ما يعزز أمان المشاة.

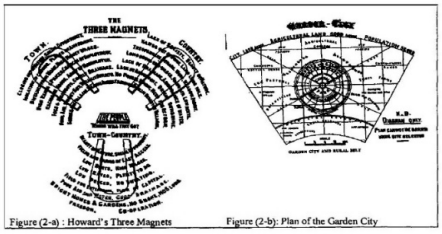
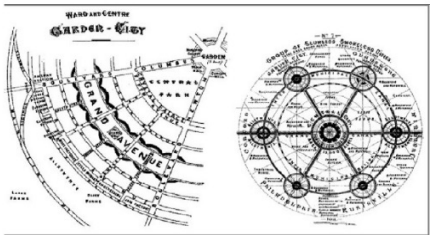


Figure (2-a) : Howard's Three Magnets Figure (2-b): Plan of the Garden City



شكل (1) مدينة الحدائق لايبنزر هوارد المصدر[3]ص36

في عام 1929، قدّم Clarence Perryمفهوم وحدة المجاورة كنهج تخطيط حضري ركّز على المجاورة كوحدة تخطيط أساسية[5]، انظر شكل (2)، طُرحت هذه الخطة في أوائل القرن العشرين عندما واجهت المدن الأمريكية مشكلات مثل النمو غير المخطط له القائم على تطوير القطاع الخاص، والاكتظاظ السكاني، والفصل العنصري، والعزلة الاجتماعية، والجنوح، وقلة المشاركة[6] وقد ركّز هذا النموذج على إعادة تنظيم المجاورة السكني كوحدة تخطيط حضرية أساسية، استجابةً لمشكلات التوسع الحضري غير المنظم في الولايات المتحدة.اقترح بيري تصميم المجاورة بوصفها وحدة متكاملة تستند إلى ستة عناصر تخطيطية، وهي: المدارس، والحدود، والمساحات المفتوحة، والمباني المؤسسية، ومناطق البيع بالتجزئة، وتخطيط الشوارع الداخلية حدد بيري نصف قطر المجاورة بـ400 متر تقريبًا، لتوفر مساكن لعدد من السكان يتراوح بين 5000 إلى 9000 نسمة، بما يسمح بالوصول السهل إلى مركز الخدمات ضمن مسافة مشي.[7] وقد جعل المدرسة الابتدائية هي النواة المركزية للمجاورة، وتُخطط الطرق والمرافق الأخرى حولها لتأمين بيئة آمنة للمشاة. كما اقترح وضع المحال التجارية والخدمات الثانوية على أطراف المجاورة، وربطها بالشوارع الرئيسة التي تحيط بها، مع إبقاء الشوارع الداخلية مخصصة للحركة المحلية فقط، بما يسهم في تقليل الازدحام المروري وتعزيز جودة المجاورة الحضرية. وعلى الرغم من أن المجاورة صُممت كآلية أساسية لتطوير إصلاح حضري شامل على مستوى المدينة، إلا أنها طُبقت بشكل رئيسي لإنشاء مجمعات سكنية، مثل ضواحي الحدائق للطبقة المتوسطة ومناطق الإسكان الاجتماعي.[7]

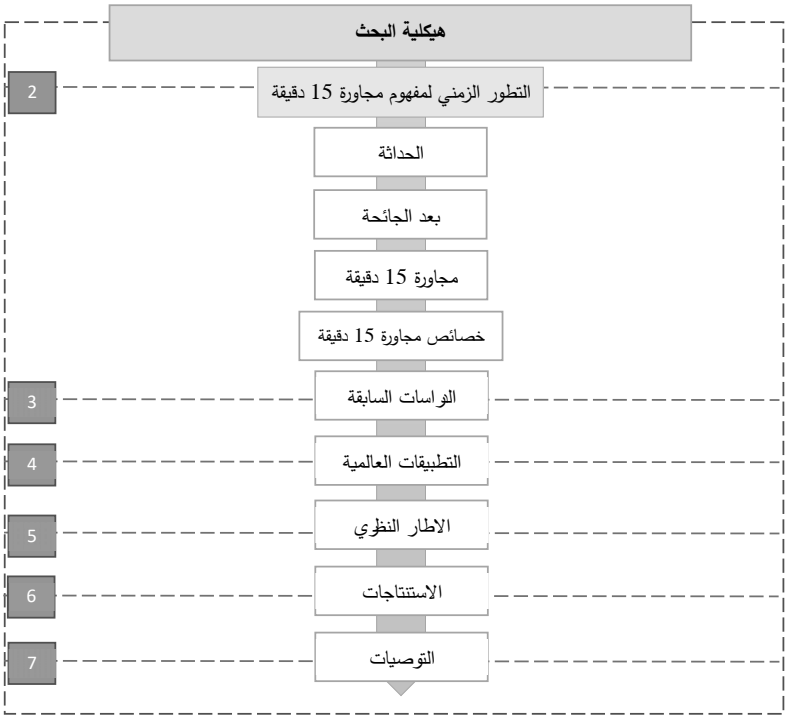
2 – التطور الزمني لمفهوم مجاورة 15 دقيقة

1.2 الحداثة

شكّل تخطيط المجاورات محورًا لجدل واسع بين المخطّطين، وقد بدأ تطبيقه بوضوح منذ حقبة الثورة الصناعية، ليتعزز بشكل أكبر خلال القرن العشرين، حين تبيّن عدم كفاءة المدن الكبرى في تلبية الاحتياجات اليومية لسكانها. ومع تطور المدن بشكل غير منتظم خلال الثورة الصناعية، برزت الحاجة إلى تنظيم حضري أكثر كفاءة يُعالج المشاكل والازمات الحضرية.[1] واستجابة لهذه القضايا فشكّلت المجاورة الحضرية[2] نقطة انطلاق لتجارب تخطيطية سعت إلى إعادة تنظيم المجال الحضري وفق أسس عقلانية ووظيفية. من أبرز هذه المحاولات ما قدّمه المفكر البريطاني إيبنيزر هاورد في أعقاب الثورة الصناعية من خلال طرحه لفكرة المدينة الحدائقية عام 1898 [3] انظر شكل (1)، والتي تقوم على إنشاء مجموعة من المدن الجديدة المترابطة والمكتفية ذاتيًا محاطة بحزام أخضر وتتوزع حول مدينة مركزية كبيرة قُسِّمت كل من هذه المدن الجديدة الدائرية إلى عدة أقسام، صُممت كل منها لاستيعاب ما يصل إلى 5000 نسمة وإن لكل قسم حدودًا مادية يمكن تحديدها، واقترح هوارد ان شبكة النقل في المدينة الخضراء تتميز بالطرق الشعاعية والمسارات المتعرجة ويتم ربط كل مدينة خضراء بالمدينة الكبرى عن طريق السكك الحديدية،[1] وتم فصل استخدامات الأراضي، ويعيش السكان في مساكن منفصلة تقع على مساحات شاسعة من الأرض، وتكون الكثافة السكانية الإجمالية منخفضة وتقع أنشطة التوظيف والتسوق على طول الشوارع المركزية بالقرب من الطرق الرئيسة لذلك يمكن اعتبار اقتراح هوارد بتقسيم المدينة إلى عدة أقسام أحد أقدم الجهود لإدخال فكرة المجاورة في سياق التخطيط وكذلك يمكن اعتبار التركيبة السكانية هي احدى القضايا المهمة التي ذكرت في مقترح هوارد فقد تصور مجتمعًا يستوعب سكانًا مختلطين اجتماعيًا يديرون شؤون المدينة بشكل تعاوني، ويشاركون بنشاط في الأنشطة المدنية، ويلتزمون بالمصلحة العامة[4]

1 – المقدمة

برز مفهوم مجاورة 15 دقيقة كنموذج تخطيطي معاصر يسعى إلى إعادة هيكلة المدن استناداً إلى مبدأ القرب المكاني والتنوع الوظيفي، مما يُمكّن السكان من الوصول إلى احتياجاتهم اليومية الأساسية كالعمل، والتعليم، والرعاية الصحية، والتسوق، والترفيه في غضون 15 دقيقة سيرًا على الأقدام أو بالدراجة. ويستند هذا المفهوم على تقليل الاعتماد على المركبات الخاصة، وخفض انبعاثات الكربون، وتشجيع وسائل النقل النشطة، فضلاً عن تعزيز التماسك الاجتماعي والشعور بالانتماء للمكان. وقد اهتم هذا البحث بتحليل التطور الزمني لمفهوم مجاورة 15 دقيقة من خلال تتبع جذوره الفكرية وتحليل التحولات التي طرأت على خصائصه ومتغيراته التخطيطية، اعتمادًا على المنهج الوصفي التحليلي القائم على مراجعة الأدبيات النظرية والتجارب الحضرية العالمية، بهدف استخلاص إطار نظري يوضح الأسس المعاصرة لهذا المفهوم ليتم اخيرا التوصل الى الاستنتاجات والتوصيات.



