

JOURNAL OF SCIENCE

PART A: ENGINEERING AND INNOVATION



Year | 2025

Volume | 12

Issue | 2

e-ISSN 2147-9542



| Owner |

on behalf of Gazi University
Rector
Prof. Dr.

Uğur ÜNAL

| Publishing Manager |

Prof. Dr.

Uğur GÖKMEN
Gazi University

| Chief Editor |

Prof. Dr.

Sema BİLGE OCAK
Gazi University

| Chief Editor |

Prof. Dr.

Mustafa Gürhan YALÇIN
Akdeniz University

| Assistant Editors |

Assoc. Prof. Dr.

Defne AKAY
Ankara University

Assoc. Prof. Dr.

Murat BİLEN
Gazi University

| Editorial Board |

Prof. Dr. Adem TATAROĞLU	Gazi University Physics
Prof. Dr. Adnan SÖZEN	Gazi University Energy Systems Engineering
Prof. Dr. Ali KESKİN	Çukurova University Automotive Engineering
Prof. Dr. Ali Osman SOLAK	Ankara University Chemistry
Prof. Dr. Alper BÜYÜKKARAGÖZ	Gazi University Civil Engineering
Prof. Dr. Atilla BIYIKOĞLU	Gazi University Mechanical Engineering
Prof. Dr. Aynur KAZAZ	Akdeniz University Civil Engineering
Prof. Dr. Cevriye GENCER	Gazi University Industrial Engineering
Prof. Dr. Çağlayan AÇIKGÖZ	Bilecik Şeyh Edebali University Chemical Engineering
Prof. Dr. Çetin ÇAKANYILDIRIM	Hitit University Chemical Engineering
Prof. Dr. Demet CANSARAN DUMAN	Ankara University The Institute of Biotechnology
Prof. Dr. Elif ORHAN	Gazi University Physics
Prof. Dr. Erdal IRMAK	Gazi University Electrical-Electronic Engineering
Prof. Dr. Fatih ÖZ	Atatürk University Food Engineering

| Editorial Board |

Prof. Dr. Feyza DİNÇER	Nevşehir Hacı Bektaş Veli University Geological Engineering
Prof. Dr. Gürhan İÇÖZ	Gazi University Mathematics
Prof. Dr. Hacer KARACAN	Gazi University Computer Engineering
Prof. Dr. Hüseyin Serdar YÜCESU	Gazi University Automotive Engineering
Prof. Dr. Meltem DOĞAN	Gazi University Chemical Engineering
Prof. Dr. Metin GÜRÜ	Gazi University Chemical Engineering
Prof. Dr. Mine TÜRKTAŞ	Gazi University Biology
Prof. Dr. Murat KAYA	Aksaray University Biotechnology and Nanotechnology
Prof. Dr. Nalan KABAY	Ege University Chemical Engineering
Prof. Dr. Nazife ASLAN	Ankara Hacı Bayram Veli University Chemistry
Prof. Dr. Niyazi Uğur KOÇKAL	Akdeniz University Civil Engineering
Prof. Dr. Nuran AY	Eskişehir Technical University Materials Science and Engineering
Prof. Dr. Nurdane İLBEYLİ	Akdeniz University Geological Engineering
Prof. Dr. Nursel AKÇAM	Gazi University Electrical-Electronic Engineering

| Editorial Board |

Prof. Dr. Ömer ŞAHİN	İstanbul Technical University Chemical Engineering
Prof. Dr. Selim ACAR	Gazi University Physics
Prof. Dr. Şükrü DURSUN	Konya Technical University Environmental Engineering
Prof. Dr. Veli ÇELİK	Ankara Yıldırım Beyazıt University Mechanical Engineering
Prof. Dr. Yılmaz ŞİMŞEK	Akdeniz University Mathematics
Prof. Dr. Yusuf URAS	Kahramanmaraş Sütçü İmam University Geological Engineering
Prof. Dr. Yücel ERCAN	TOBB University of Economics and Technology Mechanical Engineering
Prof. Dr. Zafer EVİS	Middle East Technical University Engineering Sciences
Assoc. Prof. Dr. Fusun YALÇIN	Akdeniz University Mathematics
Assoc. Prof. Dr. Yasemin LEVENTELİ	Akdeniz University Geological Engineering
Assist. Prof. Dr. Emine Şükran OKUDAN	Akdeniz University Basic Sciences Fisheries
Assist. Prof. Dr. Senai YALÇINKAYA	Marmara University Mechanical Engineering
Dr. Murat AKIN	Gazi University Computer Technologies
Dr. Silver GÜNEŞ	Gazi University Chemical Engineering



| Foreign Editorial Advisory Board |

Prof. Dr. Abdelmejid BAYAD	Université d'Évry Val d'Essonne FRANCE
Prof. Dr. Ali Behcet ALPAT	Istituto Nazionale di Fisica Nucleare (INFN) ITALY
Prof. Dr. Azra SPAGO	Dzemal Bijedic University of Mostar BOSNIA AND HERZEGOVINA
Prof. Dr. Bektay YERKIN	Satbayev University KAZAKHSTAN
Prof. Dr. Burçin BAYRAM	Miami University USA
Prof. Dr. Daeyeoul KIM	Jeonbuk National University SOUTH KOREA
Prof. Dr. Elvan AKIN	Missouri University of Science and Technology USA
Prof. Dr. Filiz DİK	Rockford University USA
Prof. Dr. Homer RAHNEJAT	Loughborough University UNITED KINGDOM
Prof. Dr. Lokesha VEERABHADRIAH	Vijayanagara Sri Krishnadevaraya University INDIA
Prof. Dr. Mehmet DİK	Rockford University USA
Prof. Dr. Nedim SULJIĆ	University of Tuzla BOSNIA AND HERZEGOVINA
Prof. Dr. Rob DWYER-JOYCE	The University of Sheffield UNITED KINGDOM
Prof. Dr. Snezana KOMATINA	University Novi Sad SERBIA
Prof. Dr. Toni NIKOLIC	University Džemal Bijedić Mostar BOSNIA AND HERZEGOVINA
Prof. Dr. Turysbekova Gaukhar SEYTKHANOVNA	Satbayev University KAZAKHSTAN
Assoc. Prof. Dr. Burcin ŞİMŞEK	Associate Director Biostatistics (Oncology) at Bristol Myers Squibb USA
Assist. Prof. Dr. Alisa BABAJIC	University of Tuzla BOSNIA AND HERZEGOVINA
Dr. Daniel Ganyi NYAMSARI	Mining Company Researcher CAMEROON



| English Language Editors |

Lecturer	Gazi University
Gizem AÇELYA AYKAN	School of Foreign Languages

Lecturer	Gazi University
Tuğçe BÜYÜKBAYRAM	School of Foreign Languages

| Technical Editors |

Research Assistant Dr.	Akdeniz University
Fatih UÇAR	

Research Assistant	Akdeniz University
Özge ÖZER ATAĞOĞLU	



| Correspondence Address |

Gazi University Graduate School of Natural and Applied Sciences
Emniyet Neighborhood, Bandırma Avenue No:6/20B, 06560,
Yenimahalle - ANKARA
B Block, Auxiliary Building

| e-mail |

gujsa06@gmail.com

| web page |

<https://dergipark.org.tr/tr/pub/gujsa>

Gazi University Journal of Science Part A: Engineering and Innovation
is a peer-reviewed journal.

| INDEXING |

TRDIZIN

ASOS
indeks

 CiteFactor
Academic Scientific Journals

 Crossref

 DRJI

 EBSCOhost

 ESJI
Eurasian Scientific
Journal Index
www.ESJIndex.org

 HARVARD
LIBRARY

INDEX  COPERNICUS
INTERNATIONAL

 INFOBASE INDEX

 Academic
Resource
Index
ResearchBib

 ROAD
DIRECTORY OF OPEN ACCESS
SCHOLARLY RESOURCES

 SCI

 Scientific Indexing Services

 Scilit

 scinapse

 SOBIAD

| ACCESSIBILITY |

 Google
Scholar

 DergiPark
AKADEMİK



This work are licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.



| CONTENTS |

Page | Articles

Evaluation of Chemical Organizations According to The Root Causes of Industrial Accidents with Analytical Hierarchy Process

347-372

S. UNVERDI  S. CETINYOKUS  T. CETINYOKUS  E. CALISKAN 

Chemical Engineering

Research Article

[10.54287/guj.1621519](https://doi.org/10.54287/guj.1621519)

Deep Learning-Based Lung Cancer Diagnosis: Data Balancing, Model Optimisation and Performance Analysis

373-391

F. ALPSALAZ 

Electrical & Electronics Engineering

Research Article

[10.54287/guj.1648772](https://doi.org/10.54287/guj.1648772)

Improved Chef-Based Optimization Algorithm with Chaos-Based Fitness Distance Balance for Frequency-Constrained Truss Structures

392-416

A. BEŞKİRLİ 

Computer Engineering

Research Article

[10.54287/guj.1667182](https://doi.org/10.54287/guj.1667182)

Proposing a Five-Phase Framework Based on ISO 23247-1 for Digital Twins in Construction

417-444

M. AYDIN 

Civil Engineering

Research Article

[10.54287/guj.1680674](https://doi.org/10.54287/guj.1680674)

A Novel Technique for Criteria Weighting in Multi-Criteria Decision-Making: Tanimoto Contrast Approach (TCA)

445-478

F.F. ALTINTAŞ 

Statistics

Research Article

[10.54287/guj.1673755](https://doi.org/10.54287/guj.1673755)

| CONTENTS |

Page | Articles

Comparison of Deep Learning Algorithms for Image Segmentation on Satellite Images

479-502 H. ACEMLI  N. KUMBASAR 

Computer Engineering

Research Article

[10.54287/gujsa.1664093](https://doi.org/10.54287/gujsa.1664093)

Removal of Boron and Boron Compounds from Wastewaters

503-540 İ. ÇELEBİ GÜRSOY  İ. AR 

Chemical Engineering

Research Article

[10.54287/gujsa.1660364](https://doi.org/10.54287/gujsa.1660364)

Optimizing Choke Size to Minimize Sand Production in Oil Wells: A Machine Learning Approach

541-561 S. AGBONS IGBINERE  I. OHENHEN  E.F. CHRISTOPHER 

Petroleum Engineering

Research Article

[10.54287/gujsa.1669814](https://doi.org/10.54287/gujsa.1669814)

Some Applications of Berezin Radius Inequalities

562-582 H. BAŞARAN 

Mathematical Sciences

Research Article

[10.54287/gujsa.1655950](https://doi.org/10.54287/gujsa.1655950)

Strength and Microstructural Analysis of Nano Powder and Basalt Fibre Reinforced RCC Pavement Concrete

583-607 Ş. SÜNBÜL  A. TORTUM 

Materials Engineering

Research Article

[10.54287/gujsa.1647353](https://doi.org/10.54287/gujsa.1647353)

| CONTENTS |

Page | Articles

Analysis and Design of a W-Band MMIC LNA with 4.2 dB Noise Figure in GaAs-based 100 nm pHEMT Technology

608-618

G.O. ARICAN  B.A. YILMAZ 

Electrical & Electronics Engineering

Research Article

[10.54287/guj.1662033](https://doi.org/10.54287/guj.1662033)

On the Structural, Electric and Optic Properties of B_{1-x}Al_xN Alloys Using Ab Initio Calculation

619-631

Ö. BAYAL  A. GÜLTEKİN  A.K. BILGİLİ  M.K. ÖZTÜRK 

Physical Sciences

Research Article

[10.54287/guj.1646889](https://doi.org/10.54287/guj.1646889)

Forecasting Natural Gas Consumption by User Type Using Machine Learning: A Comparative Study

632-651

S. TEKİN  S. PEKER  N. DUDAKLI 

Information and Computing Sciences

Research Article

[10.54287/guj.1689278](https://doi.org/10.54287/guj.1689278)

Utilization of Electromyographic Signal Classification for Predicting Hand Gestures and Grasping Forces Through Artificial Neural Networks

652-664

D. KARABULUT  S.C. DOĞRU 

Mechanical Engineering

Research Article

[10.54287/guj.1700850](https://doi.org/10.54287/guj.1700850)

Effect of Yolo-Based Vehicle Detection and Manuel Vehicle Classification Methods on Simulation-Based Intersection Analysis

665-690

B.E. ULUSOY  D. HAZAR  Y. ALBAYRAK  S. KOFTECI  M.A. YALCIN  A. DEMİRAL 

Civil Engineering

Research Article

[10.54287/guj.1647168](https://doi.org/10.54287/guj.1647168)



Gazi University

Journal of Science

PART A: ENGINEERING AND INNOVATION

<http://dergipark.org.tr/guj.1621519>

Evaluation of Chemical Organizations According to The Root Causes of Industrial Accidents with Analytical Hierarchy Process

Sehmus UNVERDI¹ Saliha CETINYOKUS^{2*} Tahsin CETINYOKUS³ Emre CALISKAN³

¹ Gazi University, Graduate School of Natural and Applied Sciences, Dep. of Environmental and Technical Research of Accidents, Ankara Türkiye

² Gazi University, Faculty of Engineering, Chemical Engineering, Ankara Türkiye

³ Gazi University, Faculty of Engineering, Industrial Engineering, Ankara Türkiye

Keywords	Abstract
Industrial Accidents	In this study, it was aimed to evaluate the accident risk of chemical organizations according to the root causes of industrial accidents with the Analytical Hierarchy Process (AHP). First of all, the data of occupational accidents with death and loss of limb in the chemical industry between 2015 and 2020 were evaluated with Pareto analysis by using the statistics of the Social Security Institution (SSI). As a result of the analysis, the main criteria that make up 80% of the accidents in the sector were obtained. Then, the final main and sub-criteria that could be the root cause of industrial accident were determined, and these criteria were weighted over the opinions of the relevant experts using AHP method. Human errors, one of the main criteria, were determined as the most important criterion with 35%. By applying AHP for the second time, three sample high-level organizations were evaluated according to the root causes of industrial accidents. Organizations showed a distinctive ranking in terms of industrial accident risk (Organization B= 0.420> Organization A= 0.354 >Organization C= 0.226). With the proposed methodology, quantitative and qualitative evaluation criteria were included in the model simultaneously, and an objective result was obtained through expert opinions.
OHS	
Pareto Analysis	
MCDM	
AHP	

Cite

Unverdi, S., Cetinyokus, S., Cetinyokus, T., & Caliskan, E. (2025). Evaluation of Chemical Organizations According to The Root Causes of Industrial Accidents with Analytical Hierarchy Process. *GU J Sci, Part A, 12(2)*, 347-372. doi:10.54287/guj.1621519

Author ID (ORCID Number)	Article Process
0000-0002-1572-5843	Submission Date 16.01.2025
0000-0001-9955-6428	Revision Date 24.03.2025
0000-0002-9963-5174	Accepted Date 08.05.2025
0000-0002-6262-7197	Published Date 30.06.2025

1. INTRODUCTION

Today, in addition to making life easier, products containing risks against the environment and human health are produced by industrial establishments. As a result of the difficulties experienced in controlling these emerging risks, major industrial accidents occur where negative effects can be seen in the short, medium and long term. Prevention of industrial accidents, minimizing their effects and measures to intervene in the accident are extremely important and the root causes of the accidents should be analyzed. There is no generally accepted definition of what root cause analysis is in the literature. The possible definition can be made as “structured research aimed at determining the actual cause of a problem and the actions necessary to eliminate it”. Root cause analysis cannot be performed using a single tool or strategy. Root cause analyzes can be grouped into 7 subgroups according to their purposes (Andersan and Fagerhaug, 2006):

1. Understanding the problem
2. Identifying ideas about the cause of the problem
3. Data collection
4. Data analysis
5. Root cause identification
6. Eliminate/remove root causes
7. Implement a solution

Accidents should be treated as a problem to be solved, as a state of difficulty. When attempting to solve a problem, one must identify the cause (or causes) of the problem and find ways to eliminate those causes and prevent it from recurring. Once you have identified the causes, eliminating them often becomes a much easier task. Zhang et al. (2020) developed a methodology for the detection of root causes of accidents in coal mines, using the characteristics of safety culture deficiencies based on accident statistics. Emphasis was placed on the role of departments and the audit environment in developing a safety culture to reduce industrial accidents. Botti et al. (2020) analyzed 118 occupational accidents that occurred in Italy between 2002 and 2015 on the root causes of accidents in the manufacturing sector. The apparent cause was shown to be voluntary adoption of an inappropriate procedure. Alekseeva et al. (2020) statistically examined the accidents in pipeline gas transport establishments in order to determine the root causes of the accidents. A functional relationship was established between the pipe manufacturer's faults and the number of corrosion-related accidents and latent defects, including defects in construction and installation work. Ghatorda et al. (2020) applied root cause analysis to increase productivity in the press manufacturing industry. Flange fixture was developed with the fishbone method and machining time was reduced by approximately 9.5 hours. Wang et al. (2018) developed a symbolic transfer entropy method with binary encoding and decimal decoding to increase resistance to noise and reduce data loss while creating a data model for complex electromechanical systems. Li et al. (2016) proposed a root cause framework, including both fixed errors and non-stationary errors, for the data-driven identification of root causes of errors in process industries. A case study on the Tennessee Eastman process was applied and the proposed framework was shown to be valid. In the study carried out by Yilmaz et al. (2016), it was emphasized that occupational health and safety was a system at both the industrial and public level, and the inadequacy of any stage of this system would reduce the level of prevention of the problem by other measures.

The root causes of industrial accidents, such as occupational accidents, can be very diverse. In addition to human errors, process, equipment, environmental effects, etc. factors must be taken into account. The problems of choosing the most suitable one among the alternatives or ranking the alternatives by considering more than one criterion are called multi-criteria problems. MCDM methods are used to choose the best decision problems

according to a wide range of concrete and intangible criteria or qualities and to ensure that decisions are made in cases where there is more than one criterion. There are a limited number of studies in the literature on the analysis of root causes with MCDM methods. Ozdemir et al. (2018) used MCDM methods regarding occupational accidents on board and indicated the priority order of the causes of occupational accidents: human factors, lack of management, ship-related problems, cargo problems and environmental factors. Wang et al. (2013) emphasized the importance of understanding the root causes in order to take effective preventive measures after the accident in their accident analysis model studies using MCDM methods. Human Factors Analysis and Classification System (HFACS) and Evidence-Based Reasoning (ER) MCDM methods were proposed in order to investigate the root causes of the accident and to rank the related prevention measures. It was seen that MCDM methods are widely used in risk assessment studies of various occupational accidents. By Yucesan and Gul (2021), probability and severity parameters of hazards in an aluminum plate manufacturing establishment were weighted using the Fuzzy Analytic Hierarchy Process (FAHP), and then the priority order of 23 different hazard groups was determined using Fuzzy TOPSIS (Technique for Order Preference by Similarity). It was shown that the three most important hazard groups for the organization were suffocation from gases, electric shock and falling objects, respectively. In another study by Gul (2020), Pythagorean fuzzy AHP and fuzzy VIKOR analyzes were applied in the gun and rifle barrel outer surface oxidation and coloring unit in occupational health and safety risk assessment. The proposed approach produced reliable results that better represent the uncertainty of the decision-making process. It was stated by Dabous et al. (2021) that electronics manufacturing was the most common application area for the integration between MCDM and FMEA, and the TOPSIS method was the most applied MCDM approach in the manufacturing industry. Fata et al. (2021) applied the VIKOR method by ranking the occupational health and safety risks including human factors. The proposed method was applied in a company that sells furniture and produces wooden products in Sicily, and a high success was achieved in distinguishing the risks. Bragatto et al. (2021) investigated the impact of the COVID-19 pandemic on the Italian chemical and process industries to which the Seveso III Directive applies. Using the AHP method, an innovative organizational “resilience” model was proposed, aiming to develop a higher capacity to face similar new crises in the future. A risk analysis model was developed with Analytical Network Process (ANP) and TOPSIS methods by Ciftci et al. (2020). The model was applied in the metal industry, providing a sequencing and interrelationships between risk factors to identify significant risks. In addition, not only traditional risk factors, but also psycho-social factors were discussed in the study. Viegas et al. (2020) performed a multi-criteria hazard and operability analysis for process safety. In practice, 40 hazards were defined; resource savings, greater focus, objectivity and a more realistic perception of danger were achieved. Oz et al. (2019) applied an extended TOPSIS model with Pythagorean fuzzy sets to prioritize hazards in a natural gas pipeline. A sensitivity analysis was also performed on the parameter weights to validate the model. Turskis et al. (2019) used the MCDM method to identify the most effective people in preventing accidents in construction Small and Medium Enterprises (SME)s. It was determined that key stakeholders, including the developer, project leader, OHS specialist, OHS coordinator, contractor and subcontractor, were the most effective people to prevent occupational accidents at SME construction sites.

Koulinas et al. (2019) stated that it would be beneficial to use a popular multi-criteria decision-making method and a quantitative risk analysis technique together to rank and prioritize risks, as a result of their work with the MCDM method in the construction industry. Gul (2018) reviewed the occupational health and safety risk assessment approaches based on MCDM methods and their fuzzy versions. It is stated that MCDM methods are widely applied for risk assessment as quantitative tools and will provide an understanding of the risk assessment process in hazardous industries. Seker et al. (2017) implemented MCDM methods in the analysis of occupational risks at construction sites. It was stated that traditional methods were not useful for solving human-centered problems due to the complexities arising from human factors, although they offered precise solutions. It has been noted by Ouédraogo et al. (2011) that MCDM methods provide a rigorous mathematical model for performing risk analysis. Ilbahar et al. (2018) proposed a new approach to risk assessment for occupational health and safety using integrated fuzzy AHP and Fine Kinney analyses. In the study by Heller et al. (2006), it was stated that MCDM methods together with risk scoring/indexing were methods used to measure subjective and objective judgments, and that the experience and knowledge of people were at least as valuable as the data itself.

It was seen that the AHP method came to the fore in risk assessment studies. AHP method is advantageous with the aspects listed below (Aycin, 2019).

- Ability of decision makers to create solutions by considering their personal preferences,
- Multiple main and sub-criteria can be used in decision making problems,
- Ease of calculation,
- Allowing multiple decision makers to make joint decisions,
- Facilitating the problem by presenting the complex problems in a hierarchical structure

Although AHP is a widely used analysis in risk analysis and risk assessment processes, it was applied for the first time in the analysis of root causes of industrial accidents in this study. Studies using AHP in the analysis of root causes of other accidents are also quite limited. The evaluation of the industrial accident risk of the organizations through root causes was carried out with the AHP method, which is prominent in the literature and is suitable for our problem. In this context, the criteria that may be the root cause of industrial accidents were determined by taking into account the Pareto analysis of the occupational accidents data with death and limb loss between 2015-2020 in workplaces operating in the chemical industry. With the application of the sequential AHP method based on expert opinions, sample industrial organizations were listed in terms of industrial accident risk.

2. MATERIAL AND METHOD

The evaluation of the organization's accident risk according to the root causes of industrial accidents with the AHP method was carried out according to the methodology presented in Figure 1.

In order to determine the main and sub-criteria that may be the root cause of industrial accidents and to realize the problem understanding stage of the root cause analysis, the "Communiqué on Occupational Health and Safety on Workplace Hazard Classes" was taken into consideration at the first stage. 19.xx.xx – 22.xx.xx NACE coded workplaces working in the chemical industry were selected from the Six Economic Activity Classification (NACE) codes. By using Social Security Institution (SSI) statistics, data on occupational accidents with death and limb loss between 2015-2020 in selected workplaces were collected and critical events were tried to be determined. Then, in order to determine the ideas about the cause of the problem, the main and sub-criteria were formed by brainstorming, taking into account the expert opinions as well as the literature research. The data collected to perform the data analysis process of root cause analysis were evaluated with Pareto analysis. After the Pareto analysis, the main criteria and sub-criteria that could be the root cause of industrial accidents were determined. The weighting of the criteria was done through AHP method and questionnaires. The questionnaire was applied to 6 experts from the Ministry of Labor and Social Security, Ministry of Environment and Urbanization, Disaster and Emergency Management Presidency, industry and university, who have roles and responsibilities within the scope of the relevant legislation. At the last stage, 3 industrial organizations within the scope of the legislation were evaluated in terms of industrial accident risk according to root causes by applying again the AHP method, and the most risky organization was determined.

2.1. Pareto Analysis

Pareto Analysis was developed by the Italian Vilfredo Pareto (1848-1923). Pareto has carried out various researches in businesses and stated that the majority of the problems are usually caused by a small number of interconnected but dominant causes. According to the Pareto analysis, which is also called the "80-20 Rule" in the literature, it is concluded that "80% of the problems should be caused by 20% activity and this important 20% should be focused on". This represents the philosophy of "apparent majority, effective minority" (Aycin, 2019).

In order to carry out the data collection process of root cause analysis, workplaces with the code 19.xx.xx – 22.xx.xx NACE were selected from the Six Economic Activity Classification (NACE) codes, taking into account the "Communiqué on Workplace Hazard Classes on Occupational Health and Safety". By using Social Security Institution (SSI) statistics, occupational accidents data with death and limb loss between 2015-2020 in selected workplaces were considered. Accidents were grouped according to the European Statistics on Accidents at Work - ESAW methodology on the basis of declared causes, and accidents in other deviation

codes were not taken into account when grouping (code 99). The causes of the accident were evaluated by the Pareto Analysis method and the 80% part was determined.

The data on occupational accidents with death and limb loss in 19.xx.xx – 22.xx.xx NACE coded workplaces working in the chemical industry are presented in Table 1.

According to the data in Table 1 and Figure 2, it was seen that the cause of the accident “acting without cooperation, unnecessary or untimely actions” constituted 15% (3 550) of 33 321 fatal and limb loss occupational accidents. In the analyzed statistics, it was shown that there were 9 367 (other) accident causes for which no accident cause was entered.

According to the Pareto approach, 80% of the accidents can be solved by eliminating 20% of the causes of these accidents. It was seen from Figure 2 that the first two causes of accidents in the sector reached 26%. Some of them were listed below:

- Action without cooperation (64),
- Loss of control (full or partial) - the hand tool (whether powered or not) or the material tool used by the tool (43),
- Slip - stumble and fall - Person fall - on the same level (52),
- Loss of control (full or partial) - machine (including unintentional start) or material tool (41).

The Organization will deal with these problems in order to the best of its ability. This list will contribute to the manager in terms of which reasons to focus on. When the results from the Pareto analysis were classified on the basis of 4 main criteria, people, management, process and equipment errors were considered as the main criteria. In order to strengthen the hand of the manager at the point of decision-making, the criteria that will be the main causes of accidents in the chemical industry were investigated through literature research and expert opinions. Dominoes, disasters and sabotage were added to the main criteria. In total, seven main criteria were reached. After the main criteria were determined, the sub-criteria for each main criterion were determined by taking the literature research and expert opinions again.

2.2. AHP Method

After determining the main and sub-criteria of the industrial accident root cause, analyzes were made with MCDM methods. In the study, the AHP method was chosen from MCDM methods because it allows to determine the relative weights of the criteria as well as its superior aspects. Various multi-criteria decision making methods are available in the literature. AHP method, which is one of these methods, stands out and is frequently used (Aycin, 2019; Ozguvenç, 2011; Dincer and Gorener, 2011).

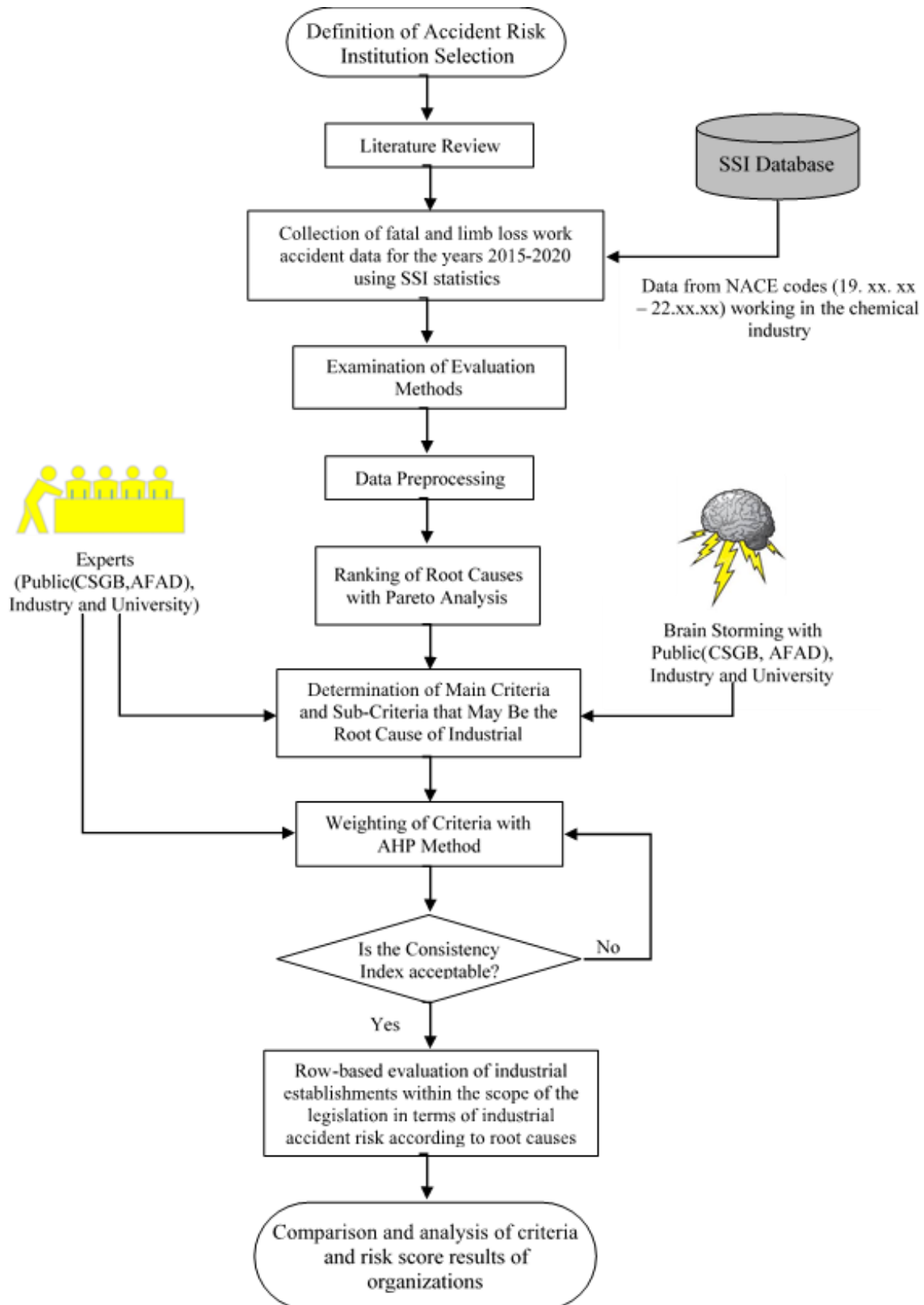
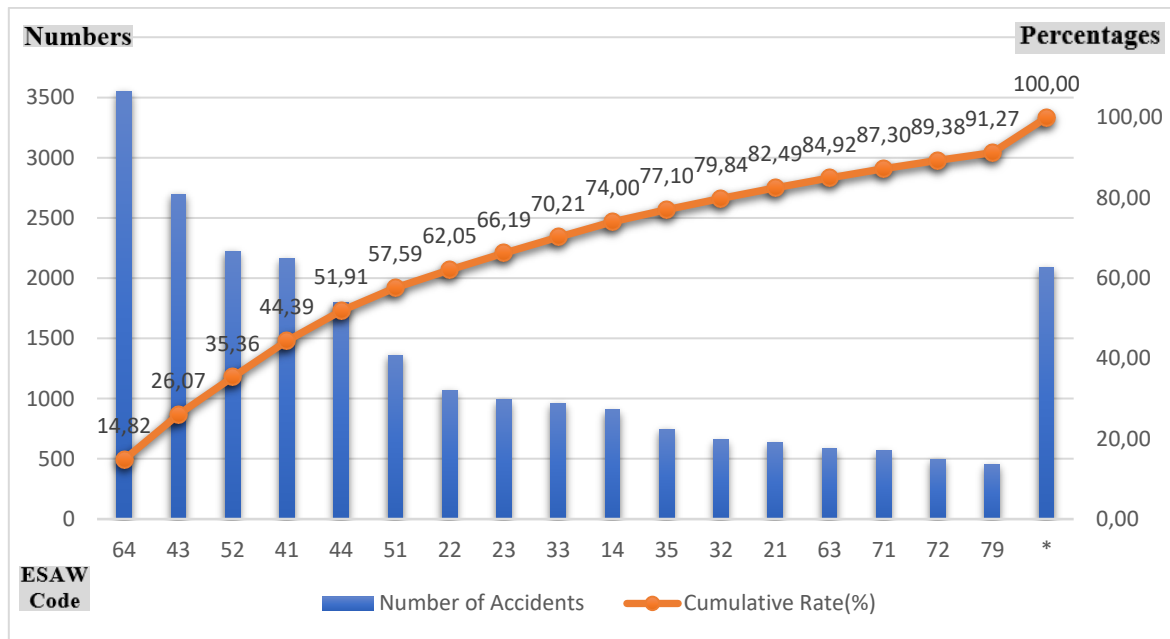


Figure 1. Flowchart of the proposed methodology

When Pareto steps were applied to the data presented above, Pareto analysis of accident causes is obtained (Figure 2).