المستخلص

أفرزت أنماط التخطيط الحضري الحداثوية القائمة على الفصل الوظيفي والاعتماد على المركبات الخاصة العديد من التحديات البيئية والوظيفية، تمثلت في ارتفاع معدلات التلوث. وقد كشفت جائحة كوفيد-19 بشكل غير مسبوق عن هشاشة هذه النماذج التخطيطية وعدم قدرتها على الاستجابة الفعالة للاحتياجات اليومية للسكان في حالات الطوارئ والأزمات الصحية. مما استدعى ضرورة مواجهة هذه السلبيات من خلال مراجعة المفاهيم وتطوير أفكار تصميمية وتخطيطية أكثر فاعلية تحقق الرفاهية للساكنين. لذلك برز مفهوم مجاورة 15 دقيقة الذي سعى إلى تعزيز إمكانية الوصول والتنقل النشط، وتحقيق مجاورة حضرية مكتفية ذاتياً وقادرة عبر تصميمها على تحقيق جودة الحياة اللازمة. وبالرغم من الانتشار الواسع لهذا المفهوم، إلا أن الأدبيات الحضرية (لاسيما ضمن النطاق المحلي) تعاني من فجوة معرفية فيما يتعلق بتطور مفهوم مجاورة 15 دقيقة من جذوره الحداثوية وصولاً إلى حقبة ما بعد الجائحة. لذا، يهدف البحث إلى تحليل التطور الزمني للمفهوم واستخلاص إطار نظري يوضح خصائصه ومتغيراته الأساسية. ولتحقيق ذلك، اعتمد البحث المنهج الوصفي التحليلي بالاستناد إلى الأدبيات النظرية والتجارب العالمية. وينطلق البحث من فرضية مفادها أن مفهوم مجاورة 15 دقيقة يمثل تحولاً جوهرياً في الفكر التخطيطي، حيث ينتقل من الاعتماد على البعد المكاني إلى التركيز على البعد الزمني، ويعد استجابة نقدية لمعالجة القصور في النماذج الحداثوية. وقد خلص البحث إلى صياغة إطار نظري متكامل يرتكز على خصائص مجاورة 15 دقيقة، والذي يمكن أن يكون بمثابة دليل مرجعي لتوجيه تصميم مجاورات حضرية أكثر استدامة ومرونة.

الكلمات المفتاحية: المجاورة الحضرية، مجاورة 15 دقيقة، البعد الحداثوي، بعد الجائحة، القرب

Abstract

Modernist urban planning patterns, based on functional separation and reliance on private vehicles, have generated numerous environmental and functional challenges. These include high pollution. The COVID-19 pandemic has unprecedentedly revealed the fragility of these planning models and their inability to effectively respond to residents' daily needs during emergencies and health crises. Furthermore, that necessitated addressing these shortcomings by re-evaluating concepts and developing more effective design and planning ideas that achieve resident well-being.

Consequently, the 15-Minute Neighborhood concept emerged, seeking to enhance accessibility and active mobility, and to realize a self-sufficient urban neighborhood capable, through its design, of achieving the necessary quality of life. Despite the widespread adoption of the concept, urban literature suffers from a knowledge gap regarding the evolution of the 15-minute neighborhood concept from its modernist origins to the post-pandemic era. Therefore, the research aims to analyze the develop an analytical theoretical framework that clarifies its basic characteristics and variables, and develop an analytical theoretical framework that clarifies its basic characteristics and variables. To achieve this, the research adopted descriptive-analytical approach, drawing on theoretical literature and global experiences.

The research proceeds from the hypothesis that the 15-minute neighborhood concept represents a fundamental shift in planning thought moving from reliance on the spatial dimension to a focus on the temporal dimension, and serves as a critical response to the deficiencies of modernist models. The research concluded with the formulation of an integrated theoretical framework based on the characteristics of the 15-minute neighborhood, which can serve as a reference guide to guide the design of more sustainable and resilient urban neighborhoods.

Keywords: Urban Neighborhood, 15-Minute Neighborhood, Modernist Dimension, Post-Pandemic, Proximity

مجاورة الـ15ـ دقيقة بين المفهوم والممارسة

دراسة تحليلية لـ تطور المفهوم من البعد الحداثوي الى مابعد الجائحة

الباحثة فتوح كاظم الوسواسي و أ. م. د. حيدر جاسم عيسى الساعدي

الجامعة التكنولوجيا، كلية هندسة العمارة – قسم هندسة التصميم الحضري ، بغداد – العراق

com.gamil@foutouhkadhum iq.edu.uotechnology@essa.j.haider

The 15-Minute Neighborhood Concept

An Analytical Study of the Concept's Evolution from the Modernist Dimension to the Post-Pandemic Era / Original article

Researcher **and Assist. Prof. Dr.**

Foutouh Kadhum Al-wiswasy **Haider Jasim Essa Al- Saaidy**

University of Technology, College of Architectural Engineering - Department of Urban Design Engineering, Baghdad - Iraq

foutouhkadhum@gamil.com haider.j.essa@uotechnology.edu.iq

[24] M. S. Nasr, I. M. Ali, A. M. Hussein, A. A. Shubbar, Q. T. Kareem, and A. T. AbdulAmeer, (2020), "Utilization of locally produced waste in the production of sustainable mortar," *Case Stud. Constr. Mater.*, vol. 13, p. e00464.

[25] Z. N. Jawad and M. A. Hassan, (2023), "Mechanical evaluation of sustainable concrete production from porcelain waste," in *IOP Conference Series: Earth and Environmental Science*, IOP Publishing, p. 12042.

[11] K. Rashid, A. Razzaq, M. Ahmad, T. Rashid, and S. Tariq, (2017), "Experimental and analytical selection of sustainable recycled concrete with ceramic waste aggregate," *Constr. Build. Mater.*, vol. 154, pp. 829–840.

[12] A. H. Allawi, A. I. AL-Hadithi, and A. S. Mohmoud, (2021), "Effects of waste plastic pet fibers on the fresh and hardened of normal concrete," *Iraqi J. Civ. Eng.*, vol. 15, no. 1, pp. 47–58.

[13] T. F. Abbas and Z. K. Abbas, (2023), "Manufacture of load bearing concrete masonry units using waste demolishing material," *J. Eng.*, vol. 29, no. 04, pp. 105–118.

[14] R. Kurda, J. D. Silvestre, and J. de Brito, (2018), "Toxicity and environmental and economic performance of fly ash and recycled concrete aggregates use in concrete: A review," *Heliyon*, vol. 4, no. 4.

[15] E. A. Khalil and M. N. AbouZeid, (2023), "Computation of the environmental performance of ready-mix concrete for reducing CO2 emissions: A case study in Egypt," *Energy Reports*, vol. 9, pp. 144–148.

[16] B. Lei, L. Yu, Z. Chen, W. Yang, C. Deng, and Z. Tang, (2022), "Carbon emission evaluation of recycled fine aggregate concrete based on life cycle assessment," *Sustainability*, vol. 14, no. 21, p. 14448.

[17] A. A. Hashim, A. I. Al-Mosawi, and S. A. Abdulsada, (2025), "Investigating the mechanical properties, durability, microstructure, and embodied CO2 emissions of silica fume-infused sustainable concrete," *Int. J. Appl. Ceram. Technol.*, vol. 22, no. 4, p. e15136.

[18] D. J. M. Flower and J. G. Sanjayan, (2007), "Green house gas emissions due to concrete manufacture," *Int. J. life cycle Assess.*, vol. 12, pp. 282–288.

[19] M. T. Hasan, A. A. Abdul-Hamead, and F. M. Othman, (2025), "Sustainable Concrete Using Porcelain and Clay Brick Waste as Partial Sand Replacement: Evaluation of Mechanical and Durability Properties," *Constr. Mater.*, vol. 5, no. 4, p. 78.

[20] M. Roig-Flores *et al.*, (2023), "Utilisation of ceramic stoneware tile waste as recycled aggregate in concrete," *Buildings*, vol. 13, no. 8, p. 1968.

[21] G. J. Khoshnaw *et al.*, (2024), "Experimental investigation on pervious recycled aggregate concrete made of waste porcelain," *Civ. Eng. J.*, vol. 10, no. 09.

[22] Z. Keshavarz and D. Mostofinejad, (2019), "Porcelain and red ceramic wastes used as replacements for coarse aggregate in concrete," *Constr. Build. Mater.*, vol. 195, pp. 218–230.

[23] W. E. Elemam, I. S. Agwa, and A. M. Tahwia, (2023), "Reusing ceramic waste as a fine aggregate and supplemental cementitious material in the manufacture of sustainable concrete," *Buildings*, vol. 13, no. 11, p. 2726.

5. المصادر

[1] A. A. Hashim and I. A. Najem, (2025), "Sustainable use of ground granulated blast-furnace slag as a partial substitute for cement: workability, mechanical properties, durability, and environmental impact," in *IOP Conference Series: Earth and Environmental Science*, IOP Publishing, p. 12027.

[2] R. B. M. Taher, (2009), "تصنيع خرسانة خالية من الركام الناعم ودراسة إمكانية استخدامها بدلا من الاعتيادية," *Al-Mukhtar J. Sci.*, vol. 23, no. 1, pp. 49–67.

[3] A. A. Hashim, R. Anaee, M. S. Nasr, A. Shubbar, and T. S. Alahmari, (2025), "Mechanical properties, corrosion resistance and microstructural analysis of recycled aggregate concrete made with ceramic wall waste and ultrafine ceria," *J. Mater. Res. Technol.*, vol. 36, pp. 627–640.

[4] A. A. Hashim, R. Anaee, and M. S. Nasr, (2025), "Enhancing the sustainability, mechanical and durability properties of recycled aggregate concrete using calcium-rich waste glass powder as a supplementary cementitious material: An experimental study and environmental assessment," *Sustain. Chem. Pharm.*, vol. 44, p. 101985.

[5] E. Sánchez, J. García-Ten, V. Sanz, and A. Moreno, (2010), "Porcelain tile: almost 30 years of steady scientific-technological evolution," *Ceram. Int.*, vol. 36, no. 3, pp. 831–845.

[6] Y. Yousefi and F. Tariku, (2022), "Hygrothermal Properties of Cladding Materials at Different Temperatures and Relative Humidities," in *International Conference on Building Energy and Environment*, Springer, pp. 613–622.

[7] K. Asante-Kyei, C. N. D. Dodoo, and A. Addae, (2025), "Resolving environmental impact of discarded ceramic artifacts on selected dumpsites in the Sekondi-Takoradi metropolis, Ghana," *Discov. Environ.*, vol. 3, no. 1, p. 129.

[8] K. A. Tariq, W. A. Adil, A. Salhi, and A. Waleed, (2026), "Experimental investigation of sustainable concrete production using ceramic waste powder as partial fine aggregate replacement," *Sci. Rep.*

[9] R. J. Daniel and S. P. Sangeetha, (2021), "Experimental study on concrete using waste ceramic as partial replacement of aggregate," *Mater. Today Proc.*, vol. 45, pp. 6603–6608.

[10] M. J. Taher, E. H. Abed, and M. S. Hashim, (2023), "Using ceramic waste tile powder as a sustainable and eco-friendly partial cement replacement in concrete production," *Mater. Today Proc.*

- **تحسين الأداء:** يمكن لركام البورسلين الناعم ان يحسن من الخواص الميكانيكية ويقلل الكثافة وانبعاثات ثنائي اوكسيد الكاربون بشكل ملحوظ.

التطبيقات المقترحة:

مدى ملاءمة الخرسانة بالبورسلين	نوع المشروع
ممتازة	الأرصفة والطرق المحلية
جيدة	الحواجز والحوائط الاستنادية
مناسبة	المباني السكنية منخفضة الارتفاع
فعالة	البنية التحتية
غير مفضلة (تحتاج مقاومة أعلى)	الجسور أو الأبراج العالية

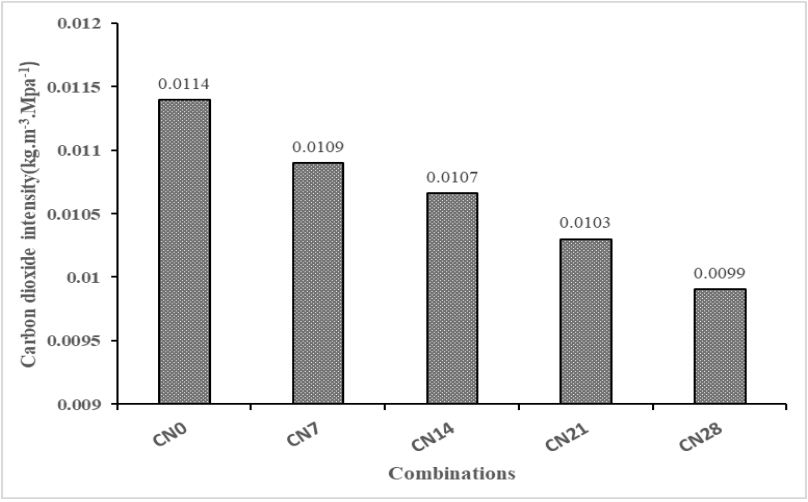
4. الاستنتاجات

سعت هذه الدراسة إلى استخدام مخلفات صناعية وهي مخلفات البورسلين كبديل جزئي للركام الناعم الطبيعي في الخرسانة، ليس فقط لتحسين خصائصها، بل أيضاً للحد من التلوث الناجم عن الكميات الهائلة من هذه المخلفات التي تُلقى يومياً في البيئة. ومن المزايا الإضافية لاستخدام هذه المخلفات أنها توفر مصدراً لمواد متوفرة بسهولة وبأسعار معقولة لاستخدامها في الخرسانة. ويمكن تلخيص نتائج هذا البحث بإيجاز على النحو التالي:

1) أشارت نتائج الاختبارات الميكانيكية، والتي شملت قياس مقاومة الانضغاط ، ومقاومة الانحناء للعينات باستخدام مسحوق ركام البورسلين الناعم إلى أن النسبة المثلى لتحسين الخصائص الميكانيكية هي 28%.
2) نتائج الكثافة الظاهرية الى أن استبدال الرمل بالبورسلين أدى إلى انخفاض واضح في كثافة الملاط حيث حقق الخليط الحاوي على 28% ركام بورسلين ناعم قيمة كثافة ظاهرية اقل ب 29% عن ماهي عليه في الخليط المرجعي.
3) قد يؤدي استبدال الرمل الطبيعي بنفايات البورسلين إلى تقليل البصمة الكربونية لخلطات الخرسانة بنسبة 13% عند استبدال الركام الناعم الطبيعي ب 28% من ركام البورسلين. ويُبرز هذا الانخفاض الميزتين المزدوجتين لاستخدام نفايات البناء والهدم: تقليل النفايات المُدفنة والتخفيف من الآثار البيئية المرتبطة باستخراج المواد الخام.
4) استخدام الخرسانة المصنعة من مخلفات البورسلين في العراق تحقق عدة فوائد اقتصادية وبيئية في السياق المحلي منها:
   - **الاستدامة البيئية:** العراق يعاني من تراكم نفايات البناء، واستخدام مخلفات البورسلين يساهم في تقليل هذه النفايات.
   - **خفض التكاليف:** توفر المواد محليًا يقلل من تكاليف النقل والشراء، خاصة في المناطق الحضرية مثل بغداد وبابل.

5.3 التقييم البيئي

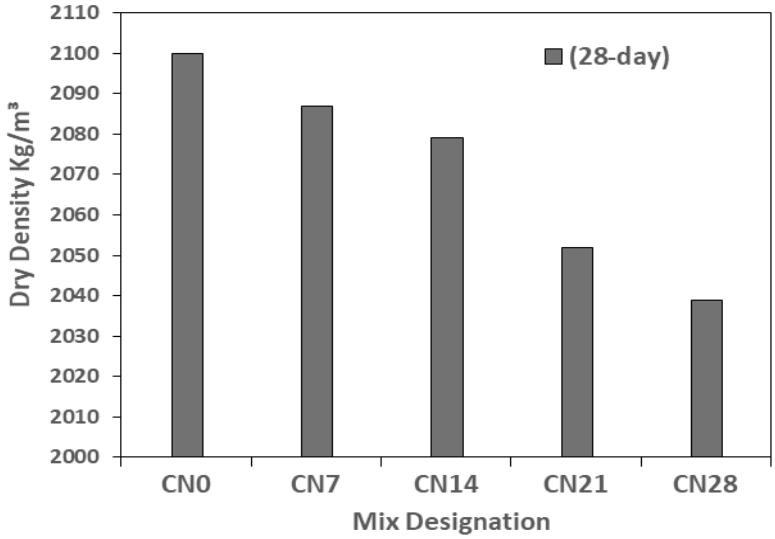
لتقييم الأداء البيئي للتركيبات التي تم تحضيرها في هذه الدراسة، تم إجراء تقدير مبسط للكربون المتضمن (مقدر بالكيلوجرام من مكافئ ثاني أكسيد الكربون لكل متر مكعب من الخرسانة) للرمل الطبيعي ومواد النفايات البديلة (ركام البورسلين الناعم). يوضح الشكل (7) كثافة ثاني أكسيد الكربون المنبعثة من مخاليط الملاط الحاوية على ركام البورسلين الناعم كبديل جزئي للركام الناعم الطبيعي والخليط المرجعي. تنخفض شدة ثنائي اوكسيد الكاربون لمنبعثة لكل وحدة من مقاومة الضغط بزيادة نسب استبدال الركام الناعم الطبيعي بركام البورسلين الناعم، مقارنةً بمزيج الخرسانة المرجعي (CN0). وقد لوحظ أكبر انخفاض في كثافة ثاني أكسيد الكربون عند عمر 28 يومًا في مزيج الملاط الحاوي على 28 % وزنيًا ركام البورسلين الناعم، حيث كان أقل بنسبة 13% من مزيج الملاط المرجعي. وبالتالي، فإن ادخال ركام البورسلين المعاد تدويره في الخرسانة يقلل من العواقب البيئية الناجمة عن التخلص غير السليم، بالإضافة إلى تخفيف انبعاثات غازات الاحتباس الحراري من قطاع البناء. تشير النتائج الموضحة في الشكل 7 إلى أن الاستبدال الجزئي للركام الناعم الطبيعي بركام البورسلين الناعم المعاد تدويره بنسبة استبدال 28 %، تُظهر كفاءة بيئية فائقة من حيث تقليل انبعاثات ثاني أكسيد الكربون.