Table 1. Causes of occupational accidents

ESAW Code	Causes of Accident	Number of Accidents
11	Electrical problem due to electrical failure - leading to direct contact	51
12	Electrical problem - leading to direct contact	59
13	Explosion	17
14	Fire, ignition	908
19	Other 10-type deviations not listed above	21
21	In solid state - leash, tipping	633
22	In the case of a liquid - seepage, leakage, spill, splash, spray	1069
23	In gaseous state - evaporation, aerosol formation, gas formation	991
24	Powdery substance - smoke generation, airborne/emitted dust and particles	44
29	Other 20-type deviations not listed above	104
31	Breakage of matter - at joints, junctions	144
32	Breakage, explosion - splinter formation (wood, glass, metal, stone, plastic, others)	657
33	Slip, fall, collapse of material vehicle - from above (falling on the victim)	962
34	Slip, fall, collapse of material vehicle - from below (pulling the casualty down)	53
35	Sliding, falling, collapsing material vehicle - at the same level	744
39	Other 30-type deviations not listed above	176
41	Loss of control (full or partial) - machine (including unintentional start) or material tool	2163
42	Loss of control (full or partial) - use of vehicle or equipment (whether powered or not)	96
43	Loss of control (full or partial) - the hand tool (whether powered or not) or the material tool used by the tool	2696
44	Loss of control (full or partial) - object (moved, played, used, etc.)	1802
45	Loss of control (full or partial) - animal	9
49	Other 40-type deviations not listed above	357
51	One's downfall - to the lower level	1361
52	Slip - stumble and fall - Person fall - on the same level	2223
59	Other 50-type deviations not listed above	191
61	Walking on a sharp object	65
62	Kneeling, sitting, leaning	107
63	Grasping or being carried away - by something or momentum	583
64	Uncooperative action, unnecessary or untimely actions	3550
69	Other 60-type deviations not listed above	153
71	Lifting, carrying, standing	570
72	Push, pull	497
73	To drop down, to bend over	84
74	Twist, turn	106
75	Chewing, twisting of leg or ankle, gliding without falling	180
79	Other 70-type deviations not listed above	454
81	Shock, fear	6
82	Brutality, assault, threat - among company employees who are subject to the employer's authority	36
83	Brutality, assault, threat - by people outside the company (bank robbery, bus drivers, etc.)	7
84	Attack, poke - by animal	11
85	A third person or a person who creates a danger to himself/herself and others	5
89	Other 80-type deviations not listed above	9
99	Other Deviations not listed in this classification	9367



* It is the bar diagram denoted as “other”. Accident causes with a cumulative contribution of 1% or less have been combined to make it readable.

Figure 2. Pareto Analysis of Accident Causes

Mathematical background of AHP method

The weights (w) of the main and sub-criteria are obtained as a result of pairwise comparisons. In this method, it is not obligatory for decision makers to make numerical comparisons. Linguistic comparisons can also be used. In the literature, Saaty's scale between 1-9 values is generally used in pairwise comparisons. In comparison matrices, the diagonals take the value 1 (Yildirim and Oner, 2014; Saaty, 1986). This is because the criterion is compared with itself (Saaty, 1990).

If the criteria used in the method are $a_1, a_2, \dots, a_n$, and the weights are $w_1, w_2, \dots, w_n$, if it is desired to compare n criteria according to their relative importance weights, the general structure of the pairwise comparison matrix will be as in Equation 1.

$$A = \begin{bmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{bmatrix} \quad (1)$$

where $a_{ij} = \frac{1}{a_{ji}}$ (according to the responsiveness) and $a_{ij} = \frac{a_{ik}}{a_{jk}}$

In real problems, the result of $\frac{w_i}{w_j}$ is often unknown. Therefore, what needs to be solved in AHP is to find a_{ij} value such that $a_{ij} \cong \frac{w_i}{w_j}$ (Yildirim and Oner, 2014).

The general form of the weight matrix is as in Equation 2.

$$W = \begin{bmatrix} w_1 \\ \vdots \\ w_i \\ \vdots \\ w_n \end{bmatrix} \begin{bmatrix} w_1 & \dots & w_j & \dots & w_n \\ w_1/w_1 & \dots & w_1/w_j & \dots & w_1/w_n \\ \vdots & & \vdots & & \vdots \\ w_i/w_1 & \dots & w_i/w_j & \dots & w_i/w_n \\ \vdots & & \vdots & & \vdots \\ w_n/w_1 & \dots & w_n/w_j & \dots & w_n/w_n \end{bmatrix} \quad (2)$$

Equation 3 is obtained by multiplying the W and w values with each other.

$$W \cdot w = \begin{bmatrix} w_1 \\ \vdots \\ w_i \\ \vdots \\ w_n \end{bmatrix} \begin{bmatrix} w_1 & \dots & w_j & \dots & w_n \\ w_1/w_1 & \dots & w_1/w_j & \dots & w_1/w_n \\ \vdots & & \vdots & & \vdots \\ w_i/w_1 & \dots & w_i/w_j & \dots & w_i/w_n \\ \vdots & & \vdots & & \vdots \\ w_n/w_1 & \dots & w_n/w_j & \dots & w_n/w_n \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_i \\ \vdots \\ w_n \end{bmatrix} = n \begin{bmatrix} w_1 \\ \vdots \\ w_i \\ \vdots \\ w_n \end{bmatrix} \quad (3)$$

The equation obtained in Equation 3 is also shown as in Equation 4.

$$(W - nI)w = 0 \quad (4)$$

The solution to this equation is the eigenvalue finding problem. Their relative weights are calculated with the w eigenvector based on λ_{Maks} , which satisfies the equation $Aw = \lambda_{Maks}w$. Here, λ_{Maks} is the largest eigenvalue of the A matrix and depending on the w eigenvector $(A - \lambda_{Maks}I)w = 0$ equation is obtained (Saaty, 1990).

In addition, two coefficients, Consistency Index (CI) and Consistency Ratio (CR), are used to determine the consistency of subjective perceptions and the accuracy of relative weights. Consistency Index (CI) can be calculated as in Equation 5.

$$CI = \frac{(\lambda_{Maks} - n)}{(n - 1)} \quad (5)$$

In the above equation, λ_{Maks} is the largest eigenvalue and n is the total number of criteria. For the results to be reliable, the CI value should not exceed 0.1. Calculation of the consistency ratio is done with the formula in Equation 6.

$$CR = \frac{CI}{RI} \quad (6)$$

In the above equation, RI stands for “Random Value Index”. It is derived from a large sample of a randomly generated cross-comparison matrix. The scale used here is 1/9, 1/8, 1/7, 1/6, ..., 1, ..., 6, 7, 8, 9. RI values obtained according to different criteria numbers are shown in Table 2.

While determining the main and sub-criteria, SSI data were used because there is no database for industrial accidents in Turkey and it provides the widest data on accidents. A comprehensive literature search was also carried out, and the seven main criteria for the problem and the sub-criteria of these main criteria were obtained. The hierarchical structure of alternative organizations are presented in Figure 3.

Table 2. *Random Value Index*

<i>n</i>	1	2	3	4	5	6	7	8	9	10	11	12
Random Value Index (RI)	0	0	0.52	0.89	1.11	1.25	1.35	1.40	1.45	1.49	1.54	1.54

Questionnaires were prepared for the main criteria and sub-criteria of root causes. Comparison matrices were created for each criterion to determine the weight levels of the main criteria and sub-criteria in the questionnaires. These prepared questionnaires were applied to experts consisting of academicians and field representatives, who have knowledge of national and international legislation on industrial accidents (1 Academician, 1 “Republic of Turkey Ministry of Labor and Social Security(CSGB), General Directorate of Occupational Health and Safety(ISGGM)” personnel, 1 “Republic of Turkey Ministry of Labor and Social Security, Guidance and Inspection Directorate” personnel, 1 “Republic of Turkey Ministry of Environment and Urbanization(CSB)” personnel, 1 “Republic of Turkey Ministry of Interior Disaster and Emergency Management Presidency(AFAD)” personnel and 1 industrial establishment personnel).

2.3. Evaluation of Organizations based on Root Causes of Industrial Accident

Three industrial organizations in different locations within the scope of the relevant legislation were also evaluated by AHP method, and as a requirement of this method, pairwise comparison matrices of the sub-criteria were created. All of the organizations are high-level and operate in the province of Ankara in Turkey.

3. RESULTS AND DISCUSSION

3.1. AHP Method

In order to perform AHP method, the main criteria and sub-criteria that may be the root cause of industrial accidents were determined by applying a questionnaire based on expert opinions. AHP method pairwise comparison matrices were prepared on the basis of both main criteria and sub-criteria. While performing the AHP method, the geometric averages for the main and sub-criteria were calculated, and pairwise comparison matrices were formed, in terms of combining the judgments of the questionnaires filled by six experts. AHP method was carried out in accordance with the equations and explanations given in Section 2.2. Weight values and consistency ratio of the main criteria are given in Table 3.

Table 3. *Weight values and consistency ratio of the main criteria*

Main Criterion	Sub Criteria	Weight Value/Priority Value
Consistency Rate $CR = 0.061 < 0.1$	Management / Organization Errors	0.090
	Process Errors	0.190
	Equipment Errors	0.200
	Domino Effect	0.080
	Human Errors	0.350
	Disasters	0.050
	Sabotage	0.040

Weight values and consistency ratio of sub-criteria are given in Table 4.

The human errors criterion came to the fore as the most important criterion in industrial accidents with 35%. This was followed by the equipment errors criterion with 20% and the process errors criterion with 19%, respectively. When determining the causes of major industrial accidents, it is important to consider possible hardware and software failures, process and design deficiencies, and human errors. It should also be stated what action is required to prevent these failures and errors (Prevention of major industrial accidents, 1991). Regulations play a critical role in reducing the occurrence of industrial accidents. The importance of addressing human errors as a proactive measure is also highlighted (Simsek et al, 2024). As a result of the pairwise comparisons of the management/organization criteria, the "improper security management system " sub-criterion was found to be the most important criterion with a rate of 26%. This result is compatible with the literature (Ozdemir et al., 2018). This was followed by the criteria of "incorrect/missing procedures and instructions" and "inadequate trainings", respectively. Botti et al. (2020) also emphasized the importance of voluntary adoption of an incorrect procedure, skipping a risk control measure, improper use of equipment and lack of coordination in accidents. There are positive and significant relationships among various factors such as procedure implementation, communication accuracy, communication satisfaction, work permit system, competence level and risk management. These factors have been shown to have a positive effect on the occurrence and frequency of maintenance related accidents (Gonyora et al., 2024). Given the critical role of employees in preventing major accidents, business management requires employees to have a broad understanding of the process used. Employees must be informed and adequately trained about the hazards of the materials used. This information and training must be in an appropriate language and format (Prevention of major industrial accidents, 1991). In process errors, "lack/insufficiency of process control" was determined as the most important criterion with a rate of 27%. "untested/experienced safety barriers" and "lack/insufficiency of passive safety barriers" were also determined at significant rates. The "vapor cloud explosion", which was under the domino effect main criterion, stood out as the most important criterion with a rate of 22%. This was followed by "BLEVE-Fireball" and "Thermal Runaway". The analyses of past domino effect accidents also supported these results (Reniers and Cozzani, 2013). In equipment errors, the criterion of "faulty design/installation/assembly" was significantly ahead (29%). "improper equipment placement/safety distances " and " inappropriate/improper maintenance " criteria are also very important. Singh and Maiti (2020) also drew attention to poorly maintained equipment in their study (Sigh and Maiti, 2020). It is recommended that organizations in the chemical and process industry continue to prioritize maintenance practices as well as human and organizational factors to ensure safe and reliable operational performance. Focusing on these issues can minimize the occurrence of accidents and promote a safe working environment (Gonyora et al., 2024). "Lack of safety culture" was found to be the most important in human errors. The importance of safety culture in preventing accidents is known (Zhang et al., 2020). Subsequently, "insufficient knowledge/ignorance about the job" and "inexperience" were determined as important. As a result of the pairwise comparisons of the disasters criteria, the "earthquake" criterion came to the fore as the most important criterion with a rate of 33%.

This was followed by the "lightening" criteria with 18% and the "storm/tornado" criteria with 12%, respectively. The importance weight of both geological and meteorological effects is remarkable. In the sabotage criterion, "sabotage with explosives" was found to be ahead with a significant rate (44%). The criterion of "tampering with fire" was determined as the second important one. Consistencies in all were determined at the appropriate value (<0.1).

The hazard analysis should include the analysis of the safety system for possible weaknesses, the determination of the residual risk with the safety system in place, and the development of the most appropriate measures for technical and organizational protection in the event of abnormal plant operation (Prevention of major industrial accidents, 1991).

Measures should also be taken to minimize the consequences of major accidents. The enterprise should plan and provide appropriate measures to mitigate the consequences of possible accidents. Mitigation should be done with security systems, alarm systems, emergency services, etc. For each major hazard facility, an on-site emergency plan should be prepared in consultation with the security team. Depending on local regulations, an external emergency plan should also be developed and implemented in cooperation with the relevant local authorities (Prevention of major industrial accidents, 1991).

Through the rigorous application of multi-criteria decision analysis (MCDA), a robust hierarchy was established by identifying specific nations as applicable examples for those aiming to increase industrial accident prevention strategies (Simsek et al., 2024). The effectiveness of the MCDM technique is also seen in the root cause analysis of other accidents. Weighting of the selection criteria for the causes of work accidents Using Fuzzy AHP, the criteria were ranked and evaluated by making pairwise comparisons of the experts' scores. The results of this study are noteworthy due to the scarcity of quantitative studies on shipboard accidents, which are considered a major problem for the maritime industry. Individuals' perception of safety culture is an important factor in accidents caused by human errors. Improving the perception of safety culture will contribute to the prevention of accidents caused by human errors (Ozdemir et al., 2018). An accident analysis model is proposed to identify the leading causes of accidents and then determine the most cost-effective safety measures to prevent accidents from occurring. In the first part of the model, HFACS and BN are utilized to investigate the causes and propose corresponding prevention measures. In the second part, the Best Fit method and ER approach are used to rank the proposed security measures in terms of cost-effectiveness. Case study of ship collision accidents demonstrated the effectiveness of the proposed model (Wang et al., 2013).

3.2. Evaluation of Organizations based on Root Causes of Industrial Accident

In order to determine the organization that is more risky in terms of industrial accident, an application was made in 3 industrial organizations in Turkey, which are within the scope of the legislation. Again, pairwise

comparison matrices were prepared on the basis of sub-criteria whose weights were determined by AHP method. The locations of the industrial organizations are shown in Figure 4.

All of the organizations are high-level organizations operating in the province of Ankara. Organization A was established in the 1990s and is an explosives production company. Within the organization; ammonium nitrate, packaged emulsion explosives, pibs, sodium thiocyanate, monoe thanol amine, sodium nitrate, sodium perchlorate, sodium nitrite, acetic acid, thiourea, EXAN etc. chemicals are used. Organization B was established in the 1960s and is a LPG filling company. Within the organization; special forklift tube, marine gas, LPGPRO, CUTPRO, IZOPRO, BETOPRO etc. products are filled. Organization C was established in the 2000s and is a LPG storage and filling company. Within the organization, there are many equipment such as tank farm, LPG pumps, fire pumps, gas compressors, etc.

The comparison matrices of the alternatives on the basis of sub-criteria and the priority vectors obtained as a result of the calculations are presented in Table 5. All ratios calculated in the comparison of alternatives on the basis of sub-criteria were less than 0.1 and were consistent. While evaluating the alternatives together with the sub-criteria, since the consistency ratio was calculated as 0 in most of them, it was not shown as a separate column in the relevant Table 5. The determination of the relative weight of each organization for the purpose was calculated as given in Equation 7 by multiplying the importance of the organizations on the basis of the sub-criterion, the relative weight of the sub-criterion and the relative weight of the main criterion.

$$S_l = \sum_{i=1, j=1}^{nm} w_i v_{ij} p_{ijl} \quad (7)$$

where

S_l : The relative weight of the 1st organization showing its ranking as a result of the AHP method

w_i : The weight of the ith main criterion

v_{ij} : The weight of the jth sub-criterion of the ith main criterion

p_{ijl} : The weight of the jth sub-criterion of the ith main criterion for the lth organization

$i = 1, 2 \dots n, j = 1, 2 \dots m$

When the calculation was made in accordance with the above equation, the scores of each organization's relative weight for the purpose determined by the AHP method were as follows.

$S_1=0.354$ (Organization A)

$S_2=0.420$ (Organization B)

$S_3=0.226$ (Organization C)

Organization B received the highest industrial accident risk score with a value of 0.420. This was followed by Organization A with a value of 0.354, and then by Organization C with a value of 0.226. Although all organizations were high-level, the potential for industrial accident risk differed. Although organizations B and C contain similar hazardous chemical and process equipment, they also differ in terms of industrial accident risk. In accordance with the relevant legislation, the organization level is determined only according to the type and amount of certain hazardous chemicals (Regulation on Industrial Accidents, 2019). A comprehensive evaluation was provided by including expert opinions on the many criteria presented in this study, and the accident risk potential of the organizations was objectively revealed on the basis of root cause.

Solutions obtained with mathematical models represent optimum results that are precise and robust. In problems like this where managers are only interested in optimum solutions, it becomes important to find the range of variation of the coefficients that does not change the result. A similar approach, namely sensitivity analysis, is applied in the study with analytical methods within the MCDM methods such as AHP. The factor affecting the sensitivity here is the people involved in the process as experts. The individual differences of more than one expert are followed in the process and in case of a significant difference/inconsistency, the opinions of the relevant expert are taken again. If necessary, it is also possible to remove the relevant expert completely. Therefore, the effect of an expert on the overall weights will be limited. In case of a different nonconformity, the solution can of course be tried again, but when we go through the results we obtained, there is no risk of changing the final ranking and sensitivity analysis was not deemed necessary for this study (Zopounidis, and Doumpos, 2017; Gal et al., 1999).

The competent authorities should define appropriate safety objectives together with the major hazard control system for their implementation. Although the control of major hazards is primarily the responsibility of the company's management, it should be established in consultation with all relevant parties. The relevant system should include the following:

- (a) establishing an infrastructure;
- (b) identifying and inventorying major hazard facilities;
- (c) obtaining and evaluating safety reports;
- (d) emergency planning and public information;
- (e) site selection and land-use planning;
- (f) inspecting facilities;
- (g) reporting major accidents;
- (h) investigating major accidents and their short- and long-term effects (Prevention of major industrial accidents, 1991).

4. CONCLUSION

In this study, a new approach based on evaluation of organizations according to the root causes of industrial accidents with AHP method was developed. The main and sub-criteria based on the root cause were determined through the Pareto Analysis of the past accident data. The weights of the main criteria and sub-criteria were determined through AHP method and expert opinions. Considering the criteria weights determined, three sample industrial organizations were ranked in terms of industrial accident risk with a second AHP method usage. As a result of the study, the original seven main criteria and sixty-nine sub-criteria that could be the root cause were determined. As a result of the pairwise comparisons of the main criteria; "human errors" stood out as the most important criterion in industrial accidents with 35%. This was followed by equipment errors with 20% and process errors with 19%, respectively. As a result of the pairwise comparisons of the sub-criteria, the rates listed below and the most important criteria were determined.

"Improper security management system" criterion of the Management / Organization criterion was 26%

"Lack/insufficiency of process control" criterion belonging to the process criterion was 27%,

"Faulty design/installation/assembly" criteria of equipment criteria was 29%,

The "Vapor cloud explosion" criterion of the domino effect criterion was 22%,

"Lack of safety culture" criterion of human error criterion was 17%,

"Earthquake" criterion of disasters criterion was 33%,

"Sabotage with Explosives" criteria belonging to the sabotage criteria was 44%

It is thought that the degree of importance obtained from a significant number of main and sub-criteria analyzed as the root causes of industrial accidents determine the elements that should be considered primarily in risk analyzes to be made in organizations where industrial accidents may occur. The high importance of the "human errors" criterion has shown that the personnel of the relevant organization should have sufficient and experience in safety culture, work and processes. The high importance of the "equipment errors" criterion has revealed the necessity of taking into consideration the faulty design/installation/assembly, improper equipment placement/safety distances, inappropriate/improper maintenances as a priority in the relevant organizations. The high importance of the "process errors" criterion has shown that the relevant organizations should especially analyze the lack / inadequacy of process control, untested / unexperienced safety barriers and passive safety barriers deficiency / inadequacy.

As a result of the analyzes made in three sample organizations, the organization B received the highest score with a value of 0.420. Organization A took the second place with a value of 0.354, and organization C took the last place with a value of 0.226. The risk potential of high-level organizations with high risks has been

rated objectively according to many criteria. It has been shown that the relevant industry and public practitioners should pay attention to which root cause-based criteria in which organizations.

Industrial accidents are evaluated within the scope of the Regulation on the Prevention of Major Industrial Accidents and Reducing Their Effects (RPMIARTE). In this regulation, organizations are categorized as upper and lower level depending on the type and amount of hazardous chemicals involved. Higher level organizations are organizations that involve higher accident risks. Therefore, the application was made on selected higher level organizations that involve high risks. The methodology proposed in the study can be applied to every organization within the scope of RPMIARTE. Our country's legislative equivalent of the EU SEVESO directives is RPMIARTE. Therefore, the proposed methodology can be applied to all Seveso organizations and is not regional in nature. However, including lower-level organizations with similar risk levels among themselves and upper-level organizations among themselves in the process will provide more reasonable results.

There are many methods used in the literature under the title of decision-making problem, such as MCDM methods, fuzzy logic and Bayesian approach. The problem addressed in the study poses a decision problem because there are alternatives. In addition, the problem is closer to solution with MCDM methods due to the criteria affecting the decision. Saaty scale was used in AHP. This scale works with all integers between 1 and 9. If survey evaluators are hesitant to give values 2-4-6-8 and have a sensitivity problem, they can manage the process with values 1-3-5-7-9. The evaluation of this situation is controlled by the consistency ratio (CR) values and the CR value is expected to be below 0.1. No negative effects were encountered in the study. It is seen that using membership degrees ranging from 0 to 1 in the fuzzy logic method produces a solution in cases where experts cannot solve the sensitivity/instability problem (Zopounidis and Doumpos, 2017; Gal et al., 1999).

It has been shown that AHP method, which is one of the MCDM methods used in the evaluation of organizations accident risk according to the root causes of industrial accidents, is an appropriate analysis for such problems aiming to compare multiple criteria. With the AHP method, quantitative and qualitative evaluation criteria were included in the model simultaneously. An objective result could be obtained by simultaneously evaluating the thoughts and judgments of decision makers with different education and experience. In future studies, analyzes can be carried out by using fuzzy MCDM methods at the stage of determining the importance levels of the criteria that may be the root cause, and by using another MCDM method in the stage of comparing organizations in terms of industrial accident risk. The results to be obtained can be compared with the results of this study.

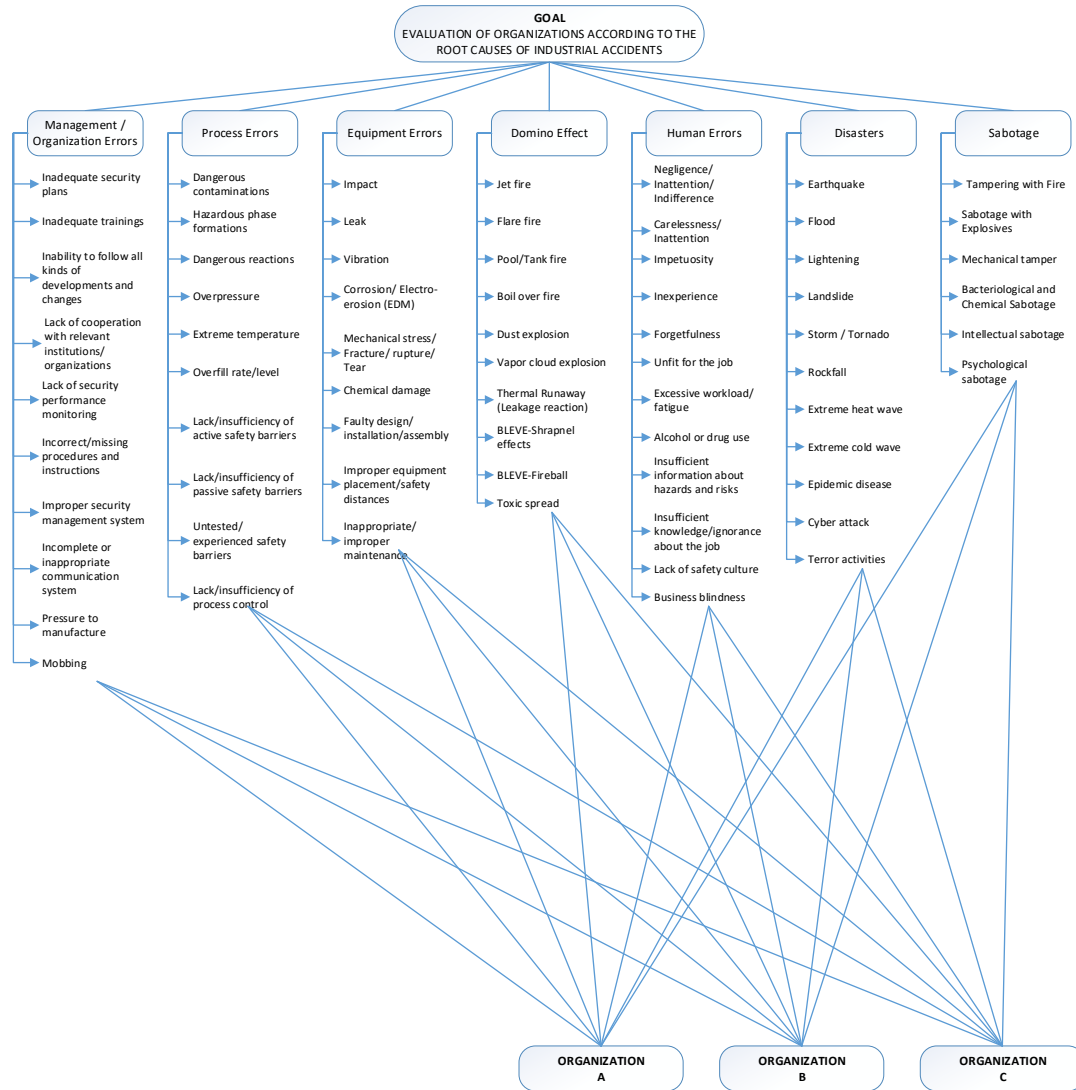


Figure 3. The hierarchical structure of alternative organizations

Table 4. Weight values and consistency ratio of sub-criteria

Main Criterion	Sub Criteria	Weight Value/ Priority Value
Management / Organization Errors CR = 0.037 < 0.1	Inadequate security plans	0.130
	Inadequate trainings	0.140
	Inability to follow all kinds of developments and changes	0.040
	Lack of cooperation with relevant institutions/organizations	0.050
	Lack of security performance monitoring	0.080
	Incorrect/missing procedures and instructions	0.150
	Improper security management system	0.260
	Incomplete or inappropriate communication system	0.080
	Pressure to manufacture	0.040
	Mobbing	0.020

Table 4. continued

Main Criterion	Sub Criteria	Weight Value/ Priority Value
Process Errors CR = 0.033 < 0.1	Dangerous contaminations	0.030
	Hazardous phase formations	0.030
	Dangerous reactions	0.040
	Overpressure	0.070
	Extreme temperature	0.070
	Overfill rate/level	0.050
	Lack/insufficiency of active safety barriers	0.120
	Lack/insufficiency of passive safety barriers	0.130
	Untested/experienced safety barriers	0.190
	Lack/insufficiency of process control	0.270
Equipment Errors CR = 0.066 < 0.1	Impact	0.080
	Leak	0.040
	Vibration	0.020
	Corrosion/ Electro-erosion (EDM)	0.070
	Mechanical stress/ Fracture/ rupture/ Tear	0.110
	Chemical damage	0.100
	Faulty design/installation/assembly	0.290
	Improper equipment placement/safety distances	0.180
	Inappropriate/improper maintenance	0.120
Domino Effect CR = 0.049 < 0.1	Jet fire	0.120
	Flare fire	0.070
	Pool/Tank fire	0.030
	Boil over fire	0.030
	Dust explosion	0.050
	Vapor cloud explosion	0.220
	Thermal Runaway	0.160
	BLEVE-Shrapnel effects	0.140
	BLEVE-Fireball	0.160
	Toxic spread	0.010

Table 4. continued

Main Criterion	Sub Criteria	Weight Value/ Priority Value
Human Errors CR = 0.029 < 0.1	Negligence/ Inattention/ Indifference	0.060
	Carelessness/ Inattention	0.050
	Impetuosity	0.030
	Inexperience	0.140
	Forgetfulness	0.040
	Unfit for the job	0.100
	Excessive workload/fatigue	0.070
	Alcohol or drug use	0.040
	Insufficient information about hazards and risks	0.120
	Insufficient knowledge/ignorance about the job	0.140
	Lack of safety culture	0.170
	Business blindness	0.040
Disasters CR = 0.056 < 0.1	Earthquake	0.330
	Flood	0.110
	Lightening	0.180
	Landslide	0.100
	Storm / Tornado	0.120
	Rockfall	0.030
	Extreme heat wave	0.030
	Extreme cold wave	0.030
	Epidemic disease	0.020
	Cyber attack	0.030
	Terror activities	0.030
Sabotage CR = 0.074 < 0.1	Tampering with Fire	0.310
	Sabotage with Explosives	0.440
	Mechanical tamper	0.120
	Bacteriological and Chemical Sabotage	0.070
	Intellectual sabotage	0.030
	Psychological sabotage	0.030



Figure 4. Google Earth image of sample industrial organizations

Table 5. Priority vectors for alternatives on basis of sub-criteria and priority vectors obtained for organizations

Main Criterion	Sub Criteria	Weight Value/ Priority Value	A	B	C
Management / Organization Errors CR = 0.037 < 0.1	Inadequate security plans	0.130	0.278	0.500	0.222
	Inadequate trainings	0.140	0.278	0.389	0.333
	Inability to follow all kinds of developments and changes	0.040	0.412	0.353	0.235
	Lack of cooperation with relevant institutions/organizations	0.050	0.357	0.357	0.286
	Lack of security performance monitoring	0.080	0.350	0.400	0.250
	Incorrect/missing procedures and instructions	0.150	0.368	0.368	0.263
	Improper security management system	0.260	0.318	0.409	0.273
	Incomplete or inappropriate communication system	0.080	0.333	0.467	0.200
	Pressure to manufacture	0.040	0.389	0.444	0.167
	Mobbing	0.020	0.438	0.375	0.188
Process Errors CR = 0.033 < 0.1	Dangerous contaminations	0.030	0.389	0.333	0.278
	Hazardous phase formations	0.030	0.389	0.389	0.222
	Dangerous reactions	0.040	0.412	0.412	0.176
	Overpressure	0.070	0.333	0.429	0.238
	Extreme temperature	0.070	0.389	0.444	0.167
	Overfill rate/level	0.050	0.389	0.444	0.167
	Lack/insufficiency of active safety barriers	0.120	0.278	0.500	0.222
	Lack/insufficiency of passive safety barriers	0.130	0.188	0.438	0.375
	Untested/experienced safety barriers	0.190	0.313	0.375	0.313
	Lack/insufficiency of process control	0.270	0.333	0.400	0.267

Table 5. continued

Main Criterion	Sub Criteria	Weight Value/ Priority Value	A	B	C
Equipment Errors CR = 0.066 < 0.1	Impact	0.080	0.300	0.600	0.100
	Leak	0.040	0.231	0.615	0.154
	Vibration	0.020	0.273	0.545	0.182
	Corrosion/ Electro-erosion (EDM)	0.070	0.200	0.600	0.200
	Mechanical stress/ Fracture/ rupture/ Tear	0.110	0.313	0.500	0.188
	Chemical damage	0.100	0.313	0.438	0.250
	Faulty design/installation/assembly	0.290	0.350	0.400	0.250
	Improper equipment placement/safety distances	0.180	0.350	0.400	0.250
	Inappropriate/improper maintenance	0.120	0.333	0.429	0.238
Domino Effect CR = 0.049 < 0.1	Jet fire	0.120	0.350	0.400	0.250
	Flare fire	0.070	0.350	0.400	0.250
	Pool/Tank fire	0.030	0.286	0.500	0.214
	Boil over fire	0.030	0.333	0.500	0.167
	Dust explosion	0.050	0.333	0.467	0.200
	Vapor cloud explosion	0.220	0.333	0.467	0.200
	Thermal Runaway (Leakage reaction)	0.160	0.333	0.467	0.200
	BLEVE-Shrapnel effects	0.140	0.263	0.421	0.316
	BLEVE-Fireball	0.160	0.313	0.500	0.188
Human Errors CR = 0.029 < 0.1	Toxic spread	0.010	0.333	0.500	0.167
	Negligence/ Inattention/ Indifference	0.060	0.450	0.400	0.150
	Carelessness/ Inattention	0.050	0.474	0.368	0.158
	Impetuosity	0.030	0.450	0.350	0.200
	Inexperience	0.140	0.474	0.263	0.263
	Forgetfulness	0.040	0.438	0.375	0.188
	Unfit for the job	0.100	0.500	0.333	0.167
	Excessive workload/fatigue	0.070	0.389	0.444	0.167
	Alcohol or drug use	0.040	0.450	0.350	0.200
Disasters CR = 0.056 < 0.1	Insufficient information about hazards and risks	0.120	0.450	0.350	0.200
	Insufficient knowledge/ignorance about the job	0.140	0.429	0.333	0.238
	Lack of safety culture	0.170	0.368	0.474	0.158
	Business blindness	0.040	0.409	0.409	0.182
	Earthquake	0.330	0.100	0.700	0.200
	Flood	0.110	0.111	0.556	0.333
	Lightening	0.180	0.412	0.353	0.235
	Landslide	0.100	0.125	0.625	0.250
	Storm / Tornado	0.120	0.125	0.625	0.250
	Rockfall	0.030	0.167	0.667	0.167
	Extreme heat wave	0.030	0.167	0.667	0.167
	Extreme cold wave	0.030	0.167	0.667	0.167
	Epidemic disease	0.020	0.385	0.462	0.154
	Cyber attack	0.030	0.385	0.385	0.231
	Terror activities	0.030	0.111	0.667	0.222

Table 5. continued

Main Criterion	Sub Criteria	Weight Value/ Priority Value	A	B	C
Sabotage CR = 0.074 < 0.1	Tampering with Fire	0.310	0.429	0.381	0.190
	Sabotage with Explosives	0.440	0.409	0.364	0.227
	Mechanical tamper	0.120	0.400	0.400	0.200
	Bacteriological and Chemical Sabotage	0.070	0.273	0.545	0.182
	Intellectual sabotage	0.030	0.538	0.308	0.154
	Psychological sabotage	0.030	0.500	0.357	0.143

AUTHOR CONTRIBUTIONS

Investigation and visualization, S.U.; conceptualization, supervision, and writing - original draft, S.C.; writing - review & editing, T.C.; methodology, E.C. All authors have read and legally accepted the final version of the article published in the journal.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Alekseeva, A., Volokhina, A., & Glebova, E. (2020). Identification of the root causes of accidents at hazardous production facilities of the fuel and energy complex. IOP Conference Series: Materials Science and Engineering, 2-5. <https://doi.org/10.1088/1757-899X/976/1/012021>
- Andersan, B., & Fagerhaug, T. (2006). Root Cause Analysis, Simplified Tools and Techniques (Second Edition). Milwaukee, Wisconsin, ASQ Quality Press.
- Aycin, E. (2019). Multi-Criteria Decision Making Computer Applied Solutions, Ankara, Nobel Publications, 333.
- Botti, L., Melloni, R., Mosconi S., & Oliva M. (2020). A Detailed Investigation on Apparent and Root Causes of Accidents in Manufacturing, Advances in Manufacturing, Production Management and Process Control, 18-25.
- Bragatto, P., Vairo, T., Maria, F. M., & Fabiano, B. (2021). The impact of the COVID-19 pandemic on the safety management in Italian Seveso industries. Journal of Loss Prevention in the Process Industries, 70. <https://doi.org/10.1016/j.jlp.2021.104393>
- Ciftci, S. E., & Arikan, F. (2020). A Multiple Criteria Risk Analysis Model and a Case Study in Metal Industry, Open Journal of Business and Management, 8, 2048-2070. <https://doi.org/10.4236/ojbm.2020.85125>
- Dabous, S. A., Ibrahim, F., Feroz, S., & Alsyounf, I. (2021). Integration of failure mode, effects, and criticality analysis with multi-criteria decision-making in engineering applications: Part I – Manufacturing industry. Engineering Failure Analysis, 122. <https://doi.org/10.1016/j.engfailanal.2021.105264>

- Dincer, H., & Gorener, A. (2011). Dynamic performance analysis with analytical hierarchy process and VIKOR technique: An application in the banking sector. *Istanbul Commerce University Journal of Social Sciences*, 19, 10, 109-127.
- Fata, C. M., Giallanza, A., Micale, R., & La Scalia G. (2021). Ranking of occupational health and safety risks by a multi-criteria perspective: Inclusion of human factors and application of VIKOR. *Safety Science*, 138. <https://doi.org/10.1016/j.ssci.2021.105234>.
- Gal, T., Stewart, T.J., & Hanne, T. (1999). *Multicriteria Decision Making Advances in MCDM Models, Algorithms, Theory, and Applications*. Frederick S. Hillier, International Series in Operations Research & Management Science (ISOR, volume 21), Springer Science+ Business Media, LLC.
- Ghatorha, K., Sharma, R., & Singh, G. (2020). Application of root cause analysis to increase material removal rate for productivity improvement: A case study of the press manufacturing industry. *Materials Today: Proceedings*, 26(2), 1780-1783. <https://doi.org/10.1016/j.matpr.2020.02.374>
- Gonyora, M., & Ventura-Medina, E. (2024). Investigating the relationship between human and organisational factors, maintenance, and accidents. The case of chemical process industry in South Africa. *Safety Science*, 176, 106530. <https://doi.org/10.1016/j.ssci.2024.106530>
- Gul, M. (2020). Application of Pythagorean fuzzy AHP and VIKOR methods in occupational health and safety risk assessment: the case of a gun and rifle barrel external surface oxidation and coloring unit. *International Journal of Occupational Safety and Ergonomics*, 26(4), 705-718. <https://doi.org/10.1080/10803548.2018.1492251>
- Gul, M. (2018). A review of occupational health and safety risk assessment approaches based on multi-criteria decision-making methods and their fuzzy versions. *Human and Ecological Risk Assessment: An International Journal*, 24, 1-38. <https://doi.org/10.1080/10807039.2018.1424531>
- Heller, S. (2006). Managing industrial risk—Having a tested and proven system to prevent and assess risk. *Journal of Hazardous Materials*, 130, 1–2, 18-63. <https://doi.org/10.1016/j.jhazmat.2005.07.067>
- Ilbahar, E., Karasan, A., Cebi, S., & Kahraman, C. (2018). A novel approach to risk assessment for occupational health and safety using Pythagorean fuzzy AHP & fuzzy inference system. *Safety Science*, 103, 124-136. <https://doi.org/10.1016/j.ssci.2017.10.025>
- Koulinas, G.K., Marhavilas, P.K., Demesouka, O.E., Vavatsikos, A.P., & Koulouriotis, D.E. (2019). Risk analysis and assessment in the worksites using the fuzzy-analytical hierarchy process and a quantitative technique – A case study for the Greek construction sector. *Safety Science*, 112, 96-104. <https://doi.org/10.1016/j.ssci.2018.10.017>
- Li, G., Joe Qin, S., & Yuan, T. (2016). Data-driven root cause diagnosis of faults in process industries. *Chemometrics and Intelligent Laboratory Systems*, 159, 1-11. <https://doi.org/10.1016/j.chemolab.2016.09.006>
- Multiple Criteria Decision Making Applications in Management and Engineering, Constantin Zopounidis, Michael Doumpos, Springer International Publishing, Switzerland, 2017.

- Ozdemir, U., Altinpinar, İ., & Demirel, F.B. (2018). A MCDM Approach with Fuzzy AHP Method for Occupational Accidents on Board, *TransNav, The International Journal on Marine Navigation and Safety of Sea Transportation*, 12(1), 93-98. <https://doi.org/10.12716/1001.12.01.10>
- Oz, N. E., Mete, S., Serin, F., & Gul, M. (2019). Risk assessment for clearing and grading process of a natural gas pipeline project: An extended TOPSIS model with Pythagorean fuzzy sets for prioritizing hazards. *Human and Ecological Risk Assessment: An International Journal*, 25(6), 1615-1632. <https://doi.org/10.1080/10807039.2018.1495057>
- Ozguvenc, D. (2011). Multi-criteria Pareto Analysis in the classification of quality problems, Master Thesis, Istanbul Technical University, Institute of Science and Technology, Istanbul.
- Ouédraogo, A., Groso, A., & Meyer, T. (2011). Risk analysis in research environment—part II: weighting lab criticality index using the analytic hierarchy process. *Safety Science*, 49, 6, 785-793. <https://doi.org/10.1016/j.ssci.2010.12.011>
- Prevention of major industrial accidents, An ILO contribution to the International Programme on Chemical Safety of UNEP, the ILO and the WHO (IPCS), International Labour Office, Geneva, 1991.
- Regulation on Prevention of Major Industrial Accidents and Mitigation of Their Effects, Official Gazette dated 02.03.2019 and numbered 30702 (2019).
- Reniers, G., & Cozzani, V. (2013). *Domino Effects in the Process Industries Modelling*. Prevention and Managing, Oxford: Elsevier.
- Saaty, T.L. (1986). Axiomatic Foundations of The Analytic Hierarchy Process. *Management Science*, 32,7, 841-855. <https://doi.org/10.1287/mnsc.32.7.841>
- Saaty, T.L. (1990). How to Make a Decision: The Analytic Hierarchy Process. *European Journal of Operational Research*, 48(1), 9-26. [https://doi.org/10.1016/0377-2217\(90\)90057-I](https://doi.org/10.1016/0377-2217(90)90057-I)
- Seker, S., & Zavadskas, E. (2017). Application of Fuzzy DEMATEL Method for Analyzing Occupational Risks on Construction Sites. *Sustainability*, 9, 2083. <https://doi.org/10.3390/su9112083>
- Simsek, A.B., Kose, G., & Goktekin, Z. (2024). Evaluating country performance in preventing industrial accidents: A multi-criteria decision analysis approach. *Journal of Loss Prevention in the Process Industries*, 87, 105241. <https://doi.org/10.1016/j.jlp.2023.105241>
- Singh, K., & Maiti J. (2020). A novel data mining approach for analysis of accident paths and performance assessment of risk control systems. *Reliability Engineering & System Safety*, 202, 107041. <https://doi.org/10.1016/j.ress.2020.107041>
- Turskis, Z., Dzitac, S., Stankiuvienė, A., & Šukys, R. (2019). A Fuzzy Group Decision-making Model for Determining the Most Influential Persons in the Sustainable Prevention of Accidents in the Construction SMEs. *International Journal of Computers Communications & Control*, 14, 90-106.
- Viegas, R. A., Mota, F. S., Costa, A. P., & Santos, F. F. (2020). A multi-criteria-based hazard and operability analysis for process safety. *Process Safety and Environmental Protection*, 144, 310-321. <https://doi.org/10.1016/j.psep.2020.07.034>

- Wang, R., Gao, X., Gao, J., Gao, Z., & Kang, J. (2018). An information transfer based novel framework for fault root cause tracing of complex electromechanical systems in the processing industry. *Mechanical Systems and Signal Processing*, 101, 121-139. <https://doi.org/10.1016/j.ymssp.2017.08.030>
- Wang, Y. F., Xie, M., Chin, K., & Fu, X. J. (2013). Accident analysis model based on Bayesian Network and Evidential Reasoning approach. *Journal of Loss Prevention in the Process Industries*, 26(1), 10-21. <https://doi.org/10.1016/j.jlp.2012.08.001>
- Yildirim, B.F., & Onder, E. (2014). *Multi-Criteria Decision Making Methods in the Solution of Operational, Managerial and Strategic Problems for Businesses, Engineers and Managers*, Bursa: Dora Publications.
- Yilmaz, F., & Alp, S. (2016). Underlying Factors of Occupational Accidents: The Case of Turkey. *Open Journal of Safety Science and Technology*, 6(1), 4-8. <https://doi.org/10.4236/ojsst.2016.61001>
- Yucesan M., & Gul M. (2021). Failure modes and effects analysis based on neutrosophic analytic hierarchy process: method and application. *Soft Computing*, 25, 11035–11052. <https://doi.org/10.1007/s00500-021-05840-z>
- Zhang, J., Fu, J., Hao, H., Fu, G., Nie, F., & Zhang, W. (2020). Root causes of coal mine accidents: Characteristics of safety culture deficiencies based on accident statistics. *Process Safety and Environmental Protection*, 136, 78-91. <https://doi.org/10.1016/j.psep.2020.01.024>



Gazi University

Journal of Science

PART A: ENGINEERING AND INNOVATION

<http://dergipark.org.tr/guj.1648772>

Deep Learning-Based Lung Cancer Diagnosis: Data Balancing, Model Optimisation and Performance Analysis

Feyyaz ALPSALAZ^{1*} ¹ Yozgat Bozok University, Yozgat, Türkiye

Keywords	Abstract
Lung Cancer	Lung cancer (LC) is one of the most lethal malignancies worldwide, and early detection is essential. This study develops a deep learning (DL) based classification model for LC diagnosis using computed tomography (CT) images. In the experiments conducted on the IQ-OTHNCCD LC dataset, the Synthetic Minority Over-sampling Technique (SMOTE) method was applied to eliminate class imbalance, data augmentation techniques were used, and an early stopping mechanism was integrated to enhance the model's generalizability. Commonly used convolutional neural network (CNN) architectures, such as ResNet101, VGG19, and DenseNet121, are compared, and the model's performance is analyzed in detail. With an accuracy of 98%, the trial results demonstrate that the suggested ResNet101 model offers the best classification performance. the DenseNet121 model exhibited a relatively lower accuracy rate in distinguishing between benign and normal classes. The study conclusively demonstrates that an optimized ResNet101-based deep learning model, enhanced with data balancing and augmentation techniques, provides the most accurate and reliable classification performance for lung cancer detection using CT images. It not only outperforms traditional CNN architectures in terms of overall accuracy (98%) but also achieves perfect classification in malignant cases. These results validate the model's potential as a robust diagnostic aid and highlight its superiority over existing methods in the literature, particularly in handling class imbalance and maintaining generalization across diverse image types.
Deep Learning	
Computed Tomography	
CNN	
ResNet101	

Cite
Alpsalaz, F. (2025). Deep Learning-Based Lung Cancer Diagnosis: Data Balancing, Model Optimisation and Performance Analysis. <i>GU J Sci, Part A, 12(2)</i> , 373-391. doi:10.54287/guj.1648772

Author ID (ORCID Number)	Article Process
0000-0002-7695-6426 Feyyaz ALPSALAZ	Submission Date
	28.02.2025
	Revision Date
	21.03.2025
	Accepted Date
	25.03.2025
	Published Date
	30.06.2025

1. INTRODUCTION

Lung cancer (LC) is one of the world's most aggressive and deadly malignancies, responsible for approximately 1.8 million deaths annually according to recent WHO data (Thandra et al., 2021). Its high mortality rate is largely due to the lack of symptoms in the early stages, which often delays diagnosis and treatment. Early detection therefore plays a critical role in improving patient survival.

Computed tomography (CT) has become the most widely used modality in the diagnosis and follow-up of LC due to its ability to provide high-resolution cross-sectional images of lung tissue. CT scans allow clinicians to detect lung nodules, differentiate between benign and malignant lesions, and assess disease progression (Tárnoki et al., 2024). However, conventional diagnostic processes rely heavily on manual interpretation by radiologists, which can be time-consuming and prone to variability due to human subjectivity and fatigue.

*Corresponding Author, e-mail: feyyaz.alpsalaz@bozok.edu.tr

In recent years, the field of medical image analysis has witnessed transformative advances with the introduction of DL techniques. DL-based approaches have enabled the development of automated diagnostic systems that significantly improve accuracy while reducing time and error rates in clinical workflows (Özdemir et al., 2025). These models, particularly convolutional neural networks (CNNs), are capable of learning complex spatial features from CT images - from low-level textures to high-level structural patterns - that aid accurate diagnosis (Wang, 2022). When trained on large datasets, DL models outperform traditional methods by providing consistent, fast and objective assessments that are often superior to human interpretation (Javed et al., 2024).

Despite these advances, many current DL-based studies focus narrowly on classification tasks, overlooking challenges such as data imbalance and overfitting. The novelty of this study lies in its integrated approach, which includes preprocessing, class balancing using the Synthetic Minority Over-sampling Technique (SMOTE), and model optimisation strategies to improve performance across classes. The study uses a dataset of 1,190 CT scan slices from the IQ-OTH/NCCD dataset, categorised into benign, malignant and normal classes, and applies three powerful CNN architectures: ResNet101, VGG19 and DenseNet121.

The rest of this paper is structured as follows: Section 2 presents a review of related studies in the literature. Section 3 describes the dataset characteristics, the preprocessing pipeline, the model configurations, and the evaluation metrics. Section 4 discusses the experimental results, including model comparisons and performance analysis. Finally, Section 5 summarises the main findings and suggests future research directions.

2. LITERATURE REVIEW

In recent years, there has been an increasing amount of research focusing on the application of DL methods for LC diagnosis, particularly using CT images. CNNs, with their high accuracy and automatic feature extraction capabilities, have become the dominant architecture in these studies.

A number of studies have proposed hybrid or explainable DL models. Wani et al. (2024) developed DeepXplainer, a DL model that incorporates explainable artificial intelligence (XAI) techniques for LC diagnosis. This model optimises cancer diagnosis with the combination of CNN and XGBoost, and makes the model outputs explainable with the SHAP method. In the study, 97.43% accuracy, 98.71% sensitivity and 98.08% F1 score were obtained (Wani et al., 2024). Similarly, Mohamed et al. (2023) propose a hybrid CNN model with Ebola Optimisation Search Algorithm (EOSA) for LC diagnosis. The model was tested on the IQ-OTH/NCCD dataset and the EOSA CNN achieved 93.21% accuracy, 90.38% sensitivity and 97.95% specificity. The hyper-parameter optimisation of the EOSA algorithm improved the classification performance of the CNN model (Mohamed et al., 2023).

Several studies have combined imaging modalities. Rajasekar et al. (2023) developed a DL based system for LC diagnosis using CT and histopathological images. The results show that CNN-based approaches provide

high accuracy in distinguishing malignant from benign lesions in CT images (Rajasekar et al., 2023). Mamatha et al. (2023) developed a DL model for LC detection using CT and histopathological images. In the study, the diagnostic performance was improved by using trained CNN models including VGG-19 and ResNet-50. The model showed superior performance compared to traditional approaches in terms of sensitivity, specificity and F1 score (Mamatha et al., 2023). Devarajan et al. (2023) developed a DL model for automatic LC detection using CT and chest X-ray images. In the study, the use of data augmentation and transfer learning techniques significantly improved model performance and demonstrated the superiority of CNN-based approaches, particularly in early detection (Devarajan et al., 2023). Davri et al. (2023) systematically reviewed DL studies using histological and cytological images for LC diagnosis, prognosis and prediction. The study analysed the effectiveness of different CNN models in differentiating between adenocarcinoma, squamous cell carcinoma and small cell LC, and highlighted that DL-based models improve diagnostic accuracy (Davri et al., 2023). Zhang et al. (2024) developed a DL model using histopathological images to predict prognosis and treatment response in patients with small cell LC (SCLC). The model predicted patient survival by identifying 50 histomorphological phenotype clusters using contrastive clustering. These results demonstrate that DL analysis of pathology images can make an important contribution to clinical decision making (Zhang et al., 2024).

Multimodal and multi-omics approaches have also gained traction. Sangeetha et al. (2024) propose an improved model for LC diagnosis using a deep neural network based on multimodal data fusion (MFDNN). By combining genetic, clinical and image data, the MFDNN model achieved 92.5% accuracy, 87.4% precision and 86.4% sensitivity, higher than previous models. The study shows that multimodal data integration can make a significant contribution to LC diagnosis (Sangeetha et al., 2024). Tran et al. (2024) developed a DL-based decision support system using genomic and proteomic data in LC diagnosis and treatment. The study contributes to LC biomarker discovery and personalised treatment planning by enabling the identification of cancer subtypes through multi-omics data fusion (Tran et al., 2024).

Advanced CNN architectures and 3D modelling have also been explored. Crasta et al. (2023) propose a new DL model based on 3D-VNet and 3D-ResNet for LC diagnosis. The model achieved successful results in segmentation and classification with 99.34%. These results demonstrate the importance of 3D CNN models in LC detection (Crasta et al., 2024). Said et al. (2023) developed a UNETR-based model for image segmentation in LC diagnosis. This model achieved 97.83% segmentation accuracy and 98.77% classification accuracy by using self-supervised learning techniques to discriminate between malignant and benign nodules in 3D CT images (Said et al., 2023).

Risk prediction and survival analysis have been studied by Mikhael et al. (2023) developed a DL model called Sybil that predicts future LC risk from a single low-dose CT (LDCT) scan. The model predicted cancer development within 1 year with high accuracy of 92% AUC (NLST dataset), 86% AUC (MGH dataset) and 94% AUC (CGMH dataset) (Mikhael et al., 2023). Sybil is able to predict individual cancer risk without the need for demographic or clinical data, contributing to personalised screening approaches (Mikhael et al., 2023).

Similarly, Huang et al. (2023) compared machine learning and DL models to predict the survival of LC patients. In the study, the DNN model achieved the highest success with an accuracy of 88.58% compared to traditional logistic regression and decision trees. These results demonstrate the superiority of DL-based approaches in predicting survival time (Huang et al., 2023).

Despite these advancements, most of the existing studies concentrate on a single stage of the diagnostic pipeline—typically classification or segmentation—without integrating key components such as preprocessing, data balancing, architectural optimization, and comprehensive performance comparison. Moreover, data imbalance and generalization across multiple classes remain underexplored challenges. To address these issues, the present study proposes a unified DL-based classification pipeline for LC diagnosis using CT images. The approach includes comprehensive preprocessing, data balancing using the SMOTE, architectural tuning, and comparison of three robust CNN architectures—ResNet101, VGG19, and DenseNet121—on the IQ-OTH/NCCD dataset across three LC categories: benign, malignant, and normal.

3. MATERIAL AND METHOD

3.1. DataSet

The success of deep learning (DL) algorithms largely depends on access to large-scale, high-quality datasets. Sufficiently diverse and well-annotated datasets allow DL models to learn robust and generalizable features, reduce the risk of overfitting or underfitting, and ensure reliable performance across different patient populations. For this reason, standardized, ethically sourced, and clinically representative data are essential in developing accurate diagnostic systems.

In this study, we utilized the IQ-OTH/NCCD lung cancer dataset, a publicly available dataset hosted on the Kaggle platform (AL-Huseiny, 2021). This dataset was collected over three months in Fall 2019 from the Iraq Oncology Teaching Hospital and the National Center for Cancer Diseases. It comprises 1,190 axial CT scan slices from 110 individuals, including both healthy subjects and patients diagnosed with lung cancer at various stages. The cases were labelled and validated by experienced radiologists and oncologists.

The dataset includes three class categories: malignant, benign, and normal. Each CT scan contains between 80 to 200 slices, and images were originally acquired in DICOM format using a Siemens SOMATOM scanner. The imaging protocol included 120 kV tube voltage, 1 mm slice thickness, window width ranging from 350 to 1200 HU, and window center between 50 and 600 HU, with breath-hold at full inspiration. For compatibility with CNN models, the images were converted to JPEG format and resized to 224×224 pixels. The class distribution is shown in Table 1.

The dataset reflects real-world imaging variability and includes individuals from diverse geographic and socioeconomic backgrounds, particularly from central Iraq (e.g., Baghdad, Wasit, Diyala, Salahuddin, and Babylon).

Table 1. *Class Distribution of the IQ-OTH/NCCD Lung Cancer Dataset*

Class	Number of Images	Description
Normal	101	Healthy lung CT slices
Benign	432	Non-cancerous pulmonary nodules
Malignant	657	Cancerous lung tissue samples
Total	1190	

However, the publicly accessible version does not contain structured demographic information such as age, gender, or clinical history in machine-readable format. As a result, subgroup analyses based on patient-level metadata could not be conducted, which is acknowledged as a limitation of this study. All data were anonymized prior to release. The dataset's usage complies with ethical guidelines, and approval was granted by the institutional review boards of the participating medical centers. Written informed consent was waived as per local ethical protocols.

3.2. DL Algorithms

Deep Learning (DL) has emerged as a powerful tool in the field of healthcare, particularly in medical imaging and cancer diagnosis. By leveraging neural networks capable of learning complex patterns, DL enables automated analysis of radiological images, reducing the burden on clinicians and enhancing diagnostic accuracy. In oncology, DL models are widely used to detect and classify tumors from medical scans such as CT, MRI, and PET images. These models can identify subtle patterns that might be overlooked by human experts, leading to earlier and more precise diagnoses.

In cancer imaging, DL-based approaches are especially valuable for tasks such as tumor segmentation, classification, and progression monitoring. For lung cancer, convolutional CNN play a crucial role in detecting pulmonary nodules and distinguishing between benign and malignant lesions. By integrating DL into clinical workflows, healthcare professionals can improve diagnostic efficiency, personalize treatment plans, and ultimately enhance patient outcomes. As advancements continue, DL is expected to further revolutionize cancer imaging by offering real-time analysis, reducing diagnostic errors, and facilitating early intervention strategies.