شكل (7): نتائج قياس شدة ثنائي اوكسيد الكاربون المنبعث للخلطة المرجعية وخلطات استبدال الرمل بمخلفات البورسلين

4.3 الكثافة الظاهرية

تم قياس الكثافة الظاهرية لكافة الخلطات التي تم اعدادها لعينات مكعبة (5 × 5 ×5) سم عند عمر 28. الشكل (6) يوضح نتائج الكثافة الظاهرية للخليط المرجعي والخلائط الحاوية على ركام البورسلين الناعم كبديل جزئي للركام الناعم الطبيعي. تتحدد الكثافة الظاهرية للملاط أو الخرسانة المتصلبة بعوامل عديدة، منها وجود فراغات هوائية محصورة أو مدمجة، وكمية وكثافة الركام، ونسبة الإسمنت والماء[24]. أظهرت النتائج انخفاضا ملحوظا في قيم الكثافة عند استبدال الركام الناعم الطبيعي جزئيا بركام البورسلين الناعم، وازداد الانخفاض بزيادة نسب الاستبدال. حيث سجل الخليط الحاوي على 28% ركام بورسلين ناعم كثافة ظاهرية مقدارها (2039) غم لكل سنتمتر مكعب وهي اقل من ماهي عليه للخليط المرجعي بـ (29 %) والتي تبلغ قيمة الكثافة الظاهرية له (2100)غم لكل سنتمتر مكعب. ويعود السبب لهذا الانخفاض الى انخفاض الكثافة النوعية والمسامية الاعلى لركام البورسلين مقارنة بالرمل العادي، مما يؤدي إلى خرسانة اقل كثافة[25] .

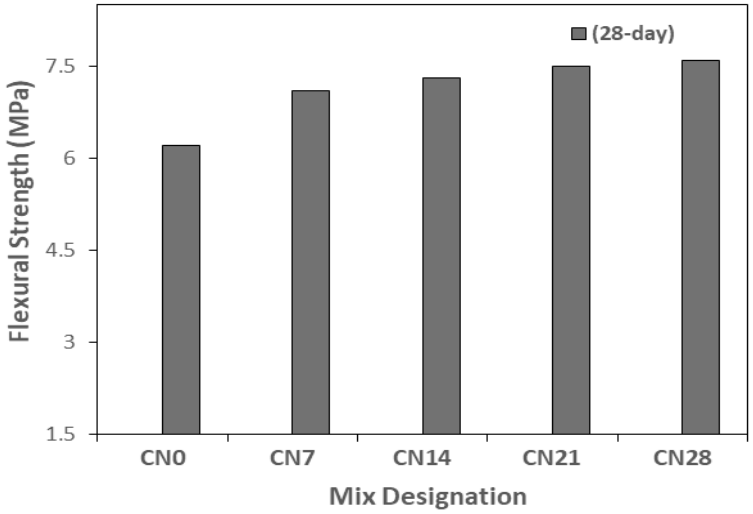


.

شكل (6): نتائج فحص الكثافة للخلطة المرجعية وخلطات استبدال الرمل بركام البورسلين الناعم

3.3 مقاومة الانثناء

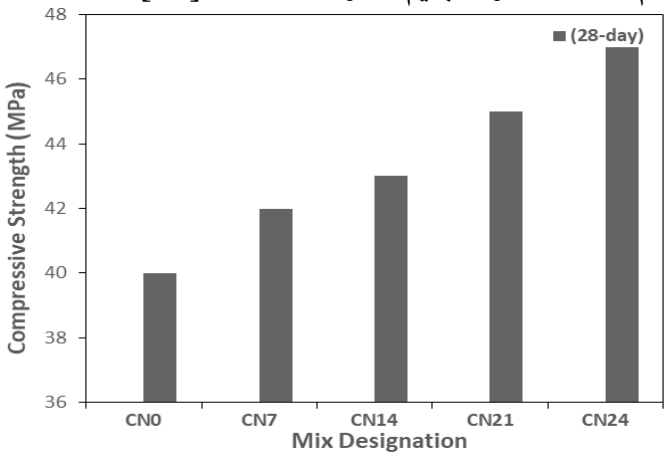
تم اجراء فحص مقاومة الانثناء بعد 28 يوم من المعالجة لعينات موشورية (4 × 4 × 16) سم مكعب، حيث تم تعريضها لمعدل تحميل قدره 1 ميجا باسكال\دقيقة اثناء الاختبار. الشكل (5) يوضح نتائج مقاومة الانثناء للخليط المرجعي والمخاليط الحاوية على ركام البورسلين الناعم كبديل جزئي للركام الناعم الطبيعي. يتضح من الشكل (5) ان مقاومة الانثناء لخلطات الملاط اتخذت مسارا مشابها لما هو عليه بالنسبة لمقاومة الانضغاط حيث ازدادت قيم مقاومة الانثناء تدريجيا بزيادة نسب استبدال الركام الناعم الطبيعي بركام البورسلين الناعم. حيث سجلت المخاليط (CN7، CN14، CN21، CN28) قيم مقاومة انثناء بلغت (7.2، 7.3، 7.5، 7.6) ميجا باسكال على التوالي وهي اعلى من ماهي عليه في الخليط المرجعي (CN0) والتي بلغت مقاومة الانثناء له (6.7) ميجا باسكال بـ (5. 14%، 7. 17%، 21%، 5.22%) على التوالي. ان هذا التحسن في مقاومة الانثناء و خاصةً عند نسبة استبدال 28% هو نتيجة للخصائص الكيميائية والفيزيائية التي يتميز بها البورسلين ومنها تركيبته الكيميائية التي يكون معظمها من السليكا والالومينا الغير متبلورة، مما يُسهّل التفاعل البو زولاني ويُعزز تكوين نواتج التفاعل الإسمنتي مثل سيليكات الكالسيوم المائية (C-S-H) والصلابة العالية التي يتميز بها، وبالتالي يُحسّن مقاومة الضغط والانثناء[23].



شكل (5): نتائج مقاومة الانثناء للخلطة المرجعية وخلطات استبدال الرمل بركام البورسلين الناعم

2.3 مقاومة الانضغاط

تم اجراء فحص مقاومة الانضغاط لكافة الخلطات وفقا للمواصفة البريطانية (BS1881–Part116). يوضح الشكل (4) نتائج اختبار مقاومة الانضغاط لعينات مكعبة (5 × 5 ×5) سم عند عمر 28، والتي خضعت للاختبار بمعدل تحميل (0.35) ميجا باسكال/ثانية. تم حساب كل نتيجة معروضة من متوسط ثلاث قراءات. كشفت النتائج أنه عند عمر 28 يومًا، سجلت قيم اختبار الانضغاط لعينات الملاط التي استبدل فيها الركام الطبيعي بركام البورسلين الناعم قيم مقاومة انضغاط اعلى من ماهي عليه في الخليط المرجعي وتزايدت القيمة تدريجيا بزيادة نسبة الاستبدال حيث سجل الخليط (CN28) قيمة مقاومة انضغاط تساوي (47) ميجا باسكال وهي اعلى بـ (17.5%) عما ماهي عليه في الخليط المرجعي حيث كانت قيمة مقاومة الانضغاط لها تساوي (40) ميجا باسكال. ان هذه الزيادة تدل على التأثير الايجابي لاستبدال الركام الناعم الطبيعي بركام البورسلين الناعم ويمكن تعليل سبب هذا التحسن الى فعالية دقائق البورسلين كونه مادة غير خاملة بفضل احتوائه على الالومينا والسليكا الفعالتين الفعاليتين مما يسهم في التفاعلات البوزولانية واكمال تفاعلات الإماهة. مضافا الى ذلك الصلادة العالية لبلاط البورسلين مقارنة بالركام الناعم الطبيعي الناتج من تصنيعه تحت ضغط عالي. ونتيجةً للأسباب التي تم ذكرها اظهرت خلطات الملاط التي استبدل فيها الركام الناعم الطبيعي بركام البورسلين الناعم تحسنا ملحوظا بقيم مقاومة الانضغاط[22] .