In this study, three deep learning models ResNet101, VGG19, and DenseNet121 were selected for performance comparison in lung cancer classification. These models were chosen based on their proven effectiveness in medical image analysis, particularly in classification tasks involving CT and histopathological data. Additionally, these models offer architectural diversity in terms of depth and parameter complexity, enabling a comprehensive evaluation of model behaviour under the same training conditions. Their wide spread use in peer-reviewed literature ensures that the results can be reliably compared to existing studies and generalized to similar diagnostic contexts.

3.2.1. Denset121

DenseNet121 is a CNN architecture in which each layer creates dense connections by using the outputs from all previous layers. This structure prevents overlearning by increasing parameter efficiency and minimises the problem of gradient loss. Instead of connecting layers sequentially as in traditional CNNs, each layer receives input directly from all previous layers and transmits its output to all subsequent layers. DenseNet121 is widely used in areas that require high accuracy, such as medical image analysis, as it provides more efficient feature extraction with fewer parameters (Aslan, 2025). In a densely connected CNN, the input of each layer is expressed as follows:

$$x_1 = H_n([x_0, x_1, x_2, \dots, x_{n-1}]) \quad (1)$$

Where x_n is the output in n layers, H_n is the activation function and $x_0, x_1, x_2, \dots, x_{n-1}$ is the combination of the outputs of all previous layers.

3.2.2. ResNet50

ResNet101 (Residual Network) is a CNN architecture developed to address the problem of gradient loss in deep neural networks. Similar to ResNet50, it employs residual connections to facilitate learning and enable the construction of much deeper networks (Eren et al., 2024). ResNet101 is a 101-layer deep CNN architecture, widely used for tasks such as medical image classification, object detection, and segmentation. Due to its increased depth, it can capture more complex features, making it suitable for high-precision image analysis applications. It offers high accuracy and generalisation capacity, increasing the trainability of deeper networks. Instead of the standard linear transformation, ResNet blocks are modelled with residual connections as follows:

$$H(x) = F(x, W) + x \quad (2)$$

Here, $H(x)$ represents the activation value extracted from the block, while $F(x, W)$ denotes the conventional convolutional transformation. In addition, this transformation, computed using W weights, represents the filtering operations performed by the model on the input data x .

3.2.3. VGG19

VGG19 is a 19-layer CNN model with a deep but simple structure. The basic principle of the model is to create deeper feature maps and perform high accuracy classifications by using small (3x3) convolutional filters. Due to its simple yet effective architecture, it is widely preferred in tasks such as medical image analysis, object recognition and segmentation (Çelik & İnik, 2023). However, due to the large number of parameters, the computational cost is higher compared to other CNN architectures. In the VGG architecture, the active feature map in each convolutional layer is computed as follows:

$$y = \sigma(W * x + b) \quad (3)$$

Where W is the convolution kernel, “*” is the convolution process, b is the bias term and σ is the activation function. The VGG architecture performs a successful classification by extracting the features of the images at different scales with the max-pooling process following successive convolutional layers.

3.3. Evaluation of Performances

The evaluation of DL models in this study was conducted using various performance metrics. Additionally, confusion matrices, as well as model accuracy and error rates, were analyzed to provide a detailed insight into prediction performance. The accuracy score indicates how well the model identifies correct classifications across different categories, whereas precision measures the proportion of correctly predicted positive cases. Recall reflects the model’s effectiveness in detecting true positive cases, while the F1 score balances these two measures, ensuring a comprehensive assessment of classification success.

The Confusion Matrix shows the correct and incorrect predictions made by the model, allowing the false positive (FP) and false negative (FN) rates to be analysed. This provides a detailed insight into which classes the model is more successful in and which errors it makes, particularly in LC diagnosis.

Model Accuracy and Model Error curves were analysed to understand the behaviour of the model during the training process. The Model Accuracy curve shows the change in the accuracy of the model during the training process, while the Model Error curve analyses the evolution of the model's error rates over time. These analyses helped to identify potential problems such as overfitting or underfitting. The results obtained provide a comprehensive performance analysis to assess how reliable the model is in LC diagnosis and whether it is suitable for clinical applications. Equations (4)-(7) allow the shortcomings of the model to be identified and areas for improvement to be identified for future studies.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$PRC = \frac{TP}{TP + FP} \quad (5)$$

$$RCL = SNS = \frac{TP}{TP + FN} \quad (6)$$

$$F1_{score} = \frac{2 * RCL * PRC}{RCL + PRC} \quad (7)$$

Accuracy (ACC), Precision (PRC), Recall (RCL) and Sensitivity (SNS).

3.4. Data Preprocessing, Balancing, and Model Optimization

Data preparation is essential for improving model performance in DL-based image classification tasks because it guarantees consistency and lowers noise in the input data. In this study, all CT images from the IQ-OTHNCCD LC dataset underwent multiple preprocessing steps before being fed into the DL model. These

steps include image resizing, grayscale conversion, noise reduction via Gaussian blur, and normalization. Each of these techniques contributes to improving model generalizability and robustness. Since the IQ-OTHNCCD LC dataset contains images of varying dimensions, it is essential to resize them to a uniform shape to maintain consistency in input size.

To ensure uniformity, all images were rescaled to 256×256 pixels through bilinear interpolation. This specific resolution was selected to optimize both processing efficiency and the retention of important details. A smaller resolution might lead to loss of crucial structural information, whereas a much higher resolution would increase computational costs without significant performance gains. Medical images often contain redundant colour information that does not contribute significantly to classification tasks. Converting CT images to grayscale helps remove unnecessary RGB colour channels, thereby reducing computational complexity and memory usage. More importantly, grayscale conversion enables the DL model to focus on essential structural patterns, such as edges, textures, and contrast variations that distinguish benign, malignant, and normal lung tissues.

Medical images frequently contain noise due to variations in scanning conditions, patient movement, or equipment differences. To enhance image quality and suppress unwanted artifacts, Gaussian blur was applied with a kernel size of (5,5). This technique smooths the image while preserving critical edges and texture structures. This technique smooths the image while preserving critical edges and texture structures. The result is a reduction in small, irrelevant variations that might otherwise lead to misclassification by the DL model. A normalization process was implemented to standardize pixel values across all images, adjusting them to fall within a 0 to 1 range.

One of the primary challenges in medical datasets is the imbalance between different classes, which can lead to biased model predictions. In the IQ-OTHNCCD LC dataset, there was a significant imbalance between benign, malignant, and normal cases. Without addressing this issue, the model might become skewed towards the majority class, thereby reducing its ability to accurately classify underrepresented classes.

To mitigate this issue, the SMOTE was employed. Unlike simple duplication of existing data points, SMOTE generates new synthetic examples for underrepresented classes, improving class balance. This method follows a structured approach: first, it randomly selects a minority class sample, then identifies its k -nearest neighbours, and finally, it synthesizes new data points by interpolating between the chosen instance and one of its neighbors. After SMOTE, all classes were balanced, resulting in an equal number of samples across categories. This adjustment allowed the DL model to learn from underrepresented cases effectively, thereby improving classification accuracy across all categories. Optimizing a DL model involves selecting appropriate architectural components, tuning hyperparameters, and employing regularization techniques to prevent overfitting. In this study, the ResNet101, VGG19 and DenseNet121 models were chosen as the base architecture, with several modifications to optimize its performance.

ResNet101 is a convolutional neural network that has been pre-trained on extensive image datasets. Rather than building the model from the ground up, transfer learning was employed to make use of features learned during prior training. Several modifications were introduced, including freezing convolutional layers and integrating custom fully connected layers. This approach allows for quicker model adaptation and improved generalization on the IQ-OTHNCCD LC dataset, as it utilizes previously extracted low-level features such as edges and textures while concentrating on learning more advanced patterns associated with lung cancer detection. To mitigate overfitting and avoid unnecessary computations, an early stopping strategy was implemented. This mechanism continuously evaluates validation loss throughout training and ceases further iterations if no improvement is detected within a predetermined number of epochs (set to 3 epochs). By doing so, the model prevents excessive training, which could otherwise result in poor generalization to new, unseen data. These techniques expanded the training dataset by introducing variations through transformations, including rotation, flipping, zooming, and contrast adjustments. Such alterations enhance the model's ability to recognize lung cancer patterns under different imaging conditions. formations such as: Rotation, Flipping, Zooming and Contrast Adjustment.

3.5. Training Strategy for the Proposed DL Model

The development process of the proposed model consists of four main stages: data preparation, data separation and normalisation, DL model construction, model training and model evaluation. In the first stage, images were processed using the IQ-OTHNCCD LC dataset, and this process is shown in Figure 1. During image resizing and preprocessing, all data were resized to 256×256 pixels and noise reduction methods were applied. In addition, the dataset was randomly shuffled by determining the class labels. In the second step, the pixel values of the images were normalised in the range of 0-1 and the dataset was divided into 75% training and 25% testing. To eliminate class imbalances, the SMOTE method was applied to ensure a balanced data distribution in the training set. In the third step, a ResNet101 based DL model was created, pre-trained layers were frozen and additional classification layers were added. The model was trained using the Adam optimization algorithm and a sparse categorical cross-entropy loss function. To improve training efficiency, an early stopping strategy was applied. Model predictions were validated against a separate dataset, and performance was evaluated using a confusion matrix and a classification report.

All model training, evaluation, and preprocessing operations were conducted using Python 3.8 in a Linux-based Google Colab Pro environment. The system was equipped with $2 \times$ NVIDIA Tesla T4 GPUs (16 GB each) and 25 GB RAM, which enabled efficient parallel computation during model training. The models were implemented using TensorFlow 2.x and the Keras API. Data handling and analysis were performed using Pandas, NumPy, and Scikit-learn 1.1.3. Image preprocessing utilized OpenCV, PIL, and imageio, while Matplotlib, Seaborn, and Plotly were used for visualization. Data balancing was applied using the Synthetic Minority Over-sampling Technique (SMOTE) from the imblearn library, and model training incorporated

early stopping and data augmentation using ImageDataGenerator. The experimental workflow was designed to be reproducible and aligned with best practices in deep learning-based medical image analysis.

This study aims to classify CT images using DL models for LC diagnosis. Advanced CNN architectures with powerful feature extraction capabilities such as ResNet101, VGG19 and DenseNet121 were used in the study. The dataset used is from the IQ-OTH/NCCD LC dataset and consists of 1,190 CT scan slices classified into three different classes as benign, malignant and normal. The main difference this study brings to the literature is that it provides a holistic approach to preprocessing, data imbalance removal and efficient optimisation of model architectures. In contrast to traditional methods, the SMOTE is used to address the class imbalance problem and the effect of the unbalanced distribution of the dataset on model performance is analysed. In addition to the final layers of the pre-trained models, additional optimisation layers were integrated to enable the model to better discriminate between different classes. In addition, early stopping mechanisms and data augmentation techniques were used in the training process to prevent overlearning and increase the generalisation capability of the model. The study comprehensively compares different DL architectures and analyses their accuracy rates, especially in the classification of malignant cases.

4. EXPERIMENTAL RESULTS

The experimental investigations carried out to assess the effectiveness of the suggested model and the outcomes are shown in this part. This study uses the IQ-OTHNCCD LC dataset to analyze the classification performance of various DL models. Data separation, model training, and evaluation procedures were applied methodically, beginning with preprocessing activities. Basic measures including accuracy, sensitivity, specificity, and F1 score were used to measure the model's performance and compare it to other studies.

Figure 2 shows examples of lung CT scans. The images are categorised as benign (a), malignant (b) and normal (c) to represent different lung conditions. The benign (a) image shows generally benign lesions or abnormal but non-cancerous tissue formations. A malignant (b) image may show marked irregularities in the lung tissue, indicating the presence of malignant cells. a normal (c) image shows no abnormal formation and the lung tissue appears to be healthy.

Figure 3 shows the confusion matrices generated to compare the performance of the DL models. Each matrix visualises the accuracy and error rates by showing the relationship between the predictions made by the model and the actual classifications. Figure 3a shows the confusion matrix for the ResNet101 model.

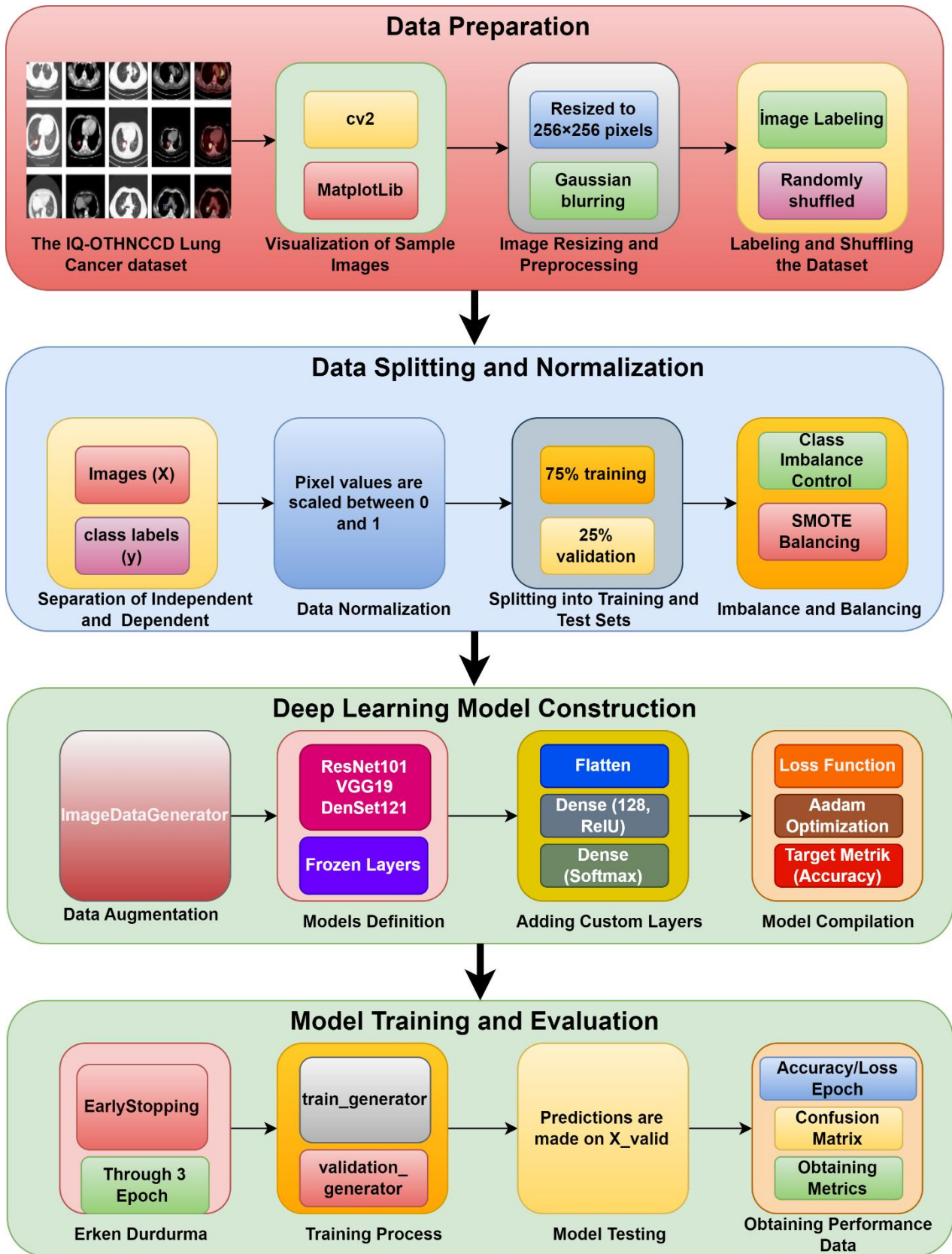


Figure 1. Flowchart of data processing and model training for LC classification model

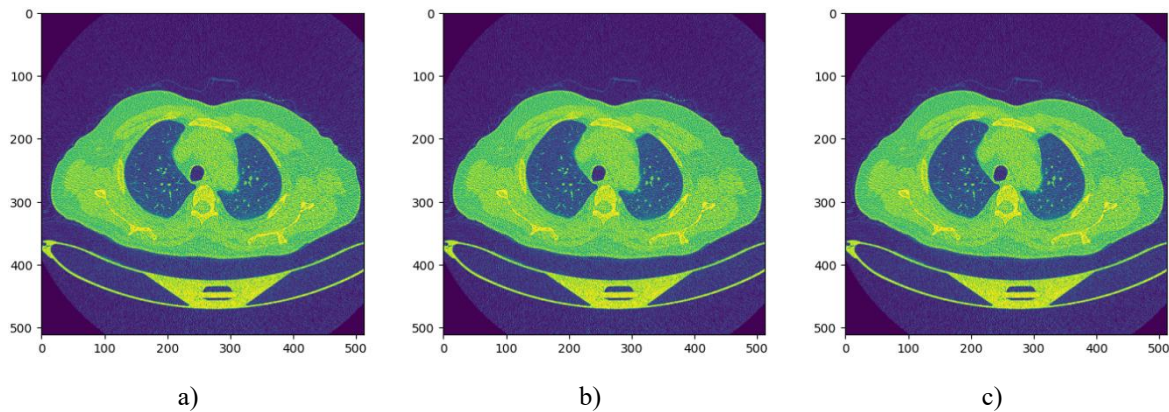


Figure 2. Examples of lung CT images. **a)** Benign, **b)** Malignant, **c)** Normal

The model classified malignant cases with 100% accuracy (143 correct predictions, 0 incorrect). However, 8 normal cases were misclassified as benign. This shows that the model is quite successful for the malignant class, but makes some mistakes in distinguishing between benign and normal classes. Figure 3b shows the performance of the VGG19 model. The model shows high accuracy for malignant cases, making only 1 error in this class. However, 7 normal cases were misclassified as benign. These results show that the model is successful in discriminating malignant cases, but is confused between benign and normal classes. Figure 3c shows the confusion matrix of the Dense121 model. Although the model correctly classified a large proportion of malignant cases (138 correct predictions, 4 errors), it misclassified more cases in the benign class than the other models. In particular, 9 benign cases were incorrectly predicted, indicating that the model performed worst in distinguishing between benign and normal classes.

The ResNet101 model stands out as the most successful model in terms of accurate prediction of malignant classes. However, incorrect prediction of normal cases as benign is a factor that can affect the overall performance of the model. While the VGG19 model has high accuracy in malignant class, it has limited performance in discriminating between normal and benign classes. The Dense121 model made more errors in predicting the benign class compared to the other models. The results show that each model has different classification abilities and that model selection should be made accordingly in applications focusing on a specific class.

Figure 4 shows the accuracy and loss values of the DL models during the training process. The performance of the ResNet101 (a), VGG19 (b) and DenseNet121 (c) models on training and validation sets are compared. The accuracy values show a steady increase as the training process progresses for all models. The fact that the training and validation accuracies are close to each other shows that the overall performance of the models is consistent. Analysing the loss values reveals that all models experience a notable decline at first, but that the subsequent epochs show a more balanced trend. Although the VGG19 model shows fluctuations in verification loss in some epochs, it generally shows a steady downward trend. The ResNet101 and DenseNet121 models, on the other hand, show more stable loss values and offer high performance in the validation set. When the results are evaluated, it can be seen that all models have successfully completed the training process. The

ResNet101 model achieved the highest accuracy value, while the DenseNet121 model showed a fast-learning process in the early epochs. The VGG19 model, on the other hand, shows a generally successful performance, although it fluctuates in verification accuracy from time to time. The results show that the models work consistently and effectively on different datasets.

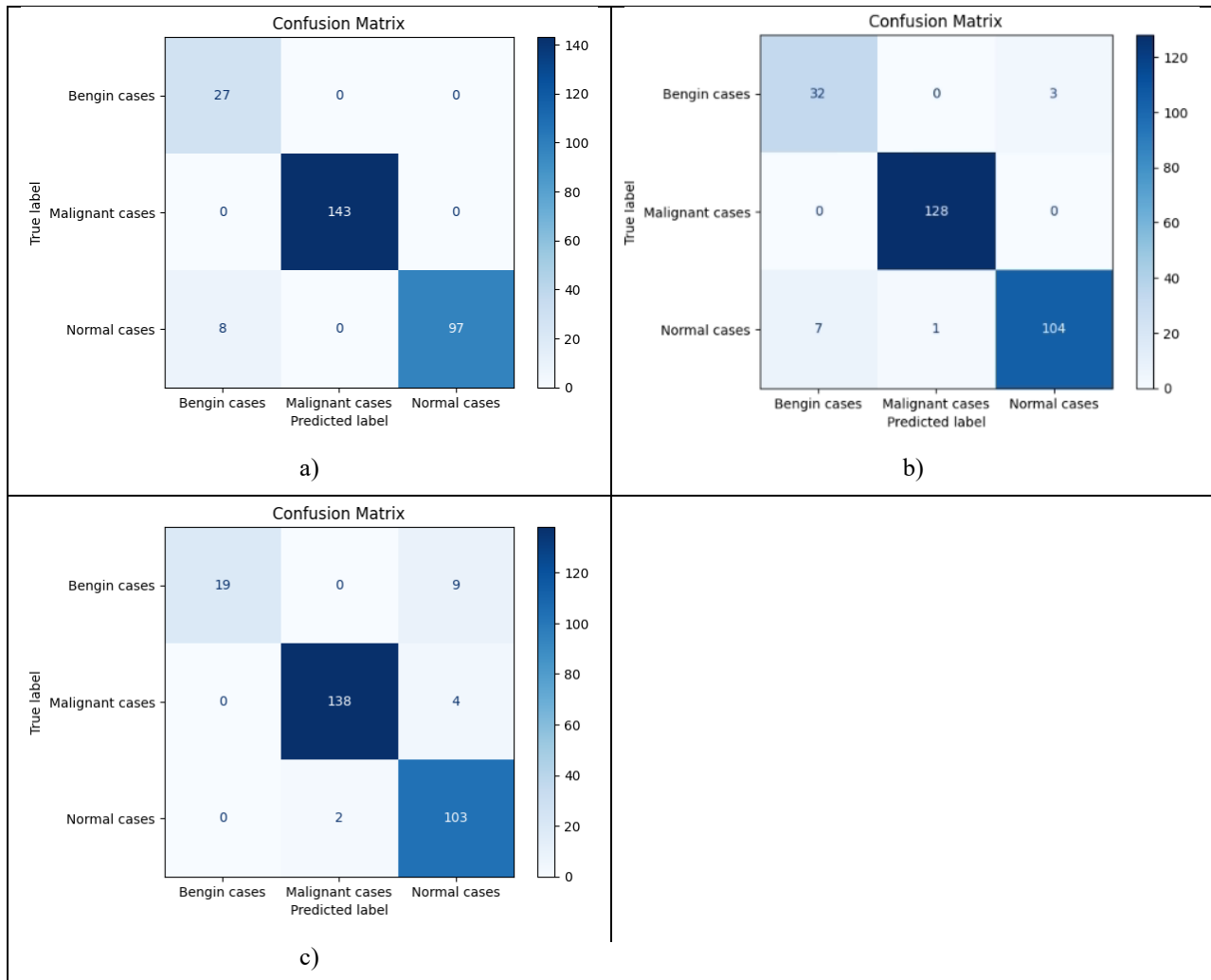


Figure 3. Confusion Matrices for the DL model. **a)** ResNet101, **b)** VGG19, **c)** DenseNet121

Figure 5 shows the ROC curve of the ResNet101 model. The curve almost reaches a true positive rate of 1 and maintains a false positive rate close to zero. The AUC (Area Under the Curve) value of 1.00 indicates that the model's classification ability is extremely high, with very few, if any, false positives or negatives. This result highlights the model's reliability, particularly in critical applications such as medical image classification, where accurate predictions are essential.

Figure 6 presents the Grad-CAM heatmaps for the normal, benign, and malignant cases, illustrating the model's focus on specific regions of the CT images for classification. In the normal case, the heatmap displays a relatively uniform distribution, highlighting the typical anatomical structures of the lungs without significant pathological changes. For the benign case, the heatmap indicates localized regions of interest, suggesting areas where benign abnormalities are detected. In contrast, the malignant case exhibits a more concentrated heatmap, with heightened intensity in certain regions, reflecting the model's identification of malignant features. These

results demonstrate how the Grad-CAM technique enables a visual interpretation of the model's attention, providing valuable insights into the decision-making process for classifying different types of pathologies in medical imaging.

Table 2 presents the classification performance of different DL models based on ACC, RCL/SNS, SPC, and F1-score.

Table 2. Performance Comparison of DL Models

Algorithm	ACC	RCL/SNS	SPC	F1 score
ResNet101	0.98	0.98	0.98	0.98
VGG19	0.96	0.96	0.96	0.96
DenseNet121	0.95	0.88	0.97	0.91

The ResNet101 model achieves the highest performance across all metrics, with an ACC of 0.98. The fact that RCL/SNS and SPC values are equal indicates that the model effectively balances positive and negative classifications. The VGG19 model also demonstrates strong performance, achieving an ACC of 0.96, along with equal RCL/SNS and SPC values, indicating stable classification performance. In contrast, the DenseNet121 model exhibits a relatively lower performance, particularly in terms of RCL/SNS, which is 0.88, despite showing competitive SPC (0.97) and F1-score (0.91). Overall, the ResNet101 model outperforms the other models, delivering the highest classification success. While the VGG19 model provides satisfactory accuracy and balanced sensitivity and specificity, the DenseNet121 model, despite strong SPC and F1-score values, falls behind in terms of sensitivity. These variations highlight critical performance differences that should be considered when selecting models for classification tasks. Table 3 presents a comparison between the classification performance of the proposed ResNet101 model and previous studies in the literature.

Table 3. Comparison of performance of DL models with literature

Algoritim	ACC	RCL/SNS	SPC	F1 score
(Asuntha & Srinivasan, 2020)	0.95	0.97	0.96	-
(Shafi et al., 2022)	0.94	0.94	0.95	0.94
(Kumar et al., 2024)	0.95	-	-	0.85
(Mohamed & Ezugwu, 2024)	0.97	0.97	0.97	0.97
ResNet101(Proposed Model)	0.98	0.98	0.98	0.98

The evaluation, based on key metrics such as ACC, RCL/SNS, SPC, and F1-score, highlights that the proposed model achieves superior results compared to existing approaches. The ResNet101 model demonstrates the highest scores across all metrics, with 98% ACC, 98% RCL/SNS, 98% SPC, and a 98% F1-score. The highest ACC rate in the literature is 97%, as reported by (Mohamed & Ezugwu, 2024). However, although this study is at the same level as the proposed model in terms of RCL/SNS and SPC, it lags behind in terms of ACC.

Studies by (Asuntha & Srinivasan, 2020) and (Shafi et al., 2022) reported ACC values of 95% and 94%, respectively, showing lower performance than the proposed model. Furthermore, (Kumar et al., 2024) did not

report RCL/SNS and SPC values, and the F1-score was the lowest at 0.85%. In comparison, the proposed ResNet101 model proves to be more successful, exceeding the best results in the literature in terms of both ACC and other classification metrics.