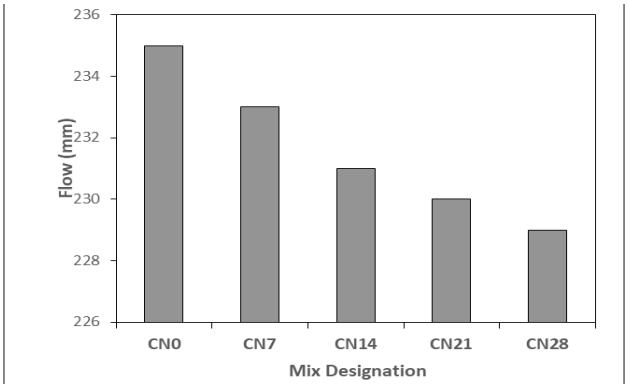


شكل (4): نتائج مقاومة الانضغاط للخلطة المرجعية وخلطات استبدال الرمل بمخلفات البورسلين

3 .النتائج والمناقشة

1.3فحص الهطول

تم اجراء اختبار الهبوط لكل من الخلطة القياسية والخلطات التي تم فيها الاستبدال الجزئي للركام الناعم الطبيعي بركام البورسلين الناعم من البورسلين المعاد تدويره، كما هو موضح في الشكل (3) وفقا للمواصفة الأمريكية (7(ASTM C 143 . وذلك بعد اكمال عملية الخلط مباشرة. وقد تم تحديد قياس هبوط الخلطة المرجعية (CN0) حيث كانت (235) ملم. لاحقًا، عند بدء عملية الاستبدال، لوحظت نقصان ملحوظ وسريع في قيمة الهبوط ليصل الى (229) ملم عند نسبة استبدال 28%، مما يشير إلى أن ركام البورسلين الناعم يوثر سلبا على قابلية التشغيل ، بما يتوافق مع النتائج التي أوردها باحثون اخرون [19]. ويعود النقصان في قابلية التشغيل لخلطات الملاط بزيادة الاستبدال الجزئي للركام الناعم بركام البورسلين الناعم الى عدة اسباب منها الاشكال الزاوية لركام البورسلين الناعم وحوافه الحادة التي ادت الى حدوث احتكاك عالٍ مع بقية مكونات الخلطة، مما ادى الى انخفاض سيولة الخلطة، وهذا بدوره قلل كثيرا من هطول الخليط. اضافة الى ذلك فان الطبيعة المسامية لمسحوق ركام البورسلين وقدرة على امتصاص الماء العاليتين مقارنة بالركام الطبيعي ايضا ساهمتا بنقصان الهطول للخلطات وخاصة عند نسب الاستبدال العالية[20]، [21].



شكل (3): نتائج فحص الهطول للخلطة المرجعية وخلطات استبدال ركام البورسلين الناعم

3.9.2 فحص مقاومة الانثناء

تم صب ثلاث مواشير قياسية بأبعاد (4×4×16) سم3 لإتمام هذا الفحص واعتماد عمر 28 يوم لكل خلطة من خلطات الملاط الإسمنتية. لكل نسبة فُحِصَت ثلاثة مواشير، واعتُمِد المتوسط الحسابي لها لتمثيل النتيجة.

4.9.3 فحص الكثافة الظاهرية

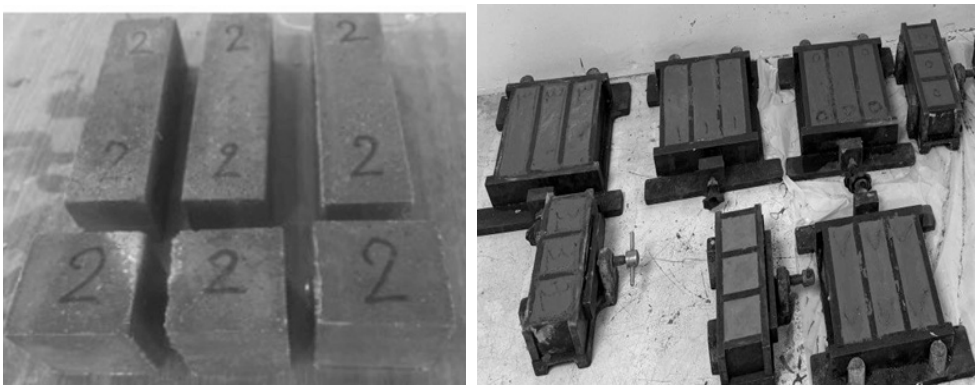
أُجري اختبار الكثافة على نفس المكعبات التي رُغِبَ في قياس مقاومتها للضغط، وذلك بقياس وزن وحجم المكعبات بدقة وهي في حالة تشبع وجفاف. حُدِّدت الكثافة الظاهرية بقسمة وزن كل مكعب على حجمه. لكل نسبة فُحِصَت ثلاثة مكعبات، واعتُمِد المتوسط الحسابي لها لتمثيل النتيجة.

5.9.2 التقييم البيئي

لحساب تأثير الاحتباس الحراري للمخاليط التي تم تحضيرها، يُقسم مؤشر كثافة ثاني أكسيد الكربون (كجم/م^3. ميجا باسكال$^{-1}$) وهو كمية ثاني أكسيد الكربون المنبعثة من المادة الرابطة (الأسمنت) والركام الناعم الطبيعي، ركام البورسلين المعاد تدويره، الماء والملدن لكل وحدة حجم من الخليط على مقاومة الضغط المحددة للخليط بالميجا باسكال عند عمر معين (28) يومًا. تم الحصول على انبعاثات ثاني أكسيد الكربون من المواد المستخدمة في تكوين خلطات البورسلين المعاد تدويرها المُصاغة في هذه الدراسة من خلال مراجعة الدراسات السابقة واعتماد القيم المذكورة فيها، كما هو موضح في الجدول (3).

الجدول (3): إحصاءات الأثر البيئي لتصنيع مكونات ملاط البورسلين المعاد تدويره

المادة	كيلوغرام مكافئ من CO2	المصدر
أسمنت بورتلاند عادي	9.27×10−1	[14]
الركام الناعم (الرمل)	9.87×10−3	[15]
الركام الناعم المعاد تدويره	5.84×10−3	[16]
الماء	1.33×10−4	[17]
الملدن الفائق	10×10−3	[18]



شكل (2): ثلاثة نماذج (مكعبات وموشور) لكل خلطة من خلطات الملاط الاسمنتية

9.2 اسلوب الانضاج

تم فك القوالب بعد 24 ساعة من انتهاء عملية الصب، بعد ذلك تم غمر جميع النماذج في الماء بدرجة حرارة (24±2 درجة سيليزية) لحين وقت الفحص.

9.2 الفحوصات

1.9.2 فحص الهطول (الانسياب)

يعتبر من اهم وابسط الفحوصات ويستخدم لتحديد ومعرفة قوام الملاطات الاسمنتية الطازجة (الطرية) بتحديد مقدار ومدى هبوطها بعد تشكيلها على هيئة مخروط ناقص، وبالتالي يساعد على التأكد من انتظامية خلطات الملاط الاسمنتية وتجانسها وضبط جودتها قبل الصب مباشرة. تم اجراء فحص الهطول لجميع الخلطات المحضرة بعد اتمام عملية الخلط مباشرة، حيث اجري الفحص وفقا للمواصفة الامريكية ASTM C1437 .

2.9.2 فحص مقاومة الانضغاط

تم صب ثلاث مكعبات قياسية بأبعاد (5×5×5) سم3 لإتمام هذا الفحص واعتماد عمر 28 يوم لكل خلطة من خلطات الملاط الإسمنتية. لكل نسبة فُحِصَت ثلاثة مكعبات، واعتُمِد المتوسط الحسابي لتمثيل النتيجة.

الخلطة بالرمز (CNO) تشير الى الخلطة المرجعية الخالية من ركام مخلفات البورسلين الناعم، اما الرموز (CN28،CN21،CN14،CN7) فتشير الى الخلطات الحاوية على ركام البورسلين الناعم بنسبة (28%،21%،14%،7%) على التوالي.

7.2 تهيئة القوالب

تم تهيئة قوالب الصب (المكعبات والمواشير) لفحص مقاومة الانضغاط ومقاومة الانثناء من خلال تنظيفها ودهنها بطبقة خفيفة من الزيت لتسهيل عملية فكها بعد التصلب كما موضح ادناه في الشكل (1).

شكل (1): قوالب فحص مقاومة الانضغاط ومقاومة الانثناء

8.2 الصب والرص

بعد اتمام عملية الخلط، تم صب نماذج المونة الاسمنتية في القوالب القياسية (مكعبات ومواشير)، بعد ذلك تم رصها باستخدام عصا الرص اليدوية، تم صب ثلاث قوالب بشكل مكعبات بأبعاد (5×5×5) سم3 وثلاث قوالب بشكل موشور بأبعاد (4×4×16) سم3 لكل خلطة من الخلطات وكما موضح في الشكل (2) لغرض إجراء فحص مقاومة الانضغاط والانثناء والكثافة وحساب انبعاثات ثنائي اوكسيد الكاربون بعمر 28 يوم.

(Glenium 54) لتحضير جميع خلطات الملاط. يصنف هذا الملدن الفائق حسب المواصفة الامريكية (ASTM C494) الى نوع (F).

5.2 ركام البورسلين الناعم

تفاصيل إدخال مخلفات البورسلين في مكونات الخرسانة

طريقة التحضير

1 - **جمع المخلفات:** تُجمع مخلفات البورسلين من مصانع البلاط أو مواقع الهدم.

2 - **الفرز والتنظيف:** تُزال الشوائب والمواد العضوية لضمان نقاء المادة.

3 - **الطحن:** تُطحن المخلفات إلى مسحوق ناعم باستخدام جهاز لوس انجلوس. واكمال التنعيم باستخدام طاحونه من الستنلس ستيل.