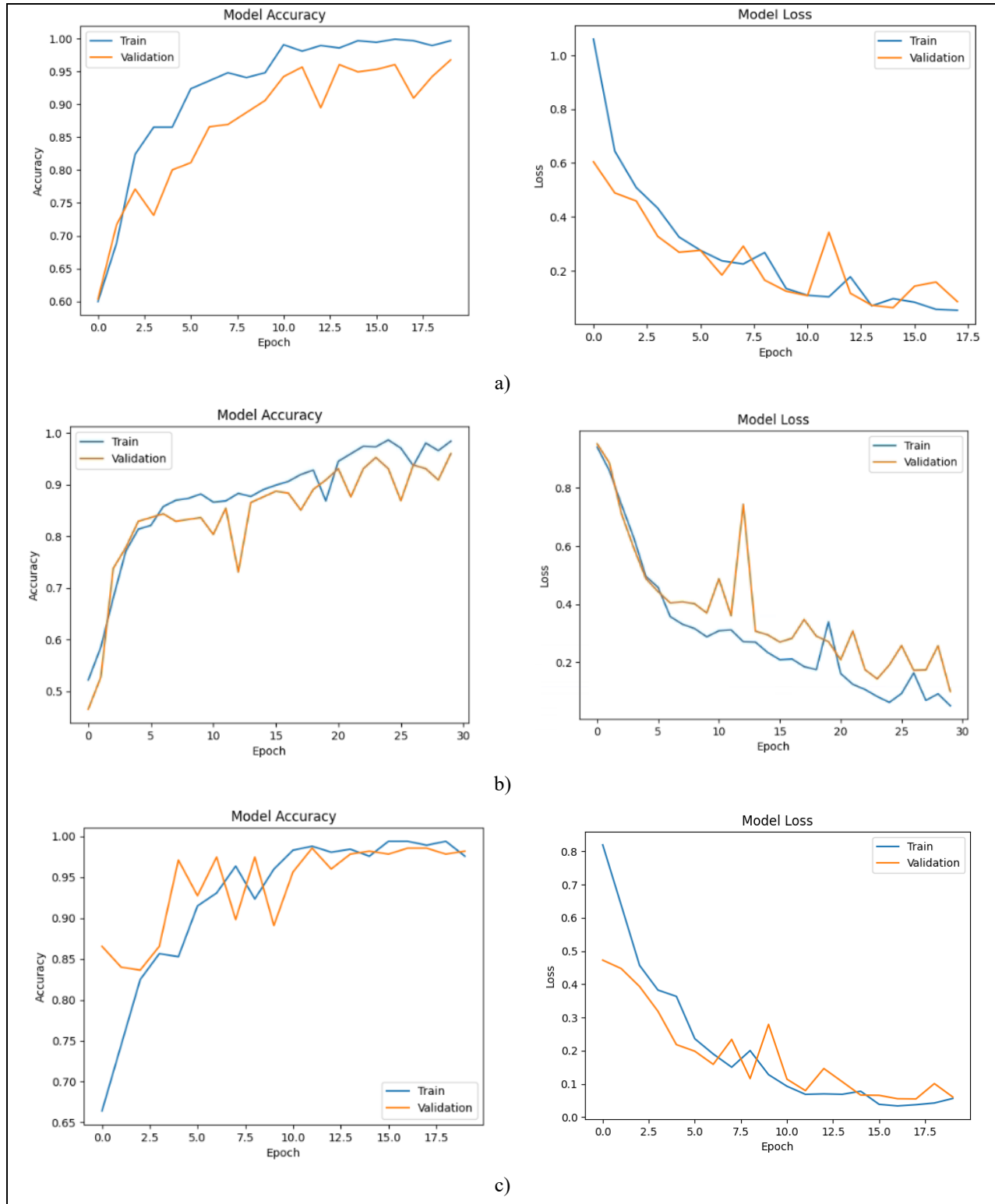


Figure 4. Accuracy/Loss curves for the training process of the DL model. **a)** ResNet101, **b)** VGG19, **c)** DenseNet121

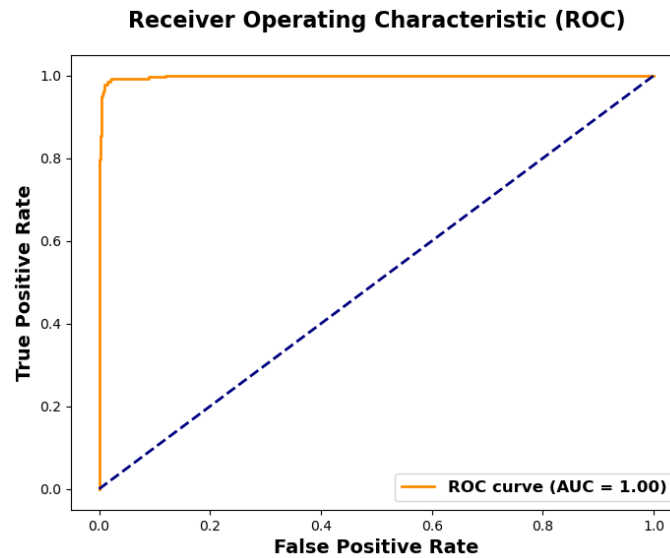


Figure 5. ROC Curve of ResNet101 Model and Performance Evaluation

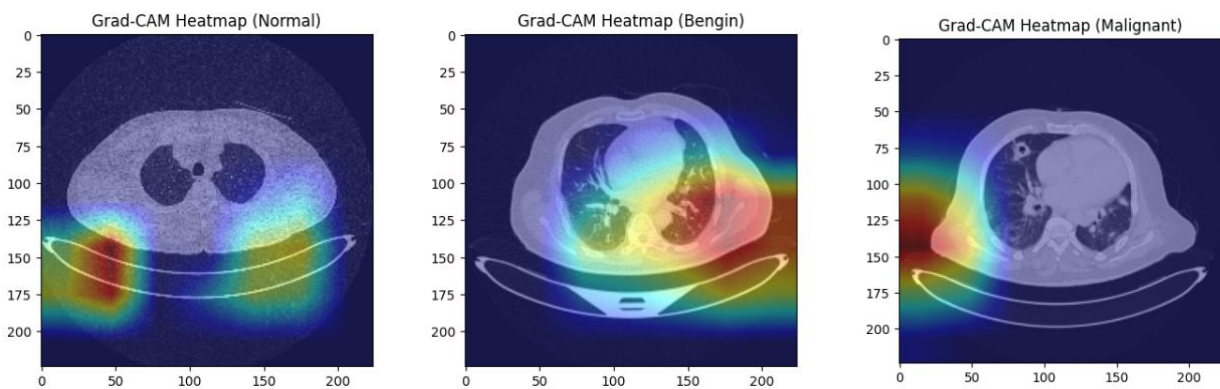


Figure 6. Grad-CAM Heatmaps for Normal, Benign, and Malignant Cases

This study demonstrated that the proposed ResNet101 model provided superior performance in classifying lung CT images, achieving the highest accuracy, sensitivity, specificity, and F1-score (all 98%) among the tested deep learning models (VGG19, DenseNet121) and existing studies in the literature. ROC analysis (AUC: 1.00) and Grad-CAM heatmaps further confirmed the model's reliability and enhanced interpretability by visualizing regions critical to classification. However, a limitation of the current study is the absence of sociodemographic information in the dataset, which may influence model generalizability.

5. CONCLUSION

This study investigated the effectiveness of DL models in LC diagnosis using CT images. Three CNN architectures—ResNet101, VGG19, and DenseNet121—were evaluated on the IQ-OTH/NCCD dataset to classify lung conditions as benign, malignant, or normal. To enhance model performance, class imbalance was addressed using SMOTE, data augmentation techniques were applied, and an early stopping mechanism was implemented to prevent overfitting.

The experimental results show that the ResNet101 model achieved the highest accuracy (98%) and outperformed other architectures in all key classification metrics. Notably, its superior performance in distinguishing malignant cases increases its potential for clinical applications. Comparisons with previous studies indicate that the proposed model offers superior accuracy and generalization capability. Despite these promising findings, some misclassifications occurred between benign and normal cases, suggesting that further improvements in feature extraction could enhance performance. Future research should explore hybrid architectures, multi-modal approaches, and the benefits of explainable AXI to improve interpretability and trust in clinical settings.

In conclusion, this study highlights the potential of DL-based models as a powerful tool for LC diagnosis, offering high accuracy, reliability, and efficiency. Integrating such models into clinical decision support systems could significantly improve early detection and patient outcomes.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- AL-Huseiny, M. (2021). National Cancer Center Database (IQ-OTH/NCCD - Lung Cancer Dataset) [<https://www.kaggle.com/datasets/adityamahimkar/iqothnccd-lung-cancer-dataset/data>]. In *Kaggle*. <https://doi.org/doi: 10.17632/bhmdr45bh2.2>
- Aslan, E. (2025). Development of malaria diagnosis with convolutional neural network architectures: a CNN-based software for accurate cell image analysis. *ITEGAM-JETIA*, 11(51), 35–42. <https://doi.org/10.5935/JETIA.V11I51.1392>
- Asuntha, A., & Srinivasan, A. (2020). Deep learning for lung Cancer detection and classification. *Multimedia Tools and Applications*, 79(11–12), 7731–7762. <https://doi.org/10.1007/s11042-019-08394-3>
- Çelik, M., & İnik, Ö. (2023). Detection of Monkeypox Among Different Pox Diseases with Different Pre-Trained Deep Learning Models. *Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 13(1), 10–21. <https://doi.org/10.21597/JIST.1206453>
- Crasta, L. J., Neema, R., & Pais, A. R. (2024). A novel Deep Learning architecture for lung cancer detection and diagnosis from Computed Tomography image analysis. *Healthcare Analytics*, 5(1) 100316. <https://doi.org/10.1016/J.HEALTH.2024.100316>
- Davri, A., Birbas, E., Kanavos, T., Ntritsos, G., Giannakeas, N., Tzallas, A. T., & Batistatou, A. (2023). Deep Learning for Lung Cancer Diagnosis, Prognosis and Prediction Using Histological and Cytological Images: A Systematic Review. *Cancers* 2023, 15(15), 3981. <https://doi.org/10.3390/CANCERS15153981>
- Devarajan, H. R., Balasubramanian, S., Kumar Swarnkar, S., Kumar, P., & Jallepalli, V. R. (2023). Deep Learning for Automated Detection of Lung Cancer from Medical Imaging Data. International Conference on Artificial Intelligence for Innovations in Healthcare Industries (ICAIIHI 2023).

<https://doi.org/10.1109/ICAIIHI57871.2023.10488962>

- Eren, B., Fen, Ü., Dergisi, B., Aslan, E., & Özüpak, Y. (2024). Classification of Blood Cells with Convolutional Neural Network Model. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 13(1), 314–326. <https://doi.org/10.17798/BITLISFEN.1401294>
- Huang, S., Arpaci, I., Al-Emran, M., Kılıçarslan, S., & Al-Sharafi, M. A. (2023). A comparative analysis of classical machine learning and deep learning techniques for predicting lung cancer survivability. *Multimedia Tools and Applications*, 82(22), 34183–34198. <https://doi.org/10.1007/s11042-023-16349-y>
- Javed, R., Abbas, T., Khan, A. H., Daud, A., Bukhari, A., & Alharbey, R. (2024). Deep learning for lungs cancer detection: a review. *Artificial Intelligence Review*, 57(8), 1–39. <https://doi.org/10.1007/S10462-024-10807-1>
- Kumar, V., Prabha, C., Sharma, P., Mittal, N., Askar, S. S., & Abouhawwash, M. (2024). Unified deep learning models for enhanced lung cancer prediction with ResNet-50–101 and EfficientNet-B3 using DICOM images. *BMC Medical Imaging*, 24(1), 1–21. <https://doi.org/10.1186/S12880-024-01241-4>
- Mamatha, B., Rashmi, D., Tiwari, K. S., Sikrant, P. A., Jovith, A. A., & Reddy, P. C. S. (2023). Lung Cancer Prediction from CT Images and using Deep Learning Techniques. 2023 2nd International Conference on Trends in Electrical, Electronics and Computer Engineering (TEECCON 2023), 263–267. <https://doi.org/10.1109/TEECCON59234.2023.10335801>
- Mikhael, P. G., Wohlgend, J., Yala, A., Karstens, L., Xiang, J., Takigami, A. K., Bourgouin, P. P., Chan, P., Mrah, S., Amayri, W., Juan, Y. H., Yang, C. T., Wan, Y. L., Lin, G., Sequist, L. V., Fintelmann, F. J., & Barzilay, R. (2023). Sybil: A Validated Deep Learning Model to Predict Future Lung Cancer Risk from a Single Low-Dose Chest Computed Tomography. *Journal of Clinical Oncology*, 41(12), 2191–2200. <https://doi.org/10.1200/JCO.22.01345>
- Mohamed, T. I. A., & Ezugwu, A. E. S. (2024). Enhancing Lung Cancer Classification and Prediction With Deep Learning and Multi-Omics Data. *IEEE Access*, 12(1), 59880–59892. <https://doi.org/10.1109/ACCESS.2024.3394030>
- Mohamed, T. I. A., Oyelade, O. N., & Ezugwu, A. E. (2023). Automatic detection and classification of lung cancer CT scans based on deep learning and ebola optimization search algorithm. *PLOS ONE*, 18(8), 0285796. <https://doi.org/10.1371/JOURNAL.PONE.0285796>
- Özdemir, B., Aslan, E., & Pacal, I. (2025). Attention Enhanced InceptionNeXt Based Hybrid Deep Learning Model for Lung Cancer Detection. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2025.3539122>
- Said, Y., Alsheikhy, A. A., Shawly, T., & Lahza, H. (2023). Medical Images Segmentation for Lung Cancer Diagnosis Based on Deep Learning Architectures. *Diagnostics* 2023, 13(3), 546. <https://doi.org/10.3390/DIAGNOSTICS13030546>
- Shafi, I., Din, S., Khan, A., Díez, I. D. L. T., Casanova, R. del J. P., Pifarre, K. T., & Ashraf, I. (2022). An Effective Method for Lung Cancer Diagnosis from CT Scan Using Deep Learning-Based Support Vector Network. *Cancers*, 14(21), 5457. <https://doi.org/10.3390/CANCERS14215457>

- Tárnoki, Á. D., Tárnoki, D. L., Dąbrowska, M., Knetki-Wróblewska, M., Frille, A., Stubbs, H., Blyth, K. G., & Juul, A. D. (2024). New developments in the imaging of lung cancer. *Breathe*, 20(1), 230176. <https://doi.org/10.1183/20734735.0176-2023>
- Thandra, K. C., Barsouk, A., Saginala, K., Aluru, J. S., & Barsouk, A. (2021). Epidemiology of lung cancer. *Contemporary Oncology*, 25(1), 45. <https://doi.org/10.5114/WO.2021.103829>
- Tran, T. O., Hoa Vo, T., & Khanh Le, N. Q. (2024). Omics-based deep learning approaches for lung cancer decision-making and therapeutics development. *Briefings in Functional Genomics*, 23(3), 181–192. <https://doi.org/10.1093/BFGP/ELAD031>
- Wang, L. (2022). Deep Learning Techniques to Diagnose Lung Cancer. *Cancers*, 14(22), 5569. <https://doi.org/10.3390/CANCERS14225569>
- Wani, N. A., Kumar, R., & Bedi, J. (2024). DeepXplainer: An interpretable deep learning based approach for lung cancer detection using explainable artificial intelligence. *Computer Methods and Programs in Biomedicine*, 243(1), 107879. <https://doi.org/10.1016/J.CMPB.2023.107879>
- Zhang, Y., Yang, Z., Chen, R., Zhu, Y., Liu, L., Dong, J., Zhang, Z., Sun, X., Ying, J., Lin, D., Yang, L., & Zhou, M. (2024). Histopathology images-based deep learning prediction of prognosis and therapeutic response in small cell lung cancer. *Npj Digital Medicine* 2024, 7(1), 1–12. <https://doi.org/10.1038/s41746-024-01003-0>



Improved Chef-Based Optimization Algorithm with Chaos-Based Fitness Distance Balance for Frequency-Constrained Truss Structures

Ayşe BEŞKİRLİ^{1*}

¹ Karamanoğlu Mehmetbey University, Faculty of Medicine, Karaman, Türkiye

Keywords	Abstract
Chaotic Maps Fitness Distance Balance Truss Structure Chef-based Optimization Algorithm	Chef-based optimization algorithm (CBOA), one of the recently proposed metaheuristic algorithms, is a population-based optimization algorithm inspired by the process of students becoming skilled chefs after receiving training from chef instructors in a culinary academy. In order to improve the performance of CBOA, seven different CBOA variants are proposed in this study, which are improved with three different chaotic maps, fitness distance balance strategy and their combinations. The effectiveness of the proposed CBOA variants is first evaluated by testing them on 16 different benchmark functions. Then, the proposed CBOA variants are applied to frequency constrained 37-bar and 52-bar truss problems to evaluate their performance on engineering problems. Thus, the success of the proposed CBOA variants on different problems was extensively investigated in three different experimental studies. Among these variants, while FC2-CBOA and FC3-CBOA variants performed well on benchmark functions, FC3-CBOA and C3-CBOA variants performed well on 37-bar and 52-bar truss problems, respectively. The results obtained from these three different experimental studies have shown that each proposed CBOA variant is able to produce effective results depending on the problem type.
Cite	
	Beşkirli, A. (2025). Improved Chef-Based Optimization Algorithm with Chaos-Based Fitness Distance Balance for Frequency-Constrained Truss Structures. <i>GU J Sci, Part A, 12(2)</i> , 392-416. doi:10.54287/guj.1667182
Author ID (ORCID Number)	Article Process
0000-0002-8694-8438 Ayşe BEŞKİRLİ	Submission Date 27.03.2025 Revision Date 17.04.2025 Accepted Date 08.05.2025 Published Date 30.06.2025

1. INTRODUCTION

In recent years, there has been a growing interest in using metaheuristic algorithms to solve real-world engineering problems (Kiran & Beskirli, 2024). In this direction, there are many new metaheuristic algorithms inspired by natural processes and the behavior of living things. Some of these metaheuristic algorithms are methods such as coati optimization algorithm (COA) (Dehghani et al., 2023), atom search algorithm (ASO) (Zhao et al., 2019), tree-seed algorithm (TSA) (Kiran, 2015), whale optimization algorithm (WOA) (Mirjalili & Lewis, 2016), gorilla troops optimizer (GTO) (Abdollahzadeh et al., 2021), honey badger algorithm (HBA) (Hashim et al., 2022), cleaner fish optimization algorithm (CFO) (Zhang et al., 2024), preschool education optimization algorithm (PEOA) (Trojovský, 2023), secretary bird optimization algorithm (SBOA) (Fu et al., 2024), pelican optimization algorithm (POA) (Trojovský & Dehghani, 2022), fungal growth optimizer (FGO) (Abdel-Basset et al., 2025), zebra optimization algorithm (ZOA) (Trojovská et al., 2022), geyser inspired algorithm (GEA) (Ghasemi et al., 2024), and chef-based optimization algorithm (CBOA) (Trojovská & Dehghani, 2022). Among these methods, CBOA (Trojovská & Dehghani, 2022) is an optimization algorithm

*Corresponding Author, e-mail: aysebes@kmu.edu.tr

inspired by the process of becoming a skilled chef by improving the cooking skills of students and young chefs who attend a culinary academy for cooking. CBOA algorithm has been effectively used in the literature to solve many problems. Some of these studies are as follows: Kutlu Onay (2023), proposed a new CBOA method equipped with various strategies to improve the performance of CBOA in engineering design problems with CEC2019 and CEC2022 test functions. They named their proposed method as CBOADP. They compared the results of CBOADP with the results of various metaheuristic algorithms. As a result of the comparison, they said that CBOADP achieved success in most of the engineering design problems solved with CBOADP. They also performed various statistical analyses to support these results. In their statistical analysis, they stated that CBOADP produced significant results and the proposed algorithm achieved success. Huang et al. (2024) proposed a method based on a hybrid SqueezeNet model and an improved CBOA algorithm for breast cancer detection from mammography images. They developed the CBOA algorithm using the sinusoidal chaotic map. They named the improved CBOA as ICBOA. They first tested the performance of their proposed ICBOA on benchmark functions. After verifying its efficiency, ICBOA was then used together with the hybrid SqueezeNet model for breast cancer detection from images. When they compared the results obtained here with the results of various state-of-the-art methods, they stated that their proposed method has a superior performance. Rajesh Kumar et al. (2024) stated that cardiovascular disease is the main cause of disability and mortality in developing countries and it is very difficult to diagnose it according to the initial determinations. For this reason, they proposed a deep learning method trained with fractional CBOA for cardiovascular risk prediction from retinal fundus images. Balasubramaniam et al. (2024) in their study, they used retinal fundus images to predict blood segmentation and cardiovascular disease with a deep residual network-based Res-UNet model trained with chronological CBOA. Aribowo et al. (2023) in his study, they performed proportional-integral-derivative (PID) parameter tuning on an automatic voltage regulator (AVR) using the CBOA algorithm. When they compared the obtained results with the results of some metaheuristic algorithms, they reported that CBOA performed better.

In the literature, the performance of metaheuristic algorithms is improved by using various strategies to produce effective results in engineering problems. Some of the studies in this direction are as follows: Joni (2024) added both chaotic and elite strategies to the mountain gazelle optimizer (MGO) algorithm for parameter estimation of photovoltaic models. They named the proposed method called CEMGO. According to the experimental results, they said that their proposed CEMGO method is better than the original MGO. Prapanca et al. (2025) added the levy flight strategy to the fata morgana algorithm in their study. They named the proposed method called MFATA. They evaluated the performance of their proposed MFATA on benchmark functions. As a result of the evaluation, they said that their proposed method is competitive. Aribowo et al. (2024) optimized the proportional integral derivative (PID) parameters using the frilled lizard optimization algorithm. Liu et al. (2024) proposed an improved marine predators algorithm (FDBCMPA) with six different chaotic maps and fitness distance balance strategy for optimization of camera calibration. They compared the results of their proposed FDBCMPA with existing methods. As a result of the comparison, they

stated that the performance of their proposed FDBCMPA improved in the optimization of camera calibration. Demirbas et al. (2025) proposed a coati optimization algorithm with opposition-based learning and fitness distance balance strategies (FDBCOA-OBL) for solving the transmission network expansion problem. As a result of the experimental results, they stated that their proposed FDBCOA-OBL has a robust performance.

In this study, seven different CBOA variants improved with three different chaotic maps, fitness distance balance strategy and their combinations are proposed to improve the performance of CBOA. The performance of the proposed CBOA variants is evaluated both on 16 different benchmark functions and on the optimization of 37-bar and 52-bar truss structures, which are real-world engineering problems.

The following sections of this paper are as follows: Section two describes the general structure of CBOA and the general operation of the algorithm. The third section describes the strategies applied to improve the performance of CBOA. Section four comparatively analyzes the performance of the proposed CBOA variants on benchmark functions. Also, the performance of the proposed CBOA variants on both 37-bar and 52-bar truss problems is evaluated. In the last section, the results of the study are summarized and suggestions for future work are given.

2. CHEF-BASED OPTIMIZATION ALGORITHM (CBOA)

Chef-based optimization algorithm (CBOA) is a population-based optimization algorithm inspired by the process of becoming a skilled chef by improving the cooking skills of students and young chefs who attend a culinary academy for cooking (Trojovská & Dehghani, 2022). In a culinary academy, which is assumed to have a certain number of chef instructors, each chef instructor is responsible for running a course. Culinary students can choose which courses to attend based on their interests. Throughout the training, chef instructors teach cooking skills and techniques to the culinary students, while at the same time developing their own skills through individual skills and instruction from the best chef instructor of the culinary academy. The culinary students, in turn, try to learn and imitate the knowledge and skills of the chef instructors and improve what they have learned through practical applications. Thus, thanks to the training they receive during the training period, each culinary student graduates from the culinary academy as a skilled chef. In this context, the mathematical modeling that played a role in the initiation and design process of CBOA is presented under the relevant headings.

2.1. Initiation of CBOA

The population of CBOA consists of two groups: culinary students and chef instructors. The population process of the algorithm is modeled by the matrix given in Equation 1.

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_i \\ \vdots \\ X_N \end{bmatrix}_{N \times d} = \begin{bmatrix} x_{1,1} & \cdots & x_{1,j} & \cdots & x_{1,d} \\ \vdots & & \vdots & & \vdots \\ x_{i,1} & \cdots & x_{i,j} & \cdots & x_{i,d} \\ \vdots & & \vdots & & \vdots \\ x_{N,1} & \cdots & x_{N,j} & \cdots & x_{N,d} \end{bmatrix}_{N \times d} \quad (1)$$

where N is the number of populations and d is the number of dimensions. The starting position of the CBOA is initialized randomly using Equation 2.

$$x_{i,j} = lb_j + r \cdot (ub_j - lb_j) \quad (2)$$

where r is a random number in the range $[0,1]$ ub_j and lb_j are the upper and lower bounds of the j th variable, respectively. The objective function is calculated using the vector given in Equation 3.

$$F = \begin{bmatrix} F_1 \\ \vdots \\ F_i \\ \vdots \\ F_N \end{bmatrix}_{N \times 1} = \begin{bmatrix} F(X_1) \\ \vdots \\ F(X_i) \\ \vdots \\ F(X_N) \end{bmatrix}_{N \times 1} \quad (3)$$

2.2. Mathematical modeling of CBOA

In CBOA, the population consists of two groups: culinary students and chef trainers. The population matrix of CBOA is given in Equations 4 and 5.

$$XS = \begin{bmatrix} XS_1 \\ \vdots \\ XS_{N_C} \\ XS_{N_C+1} \\ \vdots \\ XS_N \end{bmatrix}_{N \times d} = \begin{bmatrix} xS_{1,1} & \cdots & xS_{1,j} & \cdots & xS_{1,d} \\ \vdots & \ddots & \vdots & \cdots & \vdots \\ xS_{N_C,1} & \cdots & xS_{N_C,j} & \cdots & xS_{N_C,d} \\ xS_{N_C+1,1} & \cdots & xS_{N_C+1,j} & \cdots & xS_{N_C+1,d} \\ \vdots & \cdots & \vdots & \ddots & \vdots \\ xS_{N,1} & \cdots & xS_{N,j} & \cdots & xS_{N,d} \end{bmatrix}_{N \times d} \quad (4)$$

$$FS = \begin{bmatrix} FS_1 \\ \vdots \\ FS_{N_C} \\ FS_{N_C+1} \\ \vdots \\ FS_N \end{bmatrix}_{N \times d} \quad (5)$$

The mathematical modeling and functioning of CBOA is basically based on two phases (Trojovská & Dehghani, 2022). The first phase includes updating processes for chef trainers and the second phase includes updating processes for culinary students.

Phase I Update process for chef trainers

At this phase, a certain number of chef trainers in the culinary academy teach cooking skills to cooking students. Chef trainers develop their cooking skills based on two strategies. The first strategy is to imitate the chef instructor by trying to learn the techniques of the best chef instructor in the culinary academy. The new position of chef trainers is calculated using Equation 6. If this new position improves the objective function, it is accepted. This is expressed in Equation 7.

$$xS_{i,j}^{C/S1} = xS_{i,j} + r \cdot (BC_j - I \cdot xS_{i,j}) \quad (6)$$

$$XS_i = \begin{cases} XS_i^{C/S1}, & FS_i^{C/S1} < F_i; \\ XS_i, & \text{else,} \end{cases} \quad (7)$$

With the second strategy, each chef trainer tries to improve their skills by doing individual activities. This process is described by the mathematical formula in Equation 8. Here a random position is generated. If this randomly generated position improves the objective function, it is accepted. This is expressed in Equation 9.

$$xS_{i,j}^{C/S2} = xS_{i,j} + lb_j^{local} + r \cdot (ub_j^{local} - lb_j^{local}) \quad (8)$$

$$XS_i = \begin{cases} XS_i^{C/S2}, & FS_i^{C/S2} < F_i; \\ XS_i, & \text{else,} \end{cases} \quad (9)$$

Phase II Update process for culinary students

In the culinary academy, cooking students learn and improve their cooking skills based on three strategies. According to the first strategy, cooking students randomly choose a course of one of the chef instructors and this chef instructor teaches cooking skills to the students. The aim of this strategy is to have different chef instructors to guide the cooking students and students learn different cooking skills under the guidance of the selected chef instructor. This is done using the formula in Equation 10. New positions are calculated according to Equation 11.

$$xS_{i,j}^{S/S1} = xS_{i,j} + r \cdot (CI_{k,i,j} - I \cdot xS_{i,j}) \quad (10)$$

$$XS_i = \begin{cases} XS_i^{S/S1}, & FS_i^{S/S1} < F_i; \\ XS_i, & \text{else,} \end{cases} \quad (11)$$

According to the second strategy, the culinary student imitates the chef instructor by completely learning one of the chef instructor's skills. This situation is modeled as in Equation 12 and positions are calculated according to Equation 13.

$$xS_{i,j}^{S/S2} = \begin{cases} CI_{k,i,j}, & j = l; \\ xS_{i,j}, & \text{else,} \end{cases} \quad (12)$$

$$XS_i = \begin{cases} XS_i^{S/S2}, & FS_i^{S/S2} < F_i; \\ XS_i, & \text{else,} \end{cases} \quad (13)$$

According to the third strategy, cooking students try to improve their cooking skills through individual studies and activities. This is modeled as in Equations 14 and 15.

$$xS_{i,j}^{S/S3} = \begin{cases} xS_{i,j} + lb_j^{local} + r \cdot (ub_j^{local} - lb_j^{local}), & j = q; \\ xS_{i,j}, & j \neq q, \end{cases} \quad (14)$$

$$XS_i = \begin{cases} XS_i^{S/S3}, & FS_i^{S/S3} < F_i; \\ XS_i, & \text{else,} \end{cases} \quad (15)$$

3. STRATEGIES APPLIED FOR VARIANTS OF THE CHEF-BASED OPTIMIZATION ALGORITHM

In this study, various improvement strategies were added to the algorithm to improve the performance of CBOA. These strategies are explained under three subheadings. In the first sub-heading, Fitness distance balance strategy is explained. The second sub-heading describes the chaotic maps added to CBOA. In the third sub-heading, the combination of the fitness distance balance strategy and chaotic maps is explained.

3.1. Fitness distance balance strategy

The fitness distance balance strategy was proposed to the literature by Kahraman et al. (2020). This strategy uses a point system to find the best potential solution in the population. This strategy consists of five basic stages as presented below (Bakır et al., 2023).

1. Stage: When $x_i (i = 1, 2, \dots, k)$ P is the population, $x_{ij} = [x_{i1}, x_{i2}, \dots, x_{in}]$, $f_i (i = 1, 2, \dots, k)$ is the fitness value. The vectors P and FV are given in Equation 16.

$$P \equiv \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{k1} & \cdots & x_{kn} \end{bmatrix}_{k \times n}, FV = \begin{bmatrix} f_1 \\ \vdots \\ f_k \end{bmatrix}_{k \times 1} \quad (16)$$

2. Stage: In this study, the distance is calculated according to Equation 17 using the Euclidean distance if the problem size is n and the number of individuals in the population is k .

$$\forall x_i \neq x_{best}, D_{x,i} = \sqrt{(x_{i[1]} - x_{best[1]})^2 + (x_{i[2]} - x_{best[2]})^2 + \cdots + (x_{i[n]} - x_{best[n]})^2} \quad (17)$$

3. Stage: The D_x distance vector is given as in Equation 18 below.

$$D_x \equiv \begin{bmatrix} d_1 \\ \vdots \\ d_k \end{bmatrix}_{k \times 1} \quad (18)$$

4. Stage: $normFV$ ve $normD_x$ are the normalized fitness value and distance value, respectively, and the FDB score of individual i . is calculated as in Equation 19 below.

$$\forall x_i, S_{x[i]} = w * normFV_{x[i]} + (1 - w) * normD_{x[i]} \quad (19)$$

5. Stage: Each individual's score vector is presented in Equation 20.

$$S_x \equiv \begin{bmatrix} s_1 \\ \vdots \\ s_k \end{bmatrix}_{k \times 1} \quad (20)$$

The improved CBOA with the fitness distance balance strategy is called F-CBOA. The FDB strategy added to the CBOA algorithm is presented in Equation 21 and the pseudo code of the proposed F-CBOA is given in Algorithm 1.

$$x_{S_{i,j}}^{S/S1} = x_{S_{F,j}} + r \cdot (CI_{k_{i,j}} - I \cdot x_{S_{F,j}}) \quad (21)$$

Algorithm 1: F-CBOA

```

1: Initialize: Set population size of the CBOA(N) and iterations (T)
2:   Randomly generate an initial population matrix X.
3:   Evaluate the given objective function to obtain the vector F.
4: for t = 1 to T do
5:   Sort the X according to objective function value by Eq. 4 and 5
6:   Start Phase1:
7:     for i=1 to NC do
8:       Calculate  $XS_i^{C/S1}$  using Eq. 6
9:       Update  $XS_i$  using Eq. 7
10:      Calculate  $XS_i^{C/S2}$  using Eq. 8
11:      Update  $XS_i$  using Eq. 9
12:     end for
13:   End Phase1:
14:   Start Phase2:
15:     for t = NC + 1 to N do
16:       Calculate  $XS_i^{S/S1}$  using Eq. 20 with Fitness Distance Balance
17:       Update  $XS_i$  using Eq. 11
18:       Calculate  $XS_i^{S/S2}$  using Eq. 12
19:       Update  $XS_i$  using Eq. 13
20:       Calculate  $XS_i^{S/S3}$  using Eq. 14
21:       Update  $XS_i$  using Eq. 15
22:     end for
23:   End Phase2:
24: end for
25: Return the best solution

```

3.2. Chaotic map strategies

Chaotic maps are sensitive to initial parameters and deal with nonlinear systems (Mirjalili & Gandomi, 2017). Chaotic systems can help avoid local minima and are therefore often used to improve the performance of population-based optimization algorithms (Turkoglu et al., 2025). Therefore, chaotic maps are frequently preferred in population-based methods. In this study, duffing map (Dhopavkar et al., 2022), iterative map (Wang et al., 2014) and sinusoidal map (Wang et al., 2014) are used presented in Table 1. These three different chaotic maps were applied to the CBOA algorithm and named C1-CBOA, C2-CBOA and C3-CBOA respectively. The pseudo code of the chaos-based CBOA variants is given in Algorithm 2.

Table 1. Three different chaotic map strategies and formulas

Algorithm	Chaotic map	Formula
C1-CBOA	Duffing	$x_{k+1} = y_k$ $y_{k+1} = -bx_k y_k + ay_k - y_k^3$
C2-CBOA	Iterative	$x_{k+1} = \sin\left(\frac{a\pi}{x_k}\right), a \in (0, 1)$
C3-CBOA	Sinusoidal	$x_{k+1} = ax_k^2 \sin(\pi x_k)$

Algorithm 2: C1-CBOA, C2-CBOA, C3-CBOA

```

1: Initialize: Set population size of the CBOA(N) and iterations (T)
2:     Generate an initial population matrix X with chaotic maps.
3:     Evaluate the given objective function to obtain the vector F.
4: for t = 1 to T do
5:     Sort the X according to objective function value by Eq. 4 and 5
6:     Start Phase1:
7:         for i=1 to  $N_C$  do
8:             Calculate  $XS_i^{C/S1}$  using Eq. 6
9:             Update  $XS_i$  using Eq. 7
10:            Calculate  $XS_i^{C/S2}$  using Eq. 8
11:            Update  $XS_i$  using Eq. 9
12:        end for
13:     End Phase1:
14:     Start Phase2:
15:         for t =  $N_C + 1$  to N do
16:             Calculate  $XS_i^{S/S1}$  using Eq. 10
17:             Update  $XS_i$  using Eq. 11
18:             Calculate  $XS_i^{S/S2}$  using Eq. 12
19:             Update  $XS_i$  using Eq. 13
20:             Calculate  $XS_i^{S/S3}$  using Eq. 14
21:             Update  $XS_i$  using Eq. 15
22:         end for
23:     End Phase2:
24: end for
25: Return the best solution

```

3.3. Combination of fitness distance balance and chaotic map strategies

In this section, three different approaches, FC1-CBOA, FC2-CBOA and FC3-CBOA, are proposed in combination with the fitness distance balance strategy and chaotic maps to further improve the performance of the CBOA algorithm. The pseudo code of the proposed approach is presented in Algorithm 3.

Algorithm 3: FC1-CBOA, FC2-CBOA, FC3-CBOA

```

1: Initialize: Set population size of the CBOA(N) and iterations (T)
2:     Generate an initial population matrix X with chaotic maps.
3:     Evaluate the given objective function to obtain the vector F.
4: for t = 1 to T do
5:     Sort the X according to objective function value by Eq. 4 and 5
6:     Start Phase1:
7:         for i=1 to  $N_C$  do
8:             Calculate  $XS_i^{C/S1}$  using Eq. 6
9:             Update  $XS_i$  using Eq. 7
10:            Calculate  $XS_i^{C/S2}$  using Eq. 8
11:            Update  $XS_i$  using Eq. 9
12:        end for
13:     End Phase1:
14:     Start Phase2:
15:         for t =  $N_C + 1$  to N do
16:             Calculate  $XS_i^{S/S1}$  using Eq. 20 with Fitness Distance Balance
17:             Update  $XS_i$  using Eq. 11
18:             Calculate  $XS_i^{S/S2}$  using Eq. 12
19:             Update  $XS_i$  using Eq. 13
20:             Calculate  $XS_i^{S/S3}$  using Eq. 14
21:             Update  $XS_i$  using Eq. 15
22:         end for
23:     End Phase2:
24: end for
25: Return the best solution

```

4. EXPERIMENTAL RESULTS

In this study, the performance of the proposed CBOA variants is first tested on benchmark functions and then its performance on 37-bar and 52-bar truss problems is evaluated. The number of iterations of the original CBOA and the proposed CBOA variants is taken as 250. The benchmark functions and bar truss problems considered in the study were 30 runtimes. According to these conditions, the results obtained with the variants of the proposed CBOA and the original CBOA are presented in tables under the relevant headings.

4.1. Application of the proposed methods to benchmark functions

The names of the 16 different benchmark test functions used in the study and their equations are given in Table 2 (Bai et al., 2021; Beşkirli, 2021).

The comparison of the best, mean, std values of the original CBOA and the best, mean, std values of the proposed CBOA variants are given in Table 3, Table 4, Table 5, Table 6, Table 7, Table 8, Table 9. The results of the original CBOA are compared one by one with the results of the variants of CBOA. Also, according to the results given in the tables, a rank graph between the algorithms was created as shown in Figure 1.

The graph in Figure 1 shows the performance rankings of the original CBOA and the proposed methods on the benchmark functions. When the graph is analyzed, it is seen that the methods compared with the original CBOA show a variable performance on the test functions. When the original CBOA is compared with the F-CBOA method, it is seen that F-CBOA has a better performance ranking. However, when the original CBOA is compared with the C1-CBOA method, it is seen that the original CBOA performs better, while when the original CBOA is compared with the C2-CBOA method, both algorithms perform equally well. When the original CBOA is compared with the C3-CBOA method, it is again observed that the original CBOA has a better success ranking. In addition, when the original CBOA is compared with FC1-CBOA, FC2-CBOA, FC3-CBOA combination algorithms, it is seen that the proposed FC2-CBOA and FC3-CBOA algorithms have a higher success ranking than the original CBOA.

4.2. Application of the proposed methods to 37-bar and 52-bar truss structures

In this section, 37-bar and 52-bar truss problems with frequency constraints, which are real-world engineering problems, are considered to evaluate the performance of the proposed methods.

4.2.1. 37-bar truss structures

One of the optimization problems considered in the study is a 37-bar truss structure with frequency constraints (Öztürk & Kahraman, 2025), the details of which are presented in Figure 2. This structure has a total of 19 design variables. In order to evaluate the effectiveness of the algorithms on the 37-bar truss structure, the best, mean, standard deviation (std.) values obtained by the original CBOA and the proposed CBOA variants are presented comparatively in Table 10.

Table 2. Benchmark functions

Fn.	Function Name	Search Range	Function
F1	Sphere	$[-100, 100]^D$	$F1(x) = \sum_{i=1}^N x_i^2$
F2	Elliptic	$[-100, 100]^D$	$F2(x) = \sum_{i=1}^n (10^6)^{(i-1)/(n-1)} x_i^2$
F3	SumSquares	$[-10, 10]^D$	$F3(x) = \sum_{i=1}^n i x_i^2$
F4	SumPower	$[-10, 10]^D$	$F4(x) = \sum_{i=1}^n x_i ^{(i+1)}$
F5	Schwefel 2.22	$[-10, 10]^D$	$F5(x) = \sum_{i=1}^n x_i + \prod_{i=1}^n x_i $
F6	Schwefel 2.21	$[-100, 100]^D$	$F6(x) = \max_i \{ x_i , 1 \leq i \leq n\}$
F7	QuarticWN	$[-1.28, 1.28]^D$	$F7(x) = \sum_{i=1}^n i x_i^4 + \text{random}[0,1]$
F8	Rosenbrock	$[-10, 10]^D$	$F8(x) = \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$
F9	Non-Continuous Rastrigin	$[-5.12, 5.12]^D$	$F9(x) = \sum_{i=1}^n [y_i^2 - 10 \cos(2\pi y_i) + 10] y_i = \begin{cases} x_i, & x_i < \frac{1}{2} \\ \frac{\text{round}(2x_i)}{2}, & x_i \geq \frac{1}{2} \end{cases}$
F10	Schwefel 2.26	$[-500, 500]^D$	$F10(x) = 418.98 * n - \sum_{i=1}^n x_i \sin(\sqrt{ x_i })$
F11	Ackley	$[-32, 32]^D$	$F11(x) = 20 + e \pm 20 \exp(-0.2 \sqrt{\frac{1}{N} \sum_{i=1}^N x_i^2}) - \exp(\frac{1}{N} \sum_{i=1}^N \cos(2\pi x_i))$
F12	Penalized 1	$[-50, 50]^D$	$F12(x) = \frac{\pi}{n} \left\{ 10 \sin^2(\pi y_1) + \sum_{i=1}^{n-1} (y_i - 1)^2 [1 + 10 \sin^2(\pi y_{i+1})] + (y_n - 1)^2 \right\} + \sum_{i=1}^n u(x_i, 10, 100, 4)$
F13	Penalized 2	$[-50, 50]^D$	$F13(x) = \frac{1}{10} \left\{ \sin^2(\pi x_1) + \sum_{i=1}^{n-1} (x_i - 1)^2 [1 + \sin^2(3\pi x_{i+1})] + (x_n - 1)^2 [1 + \sin^2(2\pi x_{i+1})] \right\} + \sum_{i=1}^n u(x_i, 5, 100, 4)$
F14	Alpine	$[-10, 10]^D$	$F14(x) = \sum_{i=1}^n x_i \sin(x_i) + 0.1 x_i $
F15	Levy	$[-10, 10]^D$	$F15(x) = \sum_{i=1}^{n-1} (x_i - 1)^2 [1 + \sin^2(3\pi x_{i+1})] + \sin^2(3\pi x_1) + x_n - 1 [1 + \sin^2(3\pi x_n)]$
F16	Schwefel 1.20	$[-100, 100]^D$	$F16(x) = \sum_{i=1}^n (\sum_{j=1}^i x_j)^2$

Table 3. Results of original CBOA and proposed F-CBOA

Fnc.	CBOA			F-CBOA		
	Best	Mean	Std.	Best	Mean	Std.
F1	1.5867E-187	4.9473E-170	0.000E+00	7.6504E-238	1.0757E-216	0.000E+00
F2	3.7559E-180	4.9385E-169	0.000E+00	2.1187E-245	1.0124E-221	0.000E+00
F3	9.2550E-183	7.0487E-172	0.000E+00	2.8681E-246	4.6185E-220	0.000E+00
F4	4.1035E-249	1.3305E-228	0.000E+00	0.000E+00	0.000E+00	0.000E+00
F5	2.8999E-94	2.1182E-86	1.0822E-85	3.8192E-122	5.3787E-109	2.1577E-108
F6	2.9144E-89	5.8790E-84	1.7759E-83	8.5125E-126	1.6638E-109	8.9575E-109
F7	3.3007E-07	3.3327E-04	2.8540E-04	2.8331E-05	3.6955E-04	2.5055E-04
F8	1.3103E-02	1.4967E+01	1.3918E+01	7.3961E-03	1.0237E+01	1.3416E+01
F9	0.000E+00	1.1405E+00	3.5723E+00	0.000E+00	0.000E+00	0.000E+00
F10	1.3832E+03	3.4978E+03	9.7996E+02	2.6269E+03	5.0098E+03	1.1986E+03
F11	0.000E+00	1.4803E-16	5.5388E-16	0.000E+00	0.000E+00	0.000E+00
F12	1.1651E-05	2.8967E-05	1.1089E-05	1.2016E-05	3.6117E-05	1.5097E-05
F13	1.8638E-04	6.6179E-04	3.0371E-04	3.4117E-04	1.4731E-03	3.4799E-03
F14	1.3844E-92	5.0568E-89	1.6277E-88	9.9461E-122	2.0300E-110	9.2100E-110
F15	1.5586E-04	6.9348E-02	1.9938E-01	1.7356E-04	1.1508E-01	1.6865E-01
F16	3.0926E-166	8.3527E-156	3.5277E-155	3.4584E-226	1.3868E-199	0.000E+00

Table 4. Results of original CBOA and proposed C1-CBOA

Fnc.	CBOA			C1-CBOA		
	Best	Mean	Std.	Best	Mean	Std.
F1	1.5867E-187	4.9473E-170	0.000E+00	7.0904E-180	1.9737E-168	0.000E+00
F2	3.7559E-180	4.9385E-169	0.000E+00	1.1857E-176	2.4599E-141	1.3247E-140
F3	9.2550E-183	7.0487E-172	0.000E+00	1.7677E-180	4.7124E-169	0.000E+00
F4	4.1035E-249	1.3305E-228	0.000E+00	5.3968E-253	3.5888E-208	0.000E+00
F5	2.8999E-94	2.1182E-86	1.0822E-85	1.9829E-91	9.2268E-84	2.4200E-83
F6	2.9144E-89	5.8790E-84	1.7759E-83	3.0588E-88	8.7167E-83	3.4520E-82
F7	3.3007E-07	3.3327E-04	2.8540E-04	5.5611E-05	3.5300E-04	2.0736E-04
F8	1.3103E-02	1.4967E+01	1.3918E+01	3.5078E-02	1.6817E+01	1.3658E+01
F9	0.000E+00	1.1405E+00	3.5723E+00	0.000E+00	2.2988E+00	4.8737E+00
F10	1.3832E+03	3.4978E+03	9.7996E+02	2.3605E+03	3.5427E+03	6.6630E+02
F11	0.000E+00	1.4803E-16	5.5388E-16	0.000E+00	0.000E+00	0.000E+00
F12	1.1651E-05	2.8967E-05	1.1089E-05	7.5952E-06	3.4825E-03	1.8614E-02
F13	1.8638E-04	6.6179E-04	3.0371E-04	2.7630E-04	5.7482E-04	2.2679E-04
F14	1.3844E-92	5.0568E-89	1.6277E-88	1.4818E-93	7.4963E-86	3.1492E-85
F15	1.5586E-04	6.9348E-02	1.9938E-01	2.8324E-04	3.3956E-02	5.4541E-02
F16	3.0926E-166	8.3527E-156	3.5277E-155	3.0814E-170	5.8351E-151	3.1112E-150

Table 5. Results of original CBOA and proposed C2-CBOA

Fnc.	CBOA			C2-CBOA		
	Best	Mean	Std.	Best	Mean	Std.
F1	1.5867E-187	4.9473E-170	0.000E+00	6.8254E-183	6.5051E-171	0.000E+00
F2	3.7559E-180	4.9385E-169	0.000E+00	6.5037E-176	8.4632E-159	4.5130E-158
F3	9.2550E-183	7.0487E-172	0.000E+00	4.5305E-183	1.8571E-171	0.000E+00
F4	4.1035E-249	1.3305E-228	0.000E+00	7.3651E-251	6.3099E-224	0.000E+00
F5	2.8999E-94	2.1182E-86	1.0822E-85	8.1506E-90	1.9933E+00	1.0734E+01
F6	2.9144E-89	5.8790E-84	1.7759E-83	6.7089E-89	7.2424E-84	2.7756E-83
F7	3.3007E-07	3.3327E-04	2.8540E-04	3.3664E-05	3.4324E-04	2.4912E-04
F8	1.3103E-02	1.4967E+01	1.3918E+01	2.5863E-02	2.3347E+01	1.0399E+01
F9	0.000E+00	1.1405E+00	3.5723E+00	0.000E+00	4.4761E+00	6.8436E+00
F10	1.3832E+03	3.4978E+03	9.7996E+02	1.9391E+03	3.4671E+03	7.3962E+02
F11	0.000E+00	1.4803E-16	5.5388E-16	0.000E+00	7.4015E-17	3.9858E-16
F12	1.1651E-05	2.8967E-05	1.1089E-05	1.1085E-05	2.6678E-05	1.0540E-05
F13	1.8638E-04	6.6179E-04	3.0371E-04	8.0753E-05	5.6799E-04	2.7924E-04
F14	1.3844E-92	5.0568E-89	1.6277E-88	2.0916E-93	5.3044E-86	2.8111E-85
F15	1.5586E-04	6.9348E-02	1.9938E-01	8.8292E-05	3.3667E-02	6.6173E-02
F16	3.0926E-166	8.3527E-156	3.5277E-155	8.3507E-173	7.4692E-156	3.9926E-155

Table 6. Results of original CBOA and proposed C3-CBOA

Fnc.	CBOA			C3-CBOA		
	Best	Mean	Std.	Best	Mean	Std.
F1	1.5867E-187	4.9473E-170	0.000E+00	1.1373E-179	7.6474E-171	0.000E+00
F2	3.7559E-180	4.9385E-169	0.000E+00	3.5786E-177	7.4605E-166	0.000E+00
F3	9.2550E-183	7.0487E-172	0.000E+00	3.7614E-183	2.5247E-169	0.000E+00
F4	4.1035E-249	1.3305E-228	0.000E+00	8.3708E-248	5.3006E-228	0.000E+00
F5	2.8999E-94	2.1182E-86	1.0822E-85	3.9010E-91	4.0334E-87	1.2177E-86
F6	2.9144E-89	5.8790E-84	1.7759E-83	1.3990E-88	8.5179E-83	3.8513E-82
F7	3.3007E-07	3.3327E-04	2.8540E-04	1.0677E-04	4.5000E-04	3.1200E-04
F8	1.3103E-02	1.4967E+01	1.3918E+01	2.4128E-02	1.6830E+01	1.3667E+01
F9	0.000E+00	1.1405E+00	3.5723E+00	0.000E+00	0.000E+00	0.000E+00
F10	1.3832E+03	3.4978E+03	9.7996E+02	2.0216E+03	3.7003E+03	9.7581E+02
F11	0.000E+00	1.4803E-16	5.5388E-16	0.000E+00	0.000E+00	0.000E+00
F12	1.1651E-05	2.8967E-05	1.1089E-05	6.6017E-06	2.6277E-05	1.0415E-05
F13	1.8638E-04	6.6179E-04	3.0371E-04	2.7074E-04	1.2558E-03	3.4701E-03
F14	1.3844E-92	5.0568E-89	1.6277E-88	4.1695E-94	3.3522E-89	7.8233E-89
F15	1.5586E-04	6.9348E-02	1.9938E-01	2.9917E-04	6.3434E-02	1.0196E-01
F16	3.0926E-166	8.3527E-156	3.5277E-155	2.1340E-168	8.4216E-154	4.2894E-153

Table 7. Results of original CBOA and proposed FCI-CBOA

Fnc.	CBOA			FCI-CBOA		
	Best	Mean	Std.	Best	Mean	Std.
F1	1.5867E-187	4.9473E-170	0.000E+00	1.2520E-238	1.7924E-213	0.000E+00
F2	3.7559E-180	4.9385E-169	0.000E+00	8.1321E-250	1.1973E-216	0.000E+00
F3	9.2550E-183	7.0487E-172	0.000E+00	1.9576E-247	1.4412E-216	0.000E+00
F4	4.1035E-249	1.3305E-228	0.000E+00	0.000E+00	1.1979E-316	0.000E+00
F5	2.8999E-94	2.1182E-86	1.0822E-85	3.1489E-121	1.5183E-101	8.1762E-101
F6	2.9144E-89	5.8790E-84	1.7759E-83	4.2720E-125	1.1834E-111	4.6352E-111
F7	3.3007E-07	3.3327E-04	2.8540E-04	2.7930E-05	5.3348E-04	2.8952E-04
F8	1.3103E-02	1.4967E+01	1.3918E+01	2.4003E-03	1.5819E+01	1.3808E+01
F9	0.000E+00	1.1405E+00	3.5723E+00	0.000E+00	0.000E+00	0.000E+00
F10	1.3832E+03	3.4978E+03	9.7996E+02	3.8382E+03	5.4913E+03	1.0920E+03
F11	0.000E+00	1.4803E-16	5.5388E-16	0.000E+00	0.000E+00	0.000E+00
F12	1.1651E-05	2.8967E-05	1.1089E-05	1.6574E-05	4.1278E-05	1.5501E-05
F13	1.8638E-04	6.6179E-04	3.0371E-04	2.5477E-04	8.2295E-04	3.1033E-04
F14	1.3844E-92	5.0568E-89	1.6277E-88	1.7118E-119	1.4935E-107	8.0154E-107
F15	1.5586E-04	6.9348E-02	1.9938E-01	1.8563E-04	4.7129E-02	1.2025E-01
F16	3.0926E-166	8.3527E-156	3.5277E-155	4.5025E-230	4.5393E-198	0.000E+00

Table 8. Results of original CBOA and proposed FC2-CBOA

Fnc.	CBOA			FC2-CBOA		
	Best	Mean	Std.	Best	Mean	Std.
F1	1.5867E-187	4.9473E-170	0.000E+00	4.2149E-237	1.4247E-216	0.000E+00
F2	3.7559E-180	4.9385E-169	0.000E+00	4.3726E-253	4.4781E-209	0.000E+00
F3	9.2550E-183	7.0487E-172	0.000E+00	5.7958E-239	1.4535E-212	0.000E+00
F4	4.1035E-249	1.3305E-228	0.000E+00	0.000E+00	0.000E+00	0.000E+00
F5	2.8999E-94	2.1182E-86	1.0822E-85	2.8571E-117	1.8823E+00	1.0136E+01
F6	2.9144E-89	5.8790E-84	1.7759E-83	8.9505E-127	2.1396E-107	1.1136E-106
F7	3.3007E-07	3.3327E-04	2.8540E-04	4.0916E-05	2.7858E-04	1.7081E-04
F8	1.3103E-02	1.4967E+01	1.3918E+01	4.0228E-03	1.2098E+01	1.3790E+01
F9	0.000E+00	1.1405E+00	3.5723E+00	0.000E+00	1.000E+00	5.3852E+00
F10	1.3832E+03	3.4978E+03	9.7996E+02	3.5460E+03	5.4033E+03	1.1114E+03
F11	0.000E+00	1.4803E-16	5.5388E-16	0.000E+00	0.000E+00	0.000E+00
F12	1.1651E-05	2.8967E-05	1.1089E-05	1.6033E-05	3.4914E-03	1.8612E-02
F13	1.8638E-04	6.6179E-04	3.0371E-04	3.7583E-04	2.0549E-03	4.7325E-03
F14	1.3844E-92	5.0568E-89	1.6277E-88	1.3249E-125	3.1826E-107	1.7090E-106
F15	1.5586E-04	6.9348E-02	1.9938E-01	5.9513E-05	4.8787E-02	8.6617E-02
F16	3.0926E-166	8.3527E-156	3.5277E-155	1.2857E-225	3.8665E-206	0.000E+00

Table 9. Results of original CBOA and proposed FC3-CBOA

Fnc.	CBOA			FC3-CBOA		
	Best	Mean	Std.	Best	Mean	Std.
F1	1.5867E-187	4.9473E-170	0.000E+00	7.0016E-240	6.2062E-217	0.000E+00
F2	3.7559E-180	4.9385E-169	0.000E+00	8.2057E-256	1.5177E-221	0.000E+00
F3	9.2550E-183	7.0487E-172	0.000E+00	9.3712E-248	1.3014E-222	0.000E+00
F4	4.1035E-249	1.3305E-228	0.000E+00	0.000E+00	2.4227E-314	0.000E+00
F5	2.8999E-94	2.1182E-86	1.0822E-85	8.8458E-122	1.5093E-110	4.3088E-110
F6	2.9144E-89	5.8790E-84	1.7759E-83	4.1242E-125	5.5322E-113	2.7803E-112
F7	3.3007E-07	3.3327E-04	2.8540E-04	2.9603E-05	3.1024E-04	2.1259E-04
F8	1.3103E-02	1.4967E+01	1.3918E+01	2.9015E-03	1.0260E+01	1.3423E+01
F9	0.000E+00	1.1405E+00	3.5723E+00	0.000E+00	0.000E+00	0.0000E+00
F10	1.3832E+03	3.4978E+03	9.7996E+02	3.0314E+03	5.3814E+03	1.2002E+03
F11	0.000E+00	1.4803E-16	5.5388E-16	0.000E+00	0.000E+00	0.000E+00
F12	1.1651E-05	2.8967E-05	1.1089E-05	8.6835E-06	3.4923E-03	1.8624E-02
F13	1.8638E-04	6.6179E-04	3.0371E-04	3.8203E-04	8.4733E-04	3.5427E-04
F14	1.3844E-92	5.0568E-89	1.6277E-88	3.5821E-122	5.2003E-108	2.4297E-107
F15	1.5586E-04	6.9348E-02	1.9938E-01	2.5303E-04	8.1376E-02	1.0959E-01
F16	3.0926E-166	8.3527E-156	3.5277E-155	1.2284E-236	1.8165E-208	0.000E+00

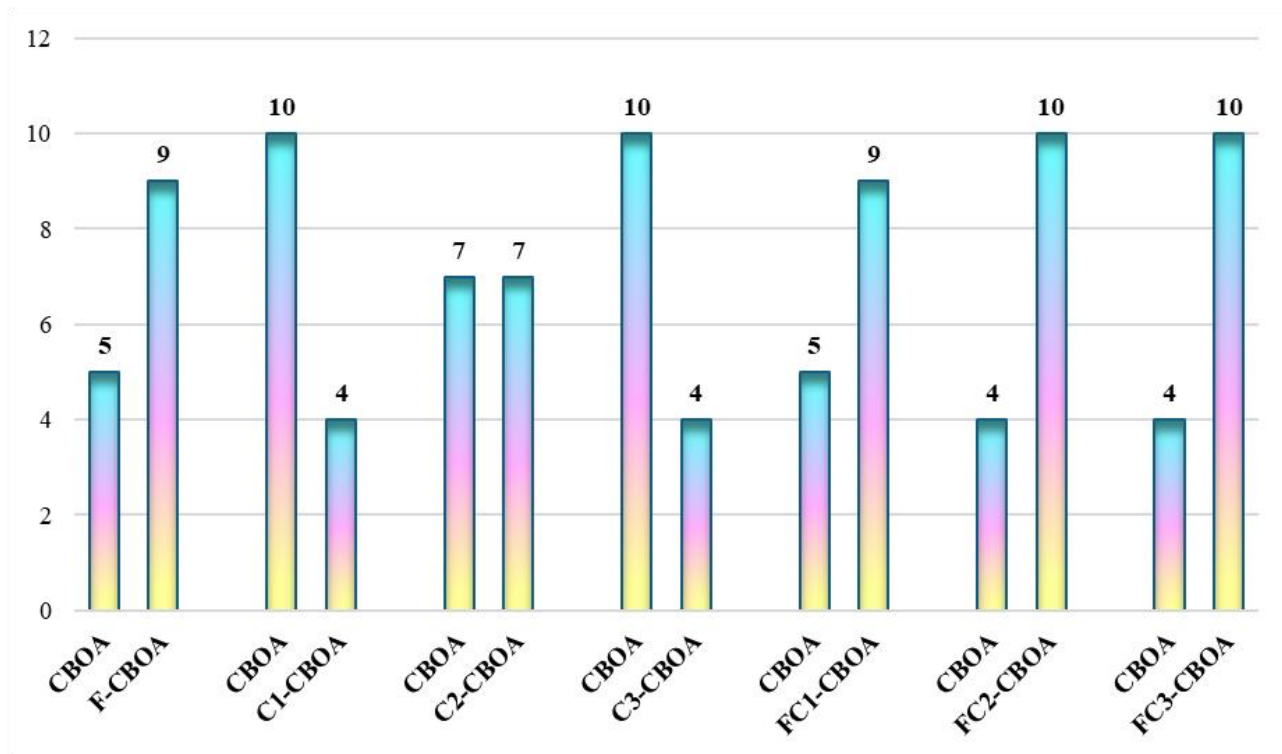


Figure 1. Success ranking of the algorithms on benchmark functions

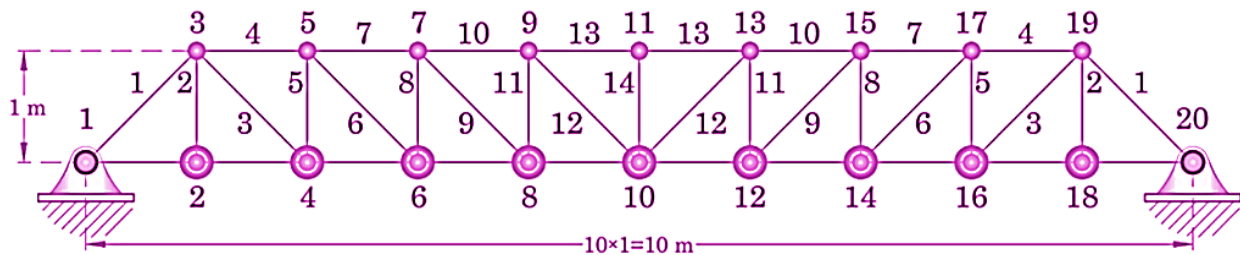


Figure 2. 37-bar truss structure (Kaveh & Mahjoubi, 2019)

When Table 10 is analyzed in terms of the best values obtained, it is seen that the proposed FC3-CBOA variant produces a better result than the original CBOA. The grouping of the design variables and the values obtained by the algorithms are presented in Table 11.

4.2.2. 52-bar truss structures

Another optimization problem considered in the study is a 52-bar truss structure with frequency constraints (Öztürk & Kahraman, 2025) as detailed in Figure 3. This structure has a total of 13 design variables. To evaluate the effectiveness of the algorithms on the 52-bar lattice structure, the best, mean, std. values obtained by the original CBOA and the proposed CBOA variants are presented in Table 12.

When Table 12 is analyzed in terms of the best values obtained, it is seen that the proposed C3-CBOA variant produces a better result than the original CBOA. The grouping of the design variables and the values obtained by the algorithms are presented in Table 13. Figure 4 shows the comparison of the best values of the proposed CBOA variants with the original CBOA for 37-bar and 52-bar truss structures. When the figure is analyzed, it

is seen that the best value for the 37-bar truss structure is obtained by the F3-CBOA variant, while the best value for the 52-bar truss structure is obtained by the C3-CBOA variant.