2. 6 تصميم الخلطات بمخلفات البورسلين

في هذه الدراسة تم اعتماد خمس خلطات للملاطات الاسمنتية وهي:

- الخلطة المرجعية الخالية من مخلفات البورسلين.
- اربع خلطات تحتوي على نسبة (7%، 14%، 21%، 28%) من ركام البورسلين الناعم كنسبة استبدال من وزن الرمل. ان نسبة الخلط المستخدمة للملاط الاسمنتى هي [اسمنت (1): رمل (3)] ونسبة الماء الى الاسمنت (W/C) مقدارها (42.0)، نسبة الملدن 1.5 % من وزن الاسمنت وكما مبينة بالجدول (2).

جدول (2): تفاصيل ونسب مكونات الملاط الاسمنتي لكل خلطة من الخلطات الخمس

رقم الخلطة	رمز الخلطة	نسبة الاستبدال %	الاسمنت (غم)	الرمل (غم)	مخلفات البورسلين (غم)	الماء (غم)	الملدن (غم)
1	CN0	0	1000	3000	0	420	15
2	CN7	7	1000	2790	210	420	15
3	CN14	14	1000	2580	420	420	15
4	CN21	21	1000	2370	630	420	15
5	CN28	28	1000	2160	840	420	15

2. المواد المستخدمة وطرق الفحص

1.2 الاسمنت

تم استخدام اسمنت بورتلاند كلسي اعتيادي (غير مقاوم للأملاح) يسمى محليا بـ (اسمنت كار وعلامته التجارية (CEM II /A-L 42.5 R). المطابق للمواصفات العراقية الخاصة بالسمنت البورتلندي م .ق. ع 1984/5 [12] .

2.2 الركام الناعم الطبيعي

تم استخدام الرمل كركام ناعم وتم تدريجه وفقا للمواصفات العراقية رقم 45 لسنة 1980 [13] وكان الرمل ضمن منطقة التدرج الثانية كما موضح في جدول (1).

جدول (1): تدرج الركام الناعم (الرمل)

رقم المنخل	المار التراكمي %	منطقة التدرج الثانية
4.75	100	100-90
2.36	93.5	100-75
1.18	73.1	90-55
0.6	58	59-35
0.3	17.5	30-8
0.15	3.6	10-0

3.2 الماء

استعمل ماء الشرب الاعتيادي لمدينة الحلة لتحضير جميع الخلطات في هذا البحث.

4.2 الملدن

ان صلابة وقوة الخرسانة والملاط الاسمنتي تتناسب بشكل عكسي مع كمية الماء المضافة للخلطة، وفي ذات الوقت لا يمكن الحصول على خلطة ذات قابلية تشغيل جيدة باستخدام نسبة ماء قليلة، لذا يفضل استخدام ملدنات تقليل الماء. تم استخدام ملدن نوع

تدهور بيئي ومشاكل في إدارة النفايات [7]. إلا أن الدراسات الحديثة المحدودة أثبتت أن هذه المخلفات، بعد معالجتها، يمكن أن تُستخدم كمادة بديلة في الخلطات الخرسانية، مما يساهم في تحسين بعض الخصائص الميكانيكية والفيزيائية للخرسانة، مثل زيادة مقاومة الانضغاط وتقليل الكثافة لذا فإن إعادة استخدام مخلفات البورسلين في الخرسانة يُمكن أن يُحدث فرقًا كبيرًا في تقليل التكاليف، وتحسين الأداء، والحد من التلوث البيئي [8]، [9]، [10]، [11] ويوجد نقص ملحوظ في الدراسات المنشورة حول فعالية إعادة تدوير نفايات البورسلين المهملة التي تُلقى في المجتمعات المحلية. لذا تهدف هذه الدراسة إلى التحري عن امكانية هذا الاستخدام، وتقديم توصيات عملية لتطبيقه في مشاريع البناء المحلية، خاصة في المناطق الحضرية والصناعية.

1 – المقدمة

تعد الخرسانة مادة البناء الاكثر انتشارا على مستوى العالم نتيجة لقوتها ومتانتها وامكانية التحكم بخواصها الهندسية عن طريق التحكم في نوعية موادها الاولية ونسب خلطها المناسبة اضافة الى التحكم في صناعتها وصبها[1] . ونتيجة لصفاتها المتميزة فان استخدامها يتزايد باستمرار[2]. حيث لوحظ ان الاستهلاك السنوي للخرسانة يتجاوز 10 مليار طن في المجتمعات الصناعية المعاصرة ومع نمو سكان العالم من المتوقع ان يصل الاستهلاك السنوي للخرسانة الى نحو حوالي 18 مليار طن [3]. ان انتاج الخرسانة مسؤول عن مجموعة من القضايا البيئية المهمة ومنها الانبعاثات الكبيرة لثنائي اوكسيد الكاربون الناتجة من صناعة الاسمنت واستخراج الركام (الناعم والخشن على حد السواء. في ظل التحديات البيئية والاقتصادية التي تواجه قطاع البناء في العراق، تبرز الحاجة إلى حلول مبتكرة ومستدامة تساهم في تحسين جودة المنشآت وتقليل الأثر البيئي. من بين هذه الحلول، يأتي استخدام المواد المعاد تدويرها كخيار فعّال وواعد[4].

يزداد استخدام كاشي البورسلين بكثرة حاليا وتوسع انتاج البورسلين بأنواع واحجام مختلفة وذلك لما يتميز به كمؤشر على التكنولوجيا العالية بإعطائه جمالية بمستوى عالي وانخفاض في قابلية امتصاص الماء وله القابلية على مقاومة التأثيرات الكيميائية ومقاومة جيدة للانجماد [5]. كما يعتبر مادة لها القابلية على مقاومة التأثيرات الفيزيائية مثل الحرارة والرطوبة او تغير في درجات الحرارة العالية وعلى ضوء ذلك لا يحدث فيه وانكماش او انزلاق للقطع وبذلك فانه يمنع أي تسرب للمياه تحته. على ضوء هذه الخصائص التي تتميز بها هذه المادة فأنها تستخدم لتغليف الواجهات وكذلك اكساء الارضيات والجدران خصوصا لما لها من قابلية امتصاص منخفضة مما يجعلها سهلة الالتصاق والتثبيت اثناء البناء[6].

تُعد مخلفات البورسلين من النفايات الشائعة في مواقع البناء والمصانع، وغالبًا ما تُهمل أو تُرمى دون الاستفادة منها. مع ازدياد إنتاج واستهلاك البورسلين نتيجةً للتحديث والتوسع الحضري، بات التخلص غير السليم من هذه المواد مصدر قلق متزايد، مما يؤدي إلى

تهدف هذه الدراسة إلى تقييم إمكانية استخدام احد انواع مخلفات البلاط وهو البورسلين كمكون كبديل للركام الناعم في الخلطات الخرسانية، وذلك لتحقيق أهداف بيئية واقتصادية وتقنية في قطاع البناء، حيث تسعى الدراسة الى تقليل النفايات الناتجة عن صناعة وتكسير البورسلين، تحسين خصائص الخرسانة، خفض تكلفة المواد الإنشائية، خفض انبعاثات ثاني اوكسيد الكاربون الناتج عن عمليات التعدين و دعم التوجه نحو البناء المستدام في العراق.تم اعداد خمس خلطات من الملاط عن طريق استبدال الركام الناعم الطبيعي بركام مخلفات البورسلين الناعم بنسب مختلفة وهي 0%، 7%، 14%، 21%، 28%. وتم التحقق من صلاحية استبدال الركام الناعم الطبيعي بركام البورسلين الناعم عن طريق اجراء مجموعة من الاختبارات الميكانيكية والفيزيائية للملاط مثل تحديد قابلية التشغيل ومقاومة الضغط والانثناء ،اضافة الى حساب كمية ثنائي اوكسيد الكاربون المنبعث من كل خلطة .اظهرت النتائج ان استبدال الركام الناعم الطبيعي بركام البورسلين الناعم ادى الى تحسن الخصائص الميكانيكية حيث ازدادت كل من مقاومة الانضغاط والانثناء بزيادة نسبة الاستبدال كما قلت نسبة ثنائي اوكسيد الكاربون المنبعث بشكل كبير عن ماهي عليه في الخلطة المرجعية . تؤكد الدراسة على الفوائد المزدوجة المتمثلة في خفض تكاليف البناء وتعزيز الإدارة المستدامة للنفايات، مما يجعل نفايات البورسلين مادة واعدة في السعي نحو ممارسات البناء الأخضر، أي البناء المستدام من خلال الحد من التدهور البيئي وتعزيز الاستخدام الدائري للنفايات.

الكلمات المفتاحية: الاستدامة، ركام مخلفات البورسلين الناعم، قابلية التشغيل، مقاومة الانضغاط، مقاومة الانثناء، التقييم البيئي

Abstract

This study aims to evaluate the feasibility of using porcelain tile waste as a component to replace fine aggregate in concrete mixes. This aims to achieve environmental, economic, and technical objectives in the construction sector. The study seeks to reduce waste generated from porcelain manufacturing and crushing, improve concrete properties, lower construction material costs, reduce carbon dioxide emissions from mining operations, and support the move towards sustainable construction in Iraq. Five mortar mixes were prepared by replacing natural fine aggregate with fine porcelain waste aggregate at different percentages: 0%, 7%, 14%, 21%, and 28%. The suitability of replacing natural fine aggregate with fine porcelain aggregate was verified by conducting a series of mechanical and physical tests on the mortar, such as determining workability, compressive strength, and flexural strength, in addition to calculating the amount of carbon dioxide emitted from each mix. The results showed that replacing natural fine aggregate with fine porcelain aggregate led to improved mechanical properties. Both compressive and flexural strength increased with increasing replacement percentages, while the carbon dioxide content decreased. The carbon dioxide emitted is significantly lower than in the reference mix. The study highlights the dual benefits of reducing construction costs and promoting sustainable waste management, making porcelain waste a promising material in the pursuit of green building practices—that is, sustainable construction through reducing environmental degradation and promoting the circular use of waste.

Keywords: Sustainability, Fine porcelain waste aggregate, Workability, Compressive strength, Flexural strength, Environmental assessment

المستخلص

الخرسانة المصنعة من مخلفات البورسلين

م.م. ريام ضياء محمد حسن[1]، ا م.د. انسام علي هاشم[2*]،
ا.م. هدى زهير عبد الغني[3]، ا.م.د. اوراس خضير عبيس[4]

3،2،1 قسم تقنيات الهندسة المدنية ، المعهد التقني بابل ، جامعة الفرات الاوسط التقنية، بابل \ العراق
4 قسم ميكانيك القدرة، المعهد التقني بابل، جامعة الفرات الاوسط التقنية، بابل \ العراق

Concrete Made from Porcelain Waste

Assist. Lect. Riyam D. Mohammed Alkafaji [1],
Assist Prof. Dr. Ansam Ali Hashim [2*],
Assist Prof . Huda Zuhair Kubba[3],
Assist Prof. Dr. Oras Khudhayer obayes[4]

1,2,3 Department of Civil Engineering, Technical Institute of Babylon, Al-Furat Al-Awsat Technical University, Babylon / /Iraq
4 Department of Power Mechanical Engineering, Technical Institute of Babylon, Al-Furat Al-Awsat Technical University, Babylon / /Iraq
(1) riyam.hassan.iba100@atu.edu.iq, (2) ansamly2@atu.edu.iq, (3) inb.huda@atu.edu.iq, (4) Oras.obais@atu.edu.iq
*** Corresponding Author**

المحتويات

تعليمات النشر
في مجلة كلية الاسراء الجامعة للعلوم الهندسية 5

الخرسانة المصنعة من مخلفات البورسلين 13

م.م. ريام ضياء محمد حسن، ا م.د. انسام علي هاشم،

ا.م. هدى زهير عبد الغني، ا.م.د. اوراس خضير عبيس

مجاورة الـ15- دقيقة بين المفهوم والممارسة
دراسة تحليلية لـ تطور المفهوم من البعد الحداثوي الى مابعد الجائحة 33

الباحثة فتوح كاظم الوسواسي ، و أ. م. د. حيدر جاسم عيسى الساعدي

(تعهد الملكية الفكرية)

إني الباحث..... صاحب البحث الموسوم (.....
.....)

أتعهد بأن البحث قد أنجز من قبلي ولم ينشر في مجلة أخرى في داخل وخارج العراق وأرغب بنشره في مجلة (**مجلة كلية الإسراء الجامعة للعلوم الهندسية**) التي تصدرها جامعة الإسراء.

التوقيع:

التاريخ:

(تعهد نقل حقوق الطبع والتوزيع)

إني الباحث..... صاحب البحث الموسوم (.....
.....)

أتعهد بنقل حقوق الطبع والتوزيع والنشر إلى مجلة (**مجلة كلية الإسراء الجامعة للعلوم الهندسية**) التي تصدرها جامعة الإسراء.

التوقيع:

التاريخ:

د- بحث في وقائع مؤتمر او ندوة علمية.

اسم الباحث أو الباحثون، (السنة)، عنوان البحث، اسم المؤتمر او الندوة العلمية، مكان الانعقاد، صفحتي البدء والانتهاء للبحث.

ترسل البحوث الى مجلة كلية الاسراء الجامعة للعلوم الهندسية على العنوان الاتي:

جامعة الاسراء– قسم التوثيق والنشر

بغداد / العراق

البريد الالكتروني:

al-esraajournal@esraa.edu.iq

7 - أن يؤشر المقيّم بشكل واضح على واحد من ثلاث اختيارات وهي:
البحث صالح للنشر بدون تعديلات.
البحث صالح للنشر بعد اجراء التعديلات.
البحث غير صالح للنشر.

8 - يجب أن يوضح المقيّم بورقة منفصلة ما هي التعديلات الأساسية التي يقترحها لغرض قبول البحث.

9 - للمقيّم حق طلب إعادة البحث إليه بعد إجراء التعديلات المطلوبة للتأكد من التزام الباحث بها.

10 -على المقيّم تسجيل اسمه ودرجته العلمية وعنوانه وتاريخ اجراء التقييم مع التوقيع على استمارة التقييم المرسلة له رفقه البحث المرسل له للتقييم.

المصادر

1 - يشار الى المصادر في متن البحث كما يلي:
اللقب او الاسم الثالث للمؤلف والسنة اذا كان البحث بإسم باحث واحد، واذا كان مؤلفين فيذكران والسنة واذا كانوا ثلاثة فاكثر فيذكر اسم الاول واخرون والسنة.

2 - ترتب المصادر حسب الصيغة العالمية (APA) وكما بالامثلة المذكورة:

أ- بحث في مجلة.
اسم الباحث أو الباحثون، (السنة)، عنوان البحث، اسم المجلة، المجلد، العدد و صفحتي البدء والانتهاء للبحث.

ب- كتب.
اسم المؤلف أو المؤلفون، (السنة) عنوان الكتاب، الطبعة، دار النشر وعدد الصفحات.

ج- الرسائل والاطاريح الجامعية.
اسم الباحث، (السنة)، عنوان الرسالة او الاطروحة، العنوان (الكلية والجامعة) وعدد الصفحات.

7 - يستعمل نظام الارقام العربية وهكذا في البحوث المرسلة للنشر وتمثل مناقشة النتائج تعبيراً موجزاً عن النتائج وتفسيراتها.

8 - تكون كتابة المصدر في القائمة المصادر متضمنة الآتي: اسم او أسماء الباحثين، سنة النشر وعنوان البحث كاملاً واسم المجلة ورقم المجلد والعدد وعدد الصفحات، مثال:

الخفاجي، جاسم محمود و حميد، محمد حسوني و كريم، حيدر حاتم، (2018) " دراسة تجريبية على الخرسانة مع استبدال جزئي للركام الخشن بواسطة المطاط غير المرغوب فيه".

مجلة كلية الاسراء الجامعة، المجلد 1 العدد 1، 217-243. و ممكن ان تكتب كالاتي: مجلة كلية الاسراء الجامعة، 1(1)،217-243.

9 - المستخلص الانكليزي يجب أن يكون وافياً ومعبراً عن البحث بصورة دقيقة وليس بالضرورة ان يكون ترجمة حرفية للمستخلص العربي و متبوعا بكلمات مفتاحية 4-6.

دليل المقيّم Reviewer Guidelines

أدناه الشروط والمتطلبات الواجب مراعاتها من قبل المقيم للبحوث المرسلة للنشر في هذه المجلة

1 - ملأ استمارة التقويم المرسلة رفقة البحث المطلوب تقييمه بشكل دقيق وعدم ترك أي فقرة بدون اجابة.

2 - على المقيّم التاكد من تطابق وتوافق عنوان البحث باللغتين العربية والانكليزية وفي حالة عدم تطابقهما اقتراح العنوان البديل.

3 - أن يبين المقيّم هل ان الجداول والاشكال التخطيطية الموجودة في البحث وافية ومعبرة.

4 - أن يبين المقيّم هل ان الباحث اتبع الاسلوب الإحصائي الصحيح.

5 - أن يوضح المقيّم هل ان مناقشة النتائج كانت كافية ومنطقية.

6 - على المقيّم تحديد مدى استخدام الباحث للمراجع العلمية الرصينة وحداثتها.

3 - يقدم البحث بثلاث نسخ ورقية ونسخة الكترونية بعد قبول البحث للنشر، يسلم البحث بشكله النهائي مطبوعا بالنظام الاعتيادي بمسافة منتظمة لكافة الصفحات عدا الصفحة الاولى التي تتضمن عنوان البحث و اسماء الباحثين وعناوينهم باللغتين العربية والإنكليزية متبوعا بالبريد الالكتروني للباحث الاول وعلى قرص مرن CD ببرنامج Microsoft Word / 2010.

4 - تقبل البحوث باللغتين العربية والانكليزية ويفضل كتابة البحث باللغة الانكليزية.

دليل المؤلف Author Guidelines

ادناه الشروط والمتطلبات الواجب مراعاتها من قبل الباحث للنشر في هذه المجلة بشرط أن لا يكون البحث قد نشر أو سينشر في أية مجلة هندسية أخرى ولم يمضِ على انجازه اكثر من أربع سنوات.

1 - يجب ان يكون عنوان البحث موجزاً قدر الامكان ومعبر عن البحث.