Table 10. Results of original CBOA and proposed CBOA variants for 37-bar truss structure

37-bar truss			
	Best	Mean	Std.
CBOA	3.7950E+02	3.8946E+02	4.9427E+00
F-CBOA	3.7735E+02	3.9333E+02	7.2335E+00
C1-CBOA	3.7291E+02	3.9117E+02	6.2613E+00
C2-CBOA	3.8613E+02	3.9286E+02	4.0322E+00
C3-CBOA	3.8164E+02	3.9028E+02	5.2331E+00
FC1-CBOA	3.8380E+02	3.9932E+02	7.8299E+00
FC2-CBOA	3.8749E+02	4.0123E+02	6.6931E+00
FC3-CBOA	3.7060E+02	3.9829E+02	7.9865E+00

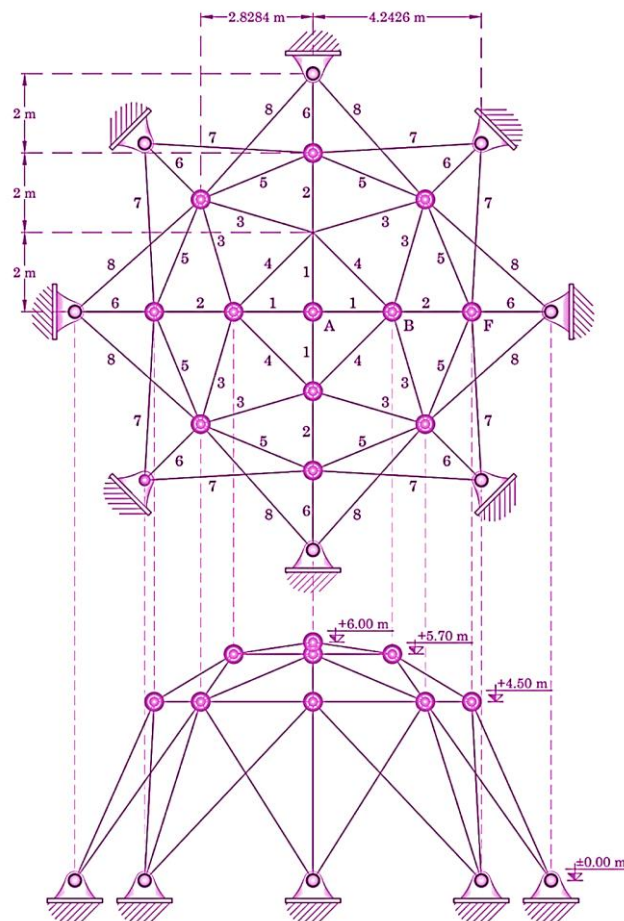


Figure 3. 52-bar truss structure (Kaveh & Mahjoubi, 2019)

Table 11. Results of the algorithms according to the grouped design variables

37-Bar Truss									
Groups	Element Group	Areas (cm ²)							
		CBOA	F-CBOA	C1-CBOA	C2-CBOA	C3-CBOA	FC1-CBOA	FC2-CBOA	FC3-CBOA
1	Y ₃ , Y ₁₉	0.7267	1.0044	0.7462	0.7428	0.6946	0.9207	0.7508	0.7273
2	Y ₅ , Y ₁₇	1.0475	1.1789	1.0779	1.0349	1.0666	1.2232	1.1902	1.1511
3	Y ₇ , Y ₁₅	1.3255	1.3301	1.3199	1.3180	1.2379	1.3749	1.3766	1.4480
4	Y ₉ , Y ₁₃	1.5144	1.4329	1.3956	1.4027	1.3691	1.4671	1.4801	1.6161
5	Y ₁₁	1.5762	1.4701	1.4052	1.4215	1.4440	1.4761	1.5000	1.6504
6	1–27	2.8890	2.9156	2.5734	3.0803	3.5238	3.0029	2.8638	2.6414
7	2–26	1.2420	3.5195	1.8876	2.8732	3.6117	3.2491	3.5697	2.2767
8	3–24	3.0044	2.4569	2.5352	3.0345	3.3928	3.0730	2.1645	1.2326
9	4–25	4.2594	3.2962	3.7985	3.1248	3.4259	2.2961	2.8660	2.8704
10	5–23	2.4697	2.2222	1.4959	2.0887	1.7522	3.0260	3.2832	1.0066
11	6–21	1.7674	2.3403	1.8722	2.2548	2.0133	2.8446	4.0986	2.0170
12	7–22	3.4931	2.6690	2.9777	3.1762	3.3955	2.7518	2.5506	2.3355
13	8–20	2.6885	1.8543	2.0791	2.7046	3.3346	2.6668	2.5793	2.5757
14	9–18	1.2856	1.4334	1.8462	3.0770	1.1559	2.1076	2.4470	2.4746
15	10–19	2.8132	4.2369	3.6442	4.0001	3.9096	2.7084	3.9198	2.8620
16	11–17	2.9817	1.4559	1.3142	1.6238	1.9856	1.7962	1.0040	1.7644
17	12–15	2.6678	1.8291	1.4765	3.4702	2.4937	1.3949	2.3830	1.6595
18	13–16	2.6322	3.0505	3.7930	3.7311	3.2573	4.6259	3.2636	3.2535
19	14	1.6276	1.3788	3.1458	1.6152	2.3171	2.2930	3.6779	1.3091
Mass (kg)		379.50	377.35	372.91	386.13	381.64	383.80	387.49	370.60

Table 12. Results of original CBOA and proposed CBOA variants for 52-bar truss structure

52-bar truss			
	Best	Mean	Std.
CBOA	2.3845E+02	2.6777E+02	2.1214E+01
F-CBOA	2.2629E+02	2.9764E+02	4.8092E+01
C1-CBOA	2.3579E+02	2.7716E+02	2.1372E+01
C2-CBOA	2.3588E+02	2.8471E+02	2.7742E+01
C3-CBOA	2.2153E+02	2.7540E+02	3.4255E+01
FC1-CBOA	2.4315E+02	3.0878E+02	5.4580E+01
FC2-CBOA	2.3345E+02	3.3966E+02	6.8234E+01
FC3-CBOA	2.4065E+02	3.0027E+02	4.4956E+01

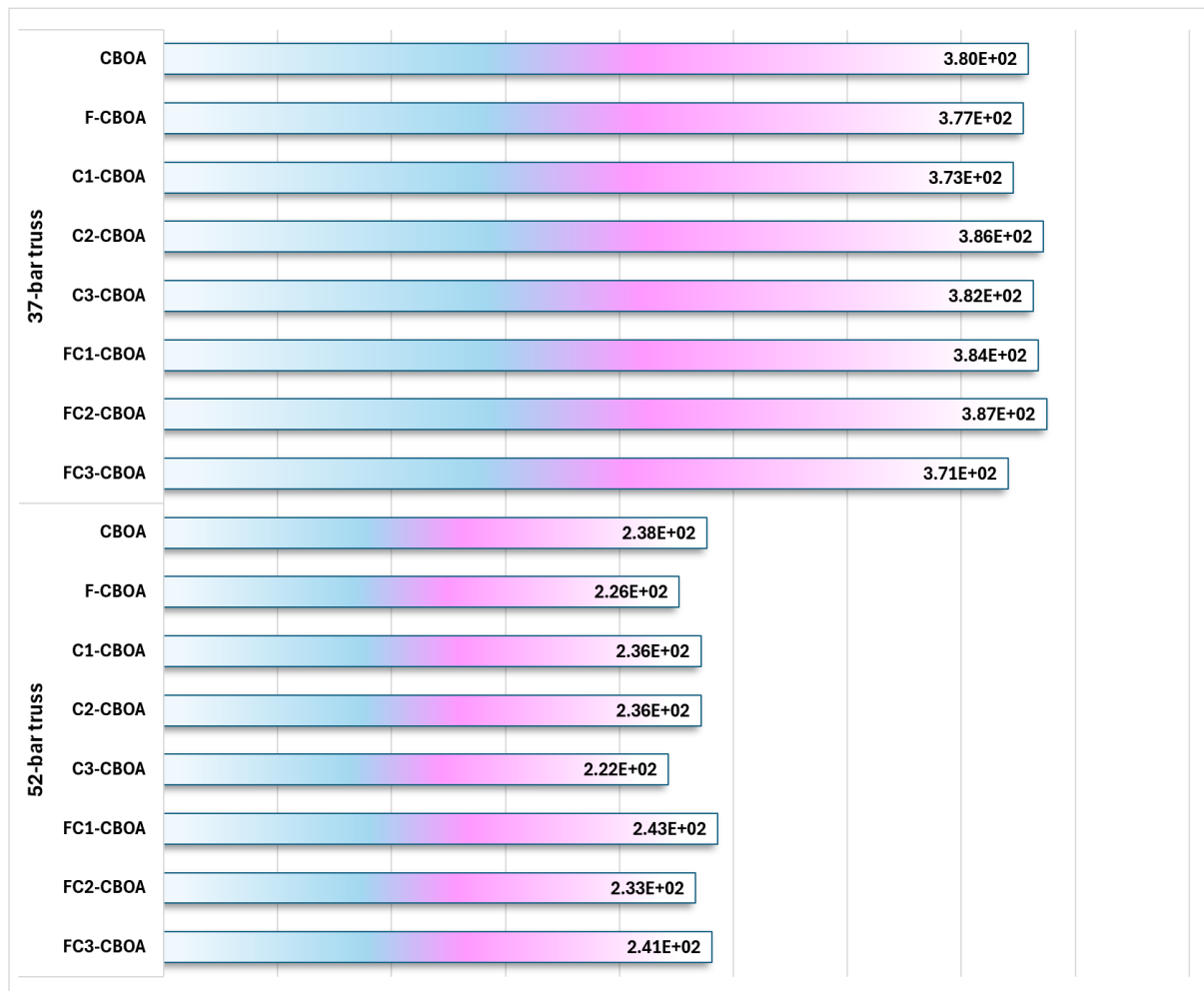
**Figure 4.** Ranking of best values for 37-bar and 52-bar truss structures

Table 13. Results of the algorithms according to the grouped design variables

52-Bar Truss									
Groups	Element Group	Areas (cm ²)							
		CBOA	F-CBOA	C1-CBOA	C2-CBOA	C3-CBOA	FC1-CBOA	FC2-CBOA	FC3-CBOA
1	Z _A	4.4145	4.7051	4.0000	4.2358	4.0000	4.0000	4.4820	4.0192
2	X _B	1.4031	1.5554	1.6453	1.0932	1.7827	1.9377	1.1868	1.7807
3	Z _B	3.7000	3.7579	3.8152	3.7473	3.7993	3.7472	3.7908	3.8365
4	X _F	3.1246	3.7467	3.7149	2.8600	3.8071	3.5574	3.0916	3.5422
5	Z _F	2.7652	2.5000	2.6189	2.8672	2.5000	2.6781	2.8157	2.7336
6	1–4	1.0000	1.4285	1.0000	1.0188	1.3863	1.0000	1.2207	1.0000
7	5–8	1.5746	2.4670	1.5030	2.3344	1.3006	1.4505	2.0581	2.6062
8	9–16	1.3901	2.0931	1.9726	1.5614	1.7855	2.0143	1.7702	2.4286
9	17–20	1.2416	1.6589	2.4156	1.9035	1.4057	1.6336	2.4155	1.4155
10	21–28	2.1040	1.9321	1.9937	1.8302	1.2389	1.3876	1.8393	1.3656
11	29–36	1.0245	1.0094	1.4014	1.0000	1.0000	1.1825	1.0000	1.0000
12	37–44	2.5847	1.2938	1.1665	2.1900	1.7038	3.0784	2.3040	1.5464
13	45–52	1.6603	1.6424	1.9847	1.8488	2.0197	1.0000	1.4029	2.1682
Mass (kg)		238.45	226.29	235.79	235.88	221.53	243.15	233.45	240.65

5. CONCLUSION

In this study, seven different variants of CBOA improved with chaos-based fitness distance balance are proposed to improve the performance of CBOA. The performance of the proposed CBOA variants on different problems is evaluated with three different experimental studies. In the improvement process, fitness distance balance is integrated into the algorithm as the first variant of CBOA. Then, duffing, iterative and sinusoidal chaotic maps were added to the CBOA algorithm as three different variants. Finally, a combination of both conformity distance balance and three different chaotic maps is proposed, and the performance of the proposed CBOA variants is investigated comparatively with three different experimental studies. In the first experimental study, the proposed CBOA variants were tested on 16 different benchmark functions. According to the results obtained, it is observed that especially the FC3-CBOA variant performs better than the original CBOA. In the second experimental study, the 37-bar truss problem with frequency constraints was considered and the best results were obtained by the FC3-CBOA variant compared to the original CBOA. In the third and

final experimental study, the 52-bar truss problem was evaluated, and the C3-CBOA variant was found to be prominent in this problem. The obtained analysis results show that each proposed CBOA variant is able to produce effective solutions depending on the problem type. In future studies, the effectiveness of the proposed CBOA variants on multi-objective optimization problems can be tested.

CONFLICT OF INTEREST

The author declares no conflict of interest.

REFERENCES

- Abdel-Basset, M., Mohamed, R., & Abouhawwash, M. (2025). Fungal growth optimizer: A novel nature-inspired metaheuristic algorithm for stochastic optimization. *Computer Methods in Applied Mechanics and Engineering*, 437, 117825. <https://doi.org/10.1016/j.cma.2025.117825>
- Abdollahzadeh, B., Soleimani Gharehchopogh, F., & Mirjalili, S. (2021). Artificial gorilla troops optimizer: A new nature-inspired metaheuristic algorithm for global optimization problems. *International Journal of Intelligent Systems*, 36(10), 5887-5958. <https://doi.org/10.1002/int.22535>
- Aribowo, W., Abualigah, L., Oliva, D., Mzili, T., Sabo, A., & A. Shehadeh, H. (2024). Frilled Lizard Optimization to optimize parameters Proportional Integral Derivative of DC Motor. *Vokasi Unesa Bulletin of Engineering, Technology and Applied Science*, 1(1), 14-21. <https://doi.org/10.26740/vubeta.v1i1.33973>
- Aribowo, W., Rahmadian, R., & Widyartono, M. (2023). Adaptive tuning of PID using chef-based optimization algorithm for improving automatic voltage regulator. *Indonesian Journal of Electrical Engineering and Computer Science*, 32(3), 1215-1223. <https://doi.org/10.11591/ijeecs.v32.i3.pp1215-1223>
- Bai, B., Zhang, J., Wu, X., wei Zhu, G., & Li, X. (2021). Reliability prediction-based improved dynamic weight particle swarm optimization and back propagation neural network in engineering systems. *Expert Systems with Applications*, 177, 114952. <https://doi.org/10.1016/j.eswa.2021.114952>
- Bakır, H., Duman, S., Guvenc, U., & Kahraman, H. T. (2023). Improved adaptive gaining-sharing knowledge algorithm with FDB-based guiding mechanism for optimization of optimal reactive power flow problem. *Electrical Engineering*, 105(5), 3121-3160. <https://doi.org/10.1007/s00202-023-01803-9>
- Balasubramaniam, S., Kadry, S., Dhanaraj, R. K., Satheesh Kumar, K., & Manthiramoorthy, C. (2024). Res-Net based blood vessel segmentation and cardio vascular disease prediction using chronological chef-based optimization algorithm based deep residual network from retinal fundus images. *Multimedia Tools and Applications*, 83(40), 87929-87958. <https://doi.org/10.1007/s11042-024-18810-y>
- Beşkirli, M. (2021). Solving continuous optimization problems using the tree seed algorithm developed with the roulette wheel strategy. *Expert Systems with Applications*, 170, 114579. <https://doi.org/10.1016/j.eswa.2021.114579>
- Dehghani, M., Montazeri, Z., Trojovská, E., & Trojovský, P. (2023). Coati Optimization Algorithm: A new bio-inspired metaheuristic algorithm for solving optimization problems. *Knowledge-Based Systems*,

259, 110011. <https://doi.org/10.1016/j.knosys.2022.110011>

- Demirbas, M., Dosoglu, M. K., & Duman, S. (2025). Enhanced Coati Optimization Algorithm for Static and Dynamic Transmission Network Expansion Planning Problems. *IEEE Access*, 13, 35068-35100. <https://doi.org/10.1109/ACCESS.2025.3544523>
- Dhopavkar, T. A., Nayak, S. K., & Roy, S. (2022). IETD: a novel image encryption technique using Tinkerbell map and Duffing map for IoT applications. *Multimedia Tools and Applications*, 81(30), 43189-43228. <https://doi.org/10.1007/s11042-022-13162-x>
- Fu, Y., Liu, D., Chen, J., & He, L. (2024). Secretary bird optimization algorithm: a new metaheuristic for solving global optimization problems. *Artificial Intelligence Review*, 57(5), 123. <https://doi.org/10.1007/s10462-024-10729-y>
- Ghasemi, M., Zare, M., Zahedi, A., Akbari, M.-A., Mirjalili, S., & Abualigah, L. (2024). Geyser Inspired Algorithm: A New Geological-inspired Meta-heuristic for Real-parameter and Constrained Engineering Optimization. *Journal of Bionic Engineering*, 21(1), 374-408. <https://doi.org/10.1007/s42235-023-00437-8>
- Hashim, F. A., Houssein, E. H., Hussain, K., Mabrouk, M. S., & Al-Atabany, W. (2022). Honey Badger Algorithm: New metaheuristic algorithm for solving optimization problems. *Mathematics and Computers in Simulation*, 192, 84-110. <https://doi.org/10.1016/j.matcom.2021.08.013>
- Huang, Q., Ding, H., & Effatparvar, M. (2024). Breast cancer diagnosis based on hybrid SqueezeNet and improved chef-based optimizer. *Expert Systems with Applications*, 237, 121470. <https://doi.org/10.1016/j.eswa.2023.121470>
- Joni, K. (2024). Parameter Estimation Of Photovoltaic based on Chaotic Elite Mountain Gazelle Optimizer. *Vokasi Unesa Bulletin of Engineering, Technology and Applied Science*, 1(1), 30-37. <https://doi.org/10.26740/vubeta.v1i1.34073>
- Kahraman, H. T., Aras, S., & Gedikli, E. (2020). Fitness-distance balance (FDB): A new selection method for meta-heuristic search algorithms. *Knowledge-Based Systems*, 190, 105169. <https://doi.org/10.1016/j.knosys.2019.105169>
- Kaveh, A., & Mahjoubi, S. (2019). Hypotrochoid spiral optimization approach for sizing and layout optimization of truss structures with multiple frequency constraints. *Engineering with Computers*, 35(4), 1443-1462. <https://doi.org/10.1007/s00366-018-0675-6>
- Kiran, M. S. (2015). TSA: Tree-seed algorithm for continuous optimization. *Expert Systems with Applications*, 42(19), 6686-6698. <https://doi.org/10.1016/j.eswa.2015.04.055>
- Kiran, M. S., & Beskirli, M. (2024). A New Approach Based on Collective Intelligence to Solve Traveling Salesman Problems. *Biomimetics*, 9(2), 118. <https://doi.org/10.3390/biomimetics9020118>
- Kutlu Onay, F. (2023). A novel improved chef-based optimization algorithm with Gaussian random walk-based diffusion process for global optimization and engineering problems. *Mathematics and Computers in Simulation*, 212, 195-223. <https://doi.org/10.1016/j.matcom.2023.04.027>
- Liu, Q., Sun, K., Tang, X., & Huo, J. (2024). Camera calibration based on lightweight fan-shaped target

- detection and fitness-distance-balance chaotic marine predators algorithm. *Optics & Laser Technology*, 176, 110883. <https://doi.org/10.1016/j.optlastec.2024.110883>
- Mirjalili, S., & Gandomi, A. H. (2017). Chaotic gravitational constants for the gravitational search algorithm. *Applied Soft Computing*, 53, 407-419. <https://doi.org/10.1016/j.asoc.2017.01.008>
- Mirjalili, S., & Lewis, A. (2016). The Whale Optimization Algorithm. *Advances in Engineering Software*, 95, 51-67. <https://doi.org/10.1016/j.advengsoft.2016.01.008>
- Öztürk, H. T., & Kahraman, H. T. (2025). Metaheuristic search algorithms in frequency constrained truss problems: Four improved evolutionary algorithms, optimal solutions and stability analysis. *Applied Soft Computing*, 171, 112854. <https://doi.org/10.1016/j.asoc.2025.112854>
- Prapanca, A., Nasreddine, B., & Imed, M. (2025). Modified FATA Morgana Algorithm Based on Levy Flight. *Vokasi Unesa Bulletin of Engineering, Technology and Applied Science*, 2(1), 1-11. <https://doi.org/10.26740/vubeta.v2i1.37066>
- Rajesh Kumar, T., Enireddy, V., Kalai Selvi, K., Shahid, M., Vijendra Babu, D., & Sudha, I. (2024). Fractional chef based optimization algorithm trained deep learning for cardiovascular risk prediction using retinal fundus images. *Biomedical Signal Processing and Control*, 94, 106269. <https://doi.org/10.1016/j.bspc.2024.106269>
- Trojovská, E., & Dehghani, M. (2022). A new human-based metaheuristic optimization method based on mimicking cooking training. *Scientific Reports*, 12(1), 14861. <https://doi.org/10.1038/s41598-022-19313-2>
- Trojovská, E., Dehghani, M., & Trojovský, P. (2022). Zebra Optimization Algorithm: A New Bio-Inspired Optimization Algorithm for Solving Optimization Algorithm. *IEEE Access*, 10, 49445-49473. <https://doi.org/10.1109/ACCESS.2022.3172789>
- Trojovský, P. (2023). A new human-based metaheuristic algorithm for solving optimization problems based on preschool education. *Scientific Reports*, 13(1), 21472. <https://doi.org/10.1038/s41598-023-48462-1>
- Trojovský, P., & Dehghani, M. (2022). Pelican Optimization Algorithm: A Novel Nature-Inspired Algorithm for Engineering Applications. *Sensors*, 22(3), 855. <https://doi.org/10.3390/s22030855>
- Turkoglu, B., Uymaz, S. A., & Kaya, E. (2025). Chaotic Artificial Algae Algorithm for Solving Global Optimization With Real-World Space Trajectory Design Problems. *Arabian Journal for Science and Engineering*, 50(2), 1279-1306. <https://doi.org/10.1007/s13369-024-09222-z>
- Wang, G.-G., Guo, L., Gandomi, A. H., Hao, G.-S., & Wang, H. (2014). Chaotic Krill Herd algorithm. *Information Sciences*, 274, 17-34. <https://doi.org/10.1016/j.ins.2014.02.123>
- Zhang, W., Zhao, J., Liu, H., & Tu, L. (2024). Cleaner fish optimization algorithm: a new bio-inspired metaheuristic optimization algorithm. *The Journal of Supercomputing*, 80(12), 17338-17376. <https://doi.org/10.1007/s11227-024-06105-w>
- Zhao, W., Wang, L., & Zhang, Z. (2019). Atom search optimization and its application to solve a hydrogeologic parameter estimation problem. *Knowledge-Based Systems*, 163, 283-304. <https://doi.org/10.1016/j.knosys.2018.08.030>



Gazi University

Journal of Science

PART A: ENGINEERING AND INNOVATION

<http://dergipark.org.tr/gujsa>

Proposing a Five-Phase Framework Based on ISO 23247-1 for Digital Twins in Construction

Murat AYDIN^{1*} ¹ Ankara University, Ankara, Türkiye

Keywords	Abstract
Digital Twin	The adoption of digital twin technology represents a significant leap forward in the construction industry, driving sustainable and efficient project workflows. Despite its transformative potential, challenges such as data integration, interoperability issues, and the absence of structured frameworks hinder broader adoption. To address these barriers, this study proposes a five-phase framework inspired by ISO 23247-1 principles, offering standardized guidelines to ensure seamless data flow, interoperability, and data-driven decision-making in digital twin applications. To evaluate its practicality, this framework was implemented in Villa EcoSmart—a hypothetical testbed simulating a sustainable residential construction project. The five phases encompass free (foundation and requirements establishment), acquire (data collection), analyze (data processing), utilize (model utilization), and update (continuous refinement). Findings demonstrate improvements in energy efficiency, material usage, and workflow optimization, underscoring the framework's value in achieving technological innovation and environmental responsibility. Additionally, this study critically assesses the scalability and real-world applicability of digital twin technologies. By bridging the gap between theoretical knowledge and industry practices, the five-phase framework advances sustainable construction methods, aligning technological solutions with ISO standards. These insights aim to guide future implementations and promote the broader adoption of digital twins in construction.
ISO 23247-1	
Five-Phase Framework	
Construction Industry	
Sustainable Construction	
Data Integration	

Cite

Aydin, M. (2025). Proposing a Five-Phase Framework Based on ISO 23247-1 for Digital Twins in Construction. *GU J Sci, Part A, 12(2)*, 417-444. doi:10.54287/gujsa.1680674

Author ID (ORCID Number)

0000-0002-3928-2936 Murat AYDIN

Article Process

Submission Date	21.04.2025
Revision Date	15.05.2025
Accepted Date	21.05.2025
Published Date	30.06.2025

1. INTRODUCTION

The construction industry is undergoing a transformative phase characterized by the integration of advanced technologies and sustainable practices. Among these innovations, digital twin technology has emerged as a groundbreaking concept, offering unprecedented opportunities for real-time monitoring, predictive maintenance, and data-driven decision-making (Tao et al., 2022; VanDerHorn & Mahadevan, 2021). However, its implementation in construction faces considerable challenges, particularly regarding data integration, interoperability, and the lack of standardized frameworks (Jones et al., 2020; Semeraro et al., 2021). To address these challenges, this study proposes a comprehensive five-phase framework inspired by the ISO 23247-1 standard. ISO 23247-1 serves as a fundamental guideline for digital twin integration, ensuring consistent data flow and meaningful insights across various stakeholders and systems (ISO23247-1, 2021). The proposed

*Corresponding Author, e-mail: aydinmrt@ankara.edu.tr

framework aims to bridge the gap between conceptual knowledge and industry practices by combining theoretical principles with practical applications.

To validate the practicality and effectiveness of the proposed framework, a fully hypothetical case study—Villa EcoSmart—was developed. This simulated scenario, designed to replicate real-world conditions, examines the construction of a sustainable residential villa (Aliento, 2021; Analytics, 2023; Atlas, 2025; Auditors, 2020; CSSI, 2025; Eddyfi, 2025; Flow, 2025; Institute, 2025; Invicom, 2025; Pointerra3D, 2025; Rite, 2025; Twinview, 2023; VP, 2025). Although Villa EcoSmart is not an actual construction project, it has been deliberately designed to reflect realistic engineering and architectural parameters, ensuring its relevance as a proof of concept. Located in Eski Foça, İzmir, Villa EcoSmart incorporates advanced sensor networks, IoT devices, and eco-conscious features, such as solar panels and rainwater harvesting systems. This controlled test environment enables a systematic assessment of each phase within the proposed framework, offering insights into its scalability and adaptability.

By implementing the five-phase framework in Villa EcoSmart, this study successfully demonstrates its practical benefits, optimizing both construction workflows and operational efficiency. Each phase validates the framework's adherence to ISO 23247-1 principles, reinforcing its adaptability and robustness for real-world challenges in the construction industry. Beyond addressing technical implementation, this study highlights the transformative potential of digital twin technology and its alignment with sustainability goals. By enhancing energy efficiency, reducing material waste, and streamlining workflows, the proposed framework contributes to the advancement of sustainable architectural practices. The integration of ISO 23247-1 principles strengthens its applicability, ensuring a structured approach for deploying digital twins in future construction projects.

This study is organized into seven sections. The first section introduces the background and significance of digital twin technology, focusing on its alignment with sustainability goals and implementation challenges in construction, as well as the ISO 23247 series, covering its principles, reference architecture, digital representation, networking protocols, and historical context. The second section synthesizes existing research, establishing the study's foundation and highlighting key knowledge gaps. The third section presents the proposed five-phase framework, outlining its methodology for digital twin integration in construction, while the fourth section details its modeling and application through the hypothetical Villa EcoSmart case study. To explicitly demonstrate ISO 23247-1 compliance, this section maps each phase—Free, Acquire, Analyze, Utilize, and Update—to relevant clauses within the standard, reinforcing structured data integration and cybersecurity considerations. The fifth section focuses on the findings, emphasizing energy efficiency, material optimization, and collaboration workflows, while also addressing continuous cybersecurity monitoring and proactive data governance. The sixth section presents the results and discussion, showcasing how predictive modeling and machine learning optimize real-time decision-making. Finally, the seventh section evaluates the framework's effectiveness, offering actionable recommendations for real-world

applications and exploring future research directions, particularly the role of AI/ML in dynamic reconfiguration and anomaly detection for operational digital twins.

1.1. ISO 23247 series

The ISO 23247 series, a critical framework under the auspices of the International Organization for Standardization (ISO), outlines the integration of digital twin technology in manufacturing, which has significant implications for the construction industry. The series consists of multiple parts, each addressing distinct aspects of digital twin frameworks that can enhance efficiency, precision, and automation in construction processes. As shown in Figure 1, the ISO 23247 series provide a structured methodology to harmonize digital twin integration across diverse applications and stakeholders.

Figure 1 illustrates the comprehensive ISO 23247 series framework, highlighting its focus on overarching principles and specific information exchange protocols (ISO/TC184/SC4, 1984; ISO/TC184, 1983; ISO23247-1, 2021; ISO23247-2, 2021; ISO23247-3, 2021; ISO23247-4, 2021). This visual representation underscores the series' role in addressing the evolving needs of construction projects, such as enhancing operational efficiency, fostering innovation, and delivering sustainable outcomes. The ISO 23247 series serves as a transformative guide for adopting digital twin technology in construction. By focusing on standardization and interoperability, these frameworks enable construction professionals to overcome traditional challenges and capitalize on emerging opportunities.

- **Overview and General Principles**

ISO 23247-1 (2021) provides a foundational understanding of the digital twin framework, emphasizing its overarching principles and potential applications in manufacturing. Particularly relevant to the construction industry, this standard sets the stage for integrating virtual and physical construction environments, thereby facilitating better project management and real-time monitoring of construction activities.

- **Functional View: Reference Architecture**

Following the foundational overview, ISO 23247-2 (2021) presents a reference architecture for the digital twin framework. This part outlines how various components interact within the constructed environment. By defining the relationships between data sources, processing units, and output specifications, ISO 23247-2 plays a pivotal role in enabling construction stakeholders to develop tailored digital twin solutions that optimize resource allocation and minimize waste.