2 - اسماء الباحثين: تكتب اسماء الباحثين وعناوين عملهم بصورة واضحة مع البريد الالكتروني للباحث الاول.

3 - يجب ان يتضمن المستخلص موجزا واضحا عن البحث مكون من 250-300 كلمة متبوعا بكلمات مفتاحية 4-6. أذا كان البحث باللغة العربية فيكون المستخلص متبوعا بالكلمات المفتاحية اولا ثم المستخلص متبوعا بالكلمات المفتاحية باللغة الانكليزية ثانيا و العكس صحيح.

4 - المقدمة: تتضمن مراجعة المعلومات وثيقة الصلة بموضوع البحث الموجودة في المصادر العلمية وتنتهي المقدمة باهداف الدراسة وأساسها المنطقي.

5 - المواد وطرائق العمل: تذكر طرائق العمل بشكل مفصل ان كانت جديدة اما اذا كانت منشورة فتذكر بشكل مختصر مع الاشارة للمصدر وتستعمل وحدات النظام العالميSystem International of Units (S.I.Us)

6 - النتائج والمناقشة: تعرض بشكل موجز وهادف وبنظام متوالي وتعرض النتائج بافضل صورة معبرة وتوضع الجداول والاشكال في أماكنها المخصصة بعد الاشارة إليها في النتائج.

تعليمات النشر

في مجلة كلية الاسراء الجامعة للعلوم الهندسية

- تصدر جامعة الاسراء (مجلة كلية الاسراء الجامعة للعلوم الهندسية) في مجلد سنوي يضم عددين.
- تقوم المجلة بنشر البحوث العلمية للباحثين في تخصصات العلوم الهندسية التالية:
  - هندسة العمارة
  - هندسة مدني
  - هندسة كيمياوية
  - هندسة الحاسوب
  - هندسة كهربائية
  - هندسة المواد
  - هندسة ميكانيكية

شروط النشر

1 - يطبع البحث بواسطة الحاسوب بمسافات مفردة بين الاسطر وبحجم خط 12 ونوع (Simplified Arabic)، اما العنوان باللغتين العربية والانكليزية فيكون بحجم خط 14 شريطة الا يزيد عدد صفحاته عن 15 صفحة بما في ذلك الجداول والاشكال والمراجع وعلى وجه واحد على ورق قياس A4 مع ترك هامش في حدود 2 سم من الاعلى والاسفل وهامش بحدود 3 سم من الجانبين الايمن والايسر.

2 - لا يفضل نشر البحوث من قبل رئيس واعضاء هيئة التحرير في المجلة سواء كان البحث منفرداً أو مشتركاً.

- **الأستاذ المشارك محنوش بيجلاري**.....جامعة الرازي\ إيران
- **الأستاذ المشارك حسن موراديي** .....جامعة الرازي\ إيران
- **الأستاذ المشارك إيمان أشابوزي**.....جامعة الرازي\ إيران

المراجعة اللغوية

- **أ. د. غالب فاضل المطلبي**.....جامعة الإسراء \ العراق.
- **أ. م. د. سعد فاضل الحسني**.....جامعة الإسراء\ العراق.

السلامة الفكرية

- **أ. م. د. أكرم علي عنبر** .....مساعد رئيس جامعة الإســـراء للشؤون الادارية\ العراق
- **م. د. محمد جبار الشمري** .....جامعة الإسراء \ العراق
- **م. م. آن سماري إبراهيم** .....جامعة الإسراء \ العراق

المسؤول المالي

- **م. م. بشار قاسم تعيب**.....جامعة الإسراء \ العراق.

رئيس هيئة التحرير

- **أ. د. عبد الرزاق جبر الماجدي** .....رئيس جامعة الإسراء \ العراق

مدير التحرير

- **أ. م. د. أحسان علي صائب الشعرباف** ..... جامعة الإسراء \ كليةالهندسة، قسم الهندسة مدنية \ العراق

هيئة التحرير

- **أ. د. موسى عزيز الموسوي** .....مستشار \ وزارة التعليم العالي والبحث العلمي \ العراق
- **أ. د. عباس محسن البدري** .....رئيس جامعة تكنولوجيا المعلومات والاتصالات \العراق
- **أ. د. ثامر خضير محمود** .....جامعة الإســراء \ كلية الهندسة، قسم الهندسة المدنية \ العراق
- **أ. د. رياض مهدي المهيدي** ..... جامعة سونبرن \ أستراليا \ هندسة مدنية
- **أ. د. مثنى حكمت الدهان**.....جامعة ميزوري \ امريكا \ هندسة ميكانيكية
- **أ. د. رمزي محمد محمود** ..... جامعة بنسلفانيا \ امريكا \ هندسة مدنية
- **أ. د. حسين الرزو**.....جامعة آركنساس \ امريكا \ هندسة إلكترونية
- **أ. د. عبد الرزاق طارش زبون العبودي**.... جامعة الإسراء\كلية الهندسة التقنية\ العراق
- **أ. م. د. كاظم عبود الماجدي** .....الجامعة المستنصرية \ هندسة كيمياوية \ العراق
- **أ. م. د. رياض عزيز الموسوي** ..... جامعة الإسراء \ هندسة مدنية \ العراق
- **أ. م. د. صباح ناصر حسن** ..... جامعة الإسراء \هندسة إلكترونية \ العراق
- **أ.م.د. محمد غازي صبري**.....جامعة الإسراء\قسم الشؤون العلمية\العراق
- **أ. م. د. عبد الناصر علك حافظ** ..... وزارة التعليم العالي و البحث العلمي \ العراق
- **م. د. إياد احمد الطويل** ..... جامعة الإسراء \ العراق
- **م.د. حيدر صادق الأعسم**.....جامعة الإسراء\كلية الهندسة التقنية\العراق

بسم الله الرحمن الرحيم

جمهورية العراق
وزارة التعليم العالي والبحث العلمي
دائرة البحث والتطوير

Republic of Iraq
Ministry of Higher Education & Scientific Research
Research & Development Department

الرقم: ب ت 4 / 5749
التاريخ: 2021/09/06

No.:
Date:

١٠٦٠
٢٠٢١/٩/٨

كلية الاسراء الجامعة / السيد العميد المحترم

م/ مجلة كلية الاسراء الجامعة للعلوم والهندسة

السلام عليكم ورحمة الله وبركاته ...

أشارة الى كتابكم المرقم م.ع/٢٣٩٨ في ٣١ / ١٢/ ٢٠٢٠ بشأن اعتماد مجلتكم واعتمادها لأغراض النشر والترقيات العلمية وتسجيلها ضمن موقع المجلات الاكاديمية العلمية العراقية ، حصلت موافقة السيد وكيل الوزارة لشؤون البحث العلمي بتاريخ ٢٠٢١/٨/٢٤ على أعتماد المجلة المذكورة في الترقيات العلمية والنشاطات العلمية المختلفة الأخرى ، واعتباراً من المجلد الثالث – العدد الثالث – لسنة ٢٠٢١ وتسجيل المجلة في موقع المجلات الاكاديمية العلمية العراقية.

للتفضل بالاطلاع وابلاغ مخول المجلة لمراجعة دائرتنا لتزويده بإسم المستخدم وكلمة المرور ليتسنى له تسجيل المجلة ضمن موقع المجلات العلمية العراقية وفهرسة اعدادها ... مع التقدير.

أ.م.د يوسف خلف يوسف
ء/ المدير العام لدائرة البحث والتطوير
٢٠٢١/٩/٦

نسخة منه الى:
- مكتب السيد وكيل الوزارة لشؤون البحث العلمي / أشارة الى موافقة سيادته المذكورة أعلاه والمثبتة على اصل مذكرتنا المرقم ب ت م٤/ ٤٥٧٦ في ٢٠٢١/٨/٢٣ / للتفضل بالاطلاع ... مع التقدير.
- قسم المشاريع الريادية / شعبة المشاريع الالكترونية / للتفضل بالعلم واتخاذ مايلزم ... مع التقدير
- قسم الشؤون العلمية / شعبة التاليف والنشر والترجمة / مع الاوليات .
- الصادرة .

مهند ابراهيم
٦ / أيلول

وزارة التعليم العالي والبحث العلمي – دائرة البحث والتطوير – القصر الأبيض – المجمع التربوي – الطابق السادس
قسم الشؤون العلمية sedep@rdd.edu.iq

مجلة كلية الإسراء
الجامعة للعلوم الهندسية

مجلة علمية محكمة تصدر عن جامعة الإسراء

رقم الايداع في دار الكتب والوثائق ببغداد (2445) لسنة (2020)
الرقم الدولي للنسخة الورقية (ISSN : 2709 - 7145)
الرقم الدولي للنسخة الإلكترونية (E-ISSN: 2790-7732)

جامعة الإسراء

(سُبْحَانَ الَّذِي أَسْرَى بِعَبْدِهِ لَيْلًا مِّنَ الْمَسْجِدِ الْحَرَامِ إِلَى الْمَسْجِدِ الْأَقْصَى)

E U

AL- Esraa University

م 2011 – 1432 هـ

المجلد 8 ـ العدد 13 ـ حزيران لسنة 2026