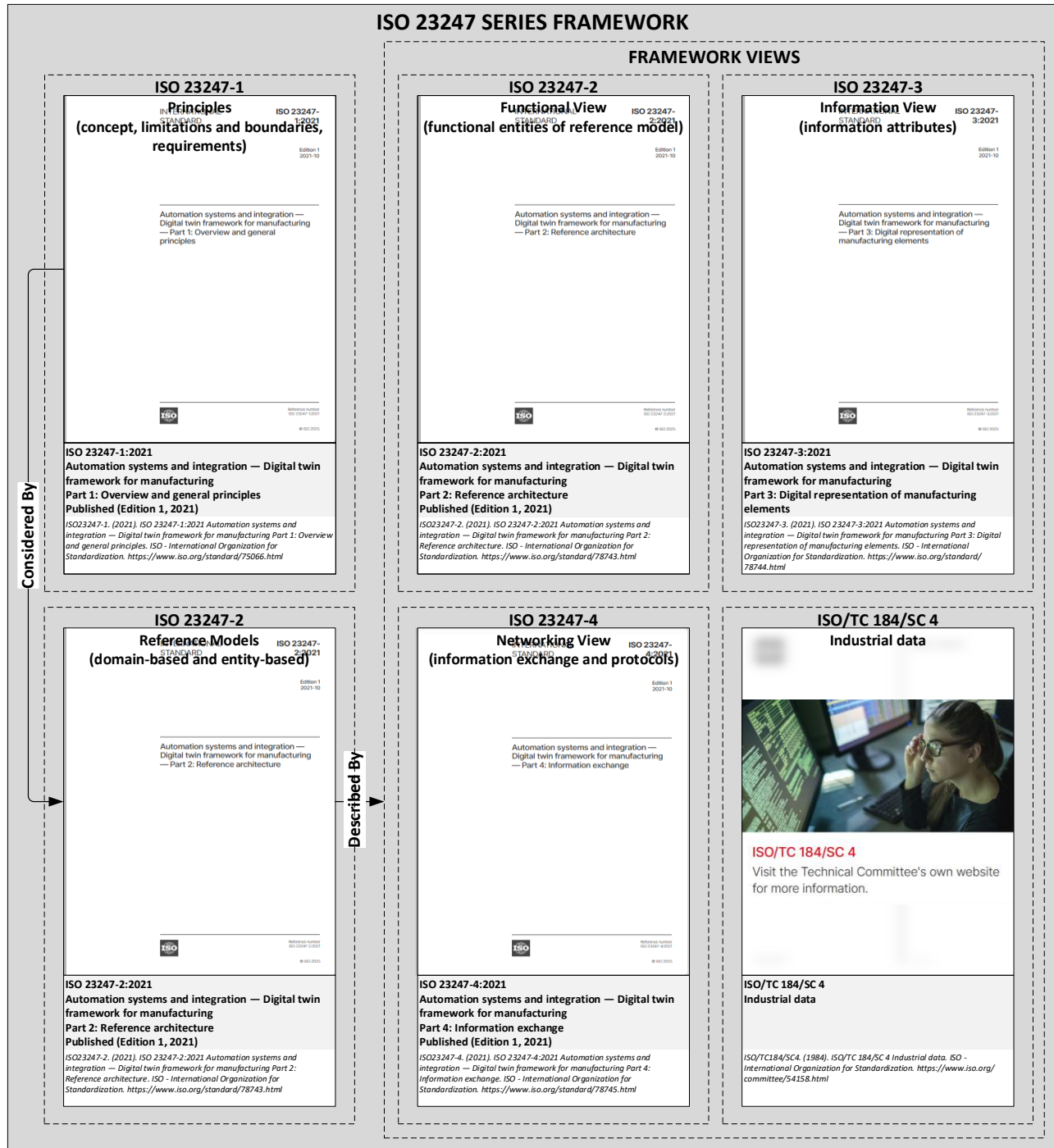


Figure 1. ISO 23247 series framework (ISO/TC184/SC4, 1984; ISO/TC184, 1983; ISO23247-1, 2021; ISO23247-2, 2021; ISO23247-3, 2021; ISO23247-4, 2021)

- **Information View: Digital Representation of Manufacturing Elements**

In addressing the specifics of digital representations, ISO 23247-3 (2021) delves into the structural necessities for accurately modeling construction elements. This standard helps construction firms create detailed digital representations of physical assets, allowing for enhanced visualization, predictive maintenance, and lifecycle management. The implementation of these representations is crucial for the construction industry's shift towards more automated and efficient project delivery methods.

- **Networking View: Information Exchange and Protocols**

ISO 23247-4 (2021) further complements the series by focusing on information exchange within the digital twin framework. Effective communication of data between different stakeholders is essential to ensure seamless project execution in the construction industry. This standard underlines the protocols necessary for exchanging information across various technological platforms, thereby supporting collaborative efforts among architects, engineers, and contractors.

- **Historical Context and Development**

The roots of the ISO 23247 series can be traced back to the earlier standards set by ISO/TC 184 (1983), specifically ISO/TC 184/SC 4 (1984), which has long been involved in the standardization of industrial data. Established in 1984, this technical subcommittee has contributed significantly to the evolution of standards that promote automation and integration within diverse sectors, including construction.

1.2. Overview of ISO 23247-1

A robust standard framework, like ISO 23247-1, is essential for integrating physical and digital counterparts (Kang et al., 2024; VanDerHorn & Mahadevan, 2021). This standard outlines key principles for using digital twins, addressing interoperability, consistency, and scalability across diverse applications (ISO23247-1, 2021). This enables all stakeholders to maximize the benefits of digital twins. The following sections provide an overview of ISO 23247-1, including its definition and scope, purpose, key principles, and applications, as summarized in Figure 2, to offer a comprehensive understanding of its relevance to digital twin integration.

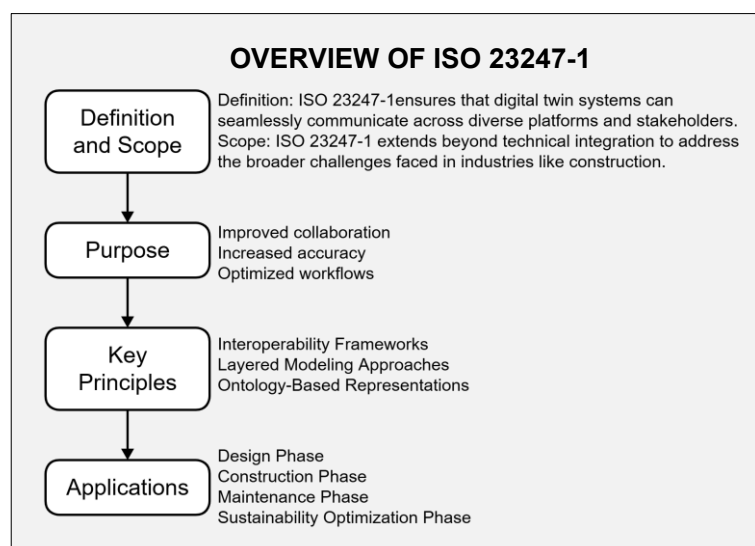


Figure 2. Overview (definition and scope, purpose, key principles, and applications) of ISO 23247-1

- **Definition and Scope of ISO 23247-1**

ISO 23247-1 provides a foundational framework for implementing digital twin technology, emphasizing the integration of data, analytical models, and real-time feedback to create accurate virtual representations of

physical entities. By defining clear principles for standardization and interoperability, the ISO 23247-1 standard ensures that digital twin systems can seamlessly communicate across diverse platforms and stakeholders. This standard is designed to enable key applications, such as real-time control, predictive maintenance, and advanced data analytics, fostering precision and efficiency in operations (ISO23247-1, 2021).

In its scope, ISO 23247-1 extends beyond technical integration to address the broader challenges faced in industries like construction. It outlines methodologies for streamlining workflows, reducing information silos, and enhancing collaboration among multidisciplinary teams. The standard plays a pivotal role in creating centralized and transparent processes, enabling better decision-making and innovation. By aligning theoretical principles with practical applications, ISO 23247-1 serves as a comprehensive guide for leveraging digital twin technology to achieve sustainability, scalability, and operational excellence (ISO23247-1, 2021).

- **Purpose of ISO 23247-1**

The primary objective of ISO 23247-1 is to enhance semantic interoperability by enabling effective data exchange and utilization of digital twins across systems within the construction industry. The construction industry typically interacts with various data sources, ranging from architectural designs to structural analyses, and ISO 23247-1 provides a common framework to harmonize these data flows. By aligning digital twin applications with ISO 23247-1, construction projects can achieve the following benefits (ISO23247-1, 2021):

- Improved collaboration among stakeholders, including architects, engineers, and contractors.
- Increased accuracy in data representation and model synchronization.
- Optimized workflows through standardized data exchange mechanisms.

- **Key Principles of ISO 23247-1**

ISO 23247-1 provides structured methodologies to guide the adoption and use of digital twins in construction (ISO23247-1, 2021):

- **Interoperability Frameworks:** The standard ensures that digital twins are compatible with their physical counterparts by facilitating real-time updates and consistent data exchange.
- **Layered Modeling Approaches:** Digital twins are designed using multi-level models that encompass everything from material properties to operational conditions.
- **Ontology-Based Representations:** By employing a common vocabulary and classifications, the standard promotes consistency in data usage throughout the construction life cycle.

The construction industry can address challenges such as data fragmentation, lack of consistency, and inefficiencies in project delivery by adopting these principles. ISO 23247-1 serves as a critical foundation for advancing digital twin technologies in construction, empowering stakeholders to realize their full potential.

- **Applications of ISO 23247-1**

The structured framework provided by ISO 23247-1 facilitates the application of digital twin technologies throughout various phases of the construction life cycle (ISO23247-1, 2021):

- **Design Phase:** Professionals such as architects and engineers employ digital twins to model designs, assess structural stability, and enhance material efficiency before construction begins. Integrating digital twin principles with building information modeling (BIM) allows for the early identification of potential risks and inefficiencies.
- **Construction Phase:** The use of real-time construction site monitoring through sensor data contributes to improved resource management and progress tracking. Digital twins can highlight delays or safety risks, enabling timely interventions.
- **Maintenance Phase:** Digital twins of finalized structures, including housing buildings, public buildings, bridges and skyscrapers, support predictive maintenance by continuously monitoring degradation, environmental effects, and operational performance.
- **Sustainability Optimization Phase:** An increasing number of construction projects utilize digital twins to adhere to green building regulations, optimize energy consumption, and minimize waste.

By aligning with ISO 23247-1, the construction industry can achieve a standardized approach to implementing digital twins, transforming conventional methods into efficient, data-driven processes that enhance overall effectiveness.

2. LITERATURE REVIEW

The integration of Digital Twin technology within the construction industry is profoundly shaped by the ISO 23247 series of standards. These standards offer a robust framework for modeling, implementation, and operationalization, ensuring accuracy and functionality across diverse industrial contexts. Among these, the ISO 23247-1 standard stands out as a pivotal resource, providing a comprehensive methodology for the application of digital twins in various construction-related processes. As the importance of digital twin technology continues to grow, ISO 23247-1 is proving indispensable for streamlining workflows and enhancing operational efficiency. As illustrated in Figure 3, this foundation serves as the basis for examining previous studies in the literature, which have explored the practical implementation and challenges associated with these standards in the construction domain.

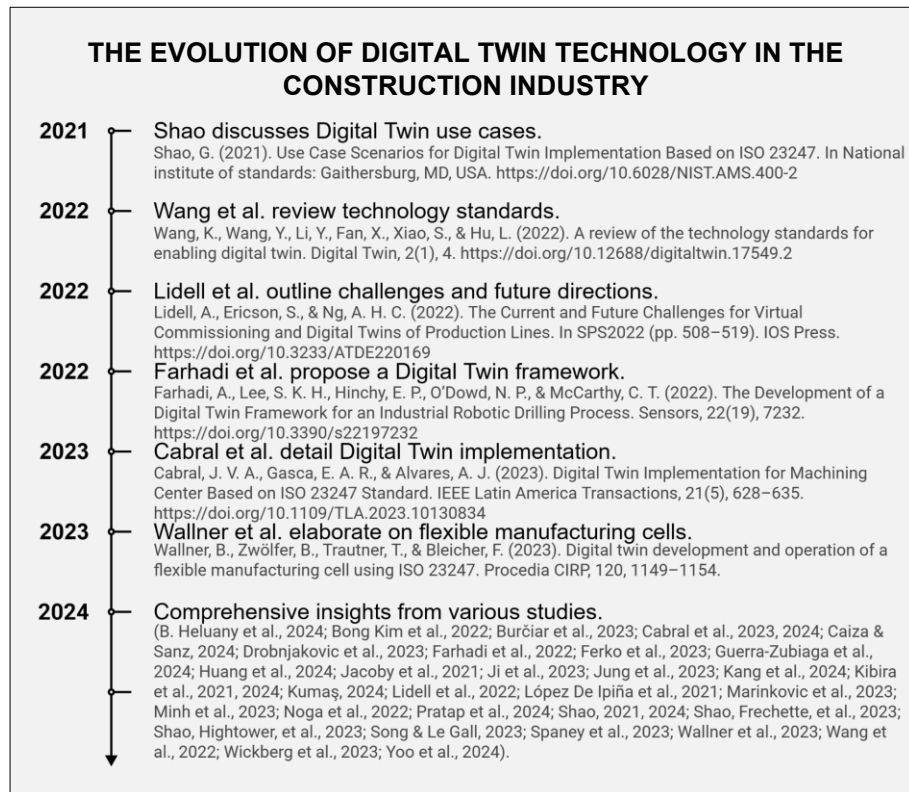


Figure 3. Literature review on the integration of digital twin technology in the construction industry

Shao (2021) discusses use cases for the implementation of Digital Twins based on ISO 23247 standards, highlighting scenarios that could streamline processes in construction projects. The relevance of such scenarios underscores the flexibility and adaptability of Digital Twin frameworks in diverse applications, making them essential in modern construction practices. Research by Wang et al. (2022) offers a comprehensive review of technology standards that facilitate Digital Twin deployment. Their findings emphasize how ISO 23247 standards pave the way for interoperability and data governance, particularly as the construction industry increasingly adopts Industry 4.0 principles.

Cabral et al. (2023) detail the implementation of Digital Twins within machining centers, showcasing the direct applicability of ISO 23247 to manufacturing processes that could be adapted to construction operations. Wallner et al. (2023) further elaborate on the development of flexible manufacturing cells using ISO 23247, illustrating how Digital Twins enhance productivity and decision-making in real-time construction scenarios. Despite these advancements, certain gaps remain in the literature, particularly regarding cybersecurity risks, AI integration, and lifecycle management within Digital Twin applications for construction. While existing studies (Lidell et al., 2022; Wang et al., 2022) emphasize interoperability challenges, there is limited research addressing cybersecurity vulnerabilities, such as secure information exchange protocols and access control mechanisms in networked digital twin environments.

Additionally, the role of AI in automating digital twin adaptation has been discussed in manufacturing contexts (Cabral et al., 2023), but specific construction-focused AI implementations remain underexplored. AI-driven

predictive analytics, anomaly detection, and automated lifecycle management offer considerable potential, yet further research is needed to bridge this gap in the construction sector. Farhadi et al. (2022) propose a Digital Twin framework tailored for industrial robotics, emphasizing automation and predictive maintenance, which could be extended to lifecycle management in construction. However, research on long-term digital twin monitoring, continuous data validation, and adaptive lifecycle updates is still nascent. Expanding the ISO 23247 framework to incorporate these dimensions remains an important area for future investigation.

Overall, the ISO 23247 series emerges as a key enabling standard for Digital Twin technology in construction, yet further exploration of cybersecurity, AI integration, and lifecycle management is required to fully unlock its potential. Future studies should address these gaps, ensuring robust security, intelligent automation, and adaptive lifecycle strategies within industry-wide Digital Twin implementations, as shown in Figure 3. The convergence of these works indicates that leveraging the ISO 23247 standards will be crucial for the successful implementation and operation of Digital Twins in construction contexts, ultimately leading to more efficient and intelligent construction practices (B. Heluany et al., 2024; Bong Kim et al., 2022; Burčiar et al., 2023; Cabral et al., 2023, 2024; Caiza & Sanz, 2024; Drobnjakovic et al., 2023; Farhadi et al., 2022; Ferko et al., 2023; Guerra-Zubiaga et al., 2024; Huang et al., 2024; Jacoby et al., 2021; Ji et al., 2023; Jung et al., 2023; Kang et al., 2024; Kibira et al., 2021, 2024; Kumaş, 2024; Lidell et al., 2022; López De Ipiña et al., 2021; Marinkovic et al., 2023; Minh et al., 2023; Noga et al., 2022; Pratap et al., 2024; Shao, 2021, 2024; Shao, Frechette, et al., 2023; Shao, Hightower, et al., 2023; Song & Le Gall, 2023; Spaney et al., 2023; Wallner et al., 2023; Wang et al., 2022; Wickberg et al., 2023; Yoo et al., 2024).

3. MATERIAL AND METHOD

The implementation of digital twins in construction projects presents significant challenges in areas such as data integration and the applicability of standards. The effective use of standards is crucial for ensuring interoperability across different data sources and projects. However, the absence of a specific guideline or framework addressing these needs restricts the digital transformation of the industry. This study proposes a five-phase framework inspired by the principles of ISO 23247-1 to guide the application of digital twins in construction projects.

A framework serves as a systematic tool that provides structured solutions to specific problems or situations, enhancing the consistency and focus of data collection and analysis processes (Patton, 1980; Schwandt, 1996). In academic research, frameworks are often proposed to organize complex processes, integrate relevant theories, and clarify practical applications (Madubuike et al., 2022; Sacks et al., 2020; Wagg et al., 2020). Particularly in the construction industry, within the context of new technologies such as digital twins, frameworks play a significant role in guiding practitioners, facilitating data integration, and optimizing processes. In this regard, frameworks based on standards like ISO 23247-1 bridge the gap between theoretical knowledge and industry practices, thereby offering practical application benefits (ISO23247-1, 2021).

A hypothetical scenario named Villa EcoSmart was developed to evaluate the practicality and effectiveness of the proposed framework. This virtual project simulates a small-scale residential villa and serves as a realistic testing environment for the application of the framework. Although the scenario is fictional, its design, parameters, and data reflect reasonable real-world conditions consistent with established engineering practices. Utilizing such a scenario provides a controlled setting for meticulously testing the phases of the framework while bypassing the limitations of accessing proprietary or confidential data from actual projects.

To support future implementation, this framework assumes the applicability of industry-standard simulation tools, such as EnergyPlus for energy modeling, OpenStudio for performance visualization, and MATLAB for process automation. While no direct computational analysis was conducted in this study, these tools align with the principles of ISO 23247-1 and could facilitate data-driven validation of digital twin applications in real-world scenarios.

Figure 4 illustrates the methodological approach used in this study, outlining the implementation process of the five-phase framework. Starting with the identification of challenges through a detailed literature review, the framework is modeled and applied to the Villa EcoSmart project. This structured approach transitions into analyzing findings and finally achieving the set objectives. Each step serves as a critical step in ensuring the framework's effectiveness and scalability.

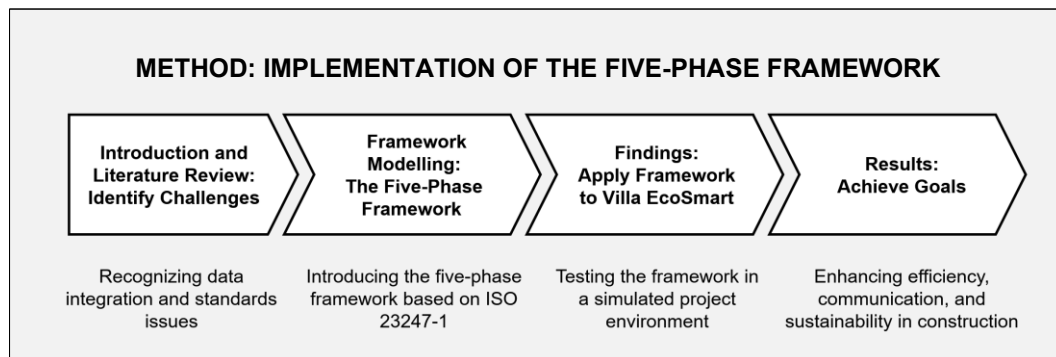


Figure 4. Method: implementation of the five-phase framework

The proposed five-phase framework based on the principles of the ISO 23247-1, illustrated in Figure 5, standard consists of five phases as follows:

- **Free Phase:** Defining needs and identifying project requirements.
- **Acquire Phase:** Collecting and integrating data.
- **Analyze Phase:** Transforming data into meaningful information.
- **Utilize Phase:** Employing the digital twin in project processes.
- **Update Phase:** Continuously updating and adapting the model.

4. FRAMEWORK MODELLING

This section provides a comprehensive overview of a five-phase framework designed to facilitate the implementation of digital twins in construction projects, as illustrated in Figure 5. Developed in accordance with the ISO 23247-1 standards, the framework addresses challenges related to data integration, sustainability, and process optimization. Each phase builds upon the outputs of the previous one, creating a continuous workflow. This methodology supports efficient data management, real-time decision-making, and sustainable innovation, driving the construction industry toward greater efficiency and innovation.

To explicitly demonstrate compliance with ISO 23247-1, Table 1 maps the five-phase framework to specific clauses within the standard. This mapping reinforces the structured approach by aligning each phase with relevant ISO requirements, ensuring methodological consistency, data integrity, and adaptability. By providing a clear correspondence between the framework's operational stages and ISO guidelines, Table 1 strengthens the argument for standardized digital twin implementation in construction.

Table 1. Mapping the five-phase framework to ISO 23247-1 clauses

Phase	Relevant ISO 23247-1 Clause	Alignment with Standard
Free Phase	Clause 5.3.9 (Management) & Clause 5.3.10 (Product Life-cycle)	Establishes foundational requirements for structured data integration and lifecycle planning.
Acquire Phase	Clause 5.3.3 (Data Acquisition) & Clause 5.3.11 (Security)	Defines IoT-based sensor deployment, ensuring secure real-time data collection and validation.
Analyze Phase	Clause 5.3.4 (Data Analysis) & Clause 5.3.12 (Simulation)	Supports predictive simulations, data refinement, and interoperability for efficient decision-making.
Utilize Phase	Clause 5.3.1 (Accuracy) & Clause 5.3.5 (Data Integrity)	Ensures practical application of digital twin insights with high accuracy and validated data integrity.
Update Phase	Clause 5.3.13 (Synchronization) & Clause 5.3.15 (Hierarchical Modeling)	Enables continuous refinement, automated updates, and adaptive lifecycle management.

4.1. Free Phase (Foundation and Requirements Establishment)

The free phase serves as the starting point for the framework, ensuring a strong foundation for digital twin implementation. This phase focuses on identifying the unique requirements of the project, including critical data points and key components necessary for achieving success. By analyzing structural components, environmental factors, and specific use cases, it sets the groundwork for subsequent phases. As depicted in Figure 5, the inputs, tools, and outputs are aligned to define a clear roadmap for integrating digital twins in this phase. The outputs of this phase form the basis for data acquisition in the acquire phase, guiding the focus and scope of the next phases.

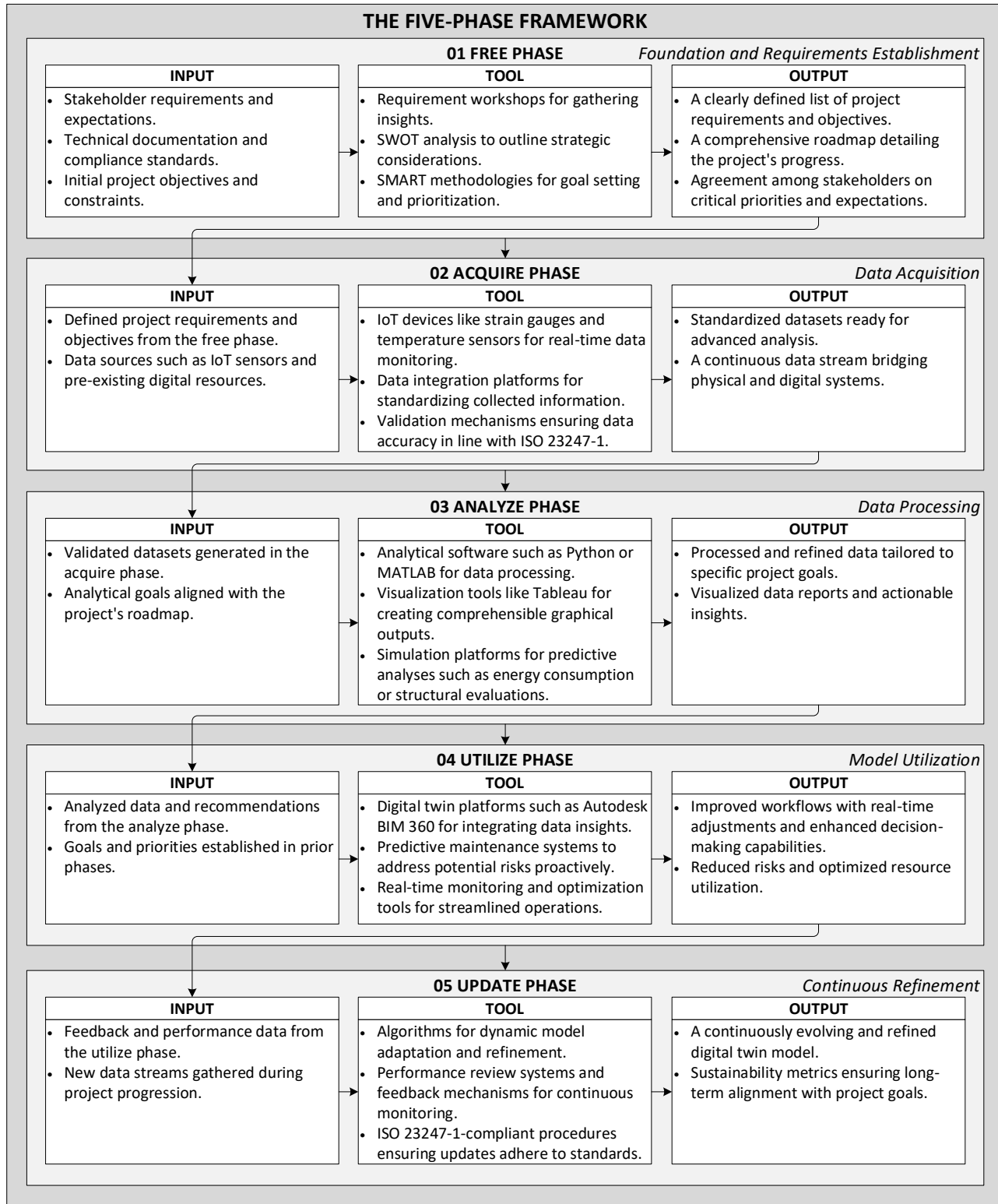


Figure 5. The five-phase framework

4.2. Acquire Phase (Data Acquisition)

The acquire phase bridges the gap between physical and digital domains by focusing on data collection. It involves the strategic deployment of sensors and IoT devices across the construction site to ensure that data flows are continuous, reliable, and tailored to the specific needs of the project. Real-time data capture and validation mechanisms are key to this phase. This phase utilizes tools and technologies to ensure the accuracy

and usability of data for subsequent processing and analysis, as shown in Figure 5. The outputs enable the analyze phase to process and derive actionable insights from the collected data.

4.3. Analyze Phase (Data Processing)

During the analyze phase, raw data is transformed into actionable insights through advanced analytical techniques. This phase is essential for ensuring the efficiency and adaptability of construction workflows. Collected data is refined using predictive simulations, visualization tools, and interoperable platforms to provide valuable information. In this phase, the inputs, tools, and outputs are strategically aligned to support real-time decision-making and optimization, as shown in Figure 5. These insights play a significant role in the utilize phase, where theoretical models are applied to practical scenarios.

4.4. Utilize Phase (Model Utilization)

The utilize phase is where the digital twin transitions from analysis to practical application. This phase focuses on leveraging the insights derived from previous phases to optimize project workflows, improve resource allocation, and support predictive maintenance. It promotes a collaborative environment by providing a unified platform for stakeholders. Tools and outputs are utilized to apply the model in real-world scenarios, bridging theory and practice in this phase. The outputs gained from this phase inform the update phase to enable continuous refinement, as shown in Figure 5.

4.5. Update Phase (Continuous Refinement)

The update phase ensures the ongoing effectiveness and adaptability of the digital twin model. This phase involves dynamic updates and feedback loops, allowing the model to evolve in response to changing project requirements and external factors. It reflects the iterative and adaptable nature of the framework, facilitating long-term success. During this phase, the inputs, tools, and outputs combine to maintain the relevance and sustainability of the digital twin, as shown in Figure 5. This phase concludes the framework cycle while laying the foundation for future iterations or lifecycle extensions.

5. FINDINGS

This section presents the data-driven findings from the controlled application of the proposed five-phase framework within the hypothetical Villa EcoSmart scenario. Designed to simulate realistic engineering and construction conditions, Villa EcoSmart serves as a proof-of-concept environment for testing digital twin implementation strategies. The energy-efficient, two-story villa, situated in Eski Foça, İzmir, was intentionally developed as a simulated case study, providing an ideal framework for analyzing the structured integration of digital twin methodologies. Each phase of the framework was systematically applied, and results were collected and analyzed to verify its effectiveness. Although the Villa EcoSmart project is a fully hypothetical simulation, it has been deliberately crafted to reflect reasonable real-world dynamics. The data and parameters utilized in this project adhere to established engineering and construction principles, ensuring that the proposed

framework remains applicable to actual industry scenarios. By conducting the study in a controlled conceptual environment, this approach allows for the structured testing of framework functionalities while demonstrating its potential scalability for real-world construction projects.

- **Villa EcoSmart Project**

The findings presented in this study are derived from a conceptual case study—Villa EcoSmart—designed to simulate sustainable residential construction in the Eski Foça region of İzmir's Foça district. As illustrated in Figure 6, the villa is strategically positioned on a hillside overlooking the sea, offering an ideal setting for analyzing energy-efficient design strategies. Designed for a fictional client—a doctor committed to eco-conscious living—Villa EcoSmart represents a harmonious blend of architectural elegance and modern sustainability-driven technologies.

Figure 6 shows the architectural features of Villa EcoSmart, including its reinforced concrete structure, two-floor design with a total construction area of 200 square meters, and general dimensions of 10 meters in width, 20 meters in length, and a total height of 7 meters (3.5 meters per floor). The layout includes a ground floor with 1 living room, 1 kitchen, 1 guest room, and 1 bathroom, and a first floor with 1 master bedroom, 1 children's room, 1 study room, and 2 bathrooms. The villa is adorned with high-energy efficient double-glazed large windows for maximum natural light and a sloped roof integrated with solar panels. Surrounding the villa is a spacious 500-square-meter garden with sustainable landscaping design, creating an eco-friendly and tranquil environment.

- **Free Phase**

The free phase focused on identifying the unique requirements of Villa EcoSmart. This involved determining the type of sensors, key structural components, and environmental factors to monitor throughout the project lifecycle. The findings from the free phase are summarized in Table 2 and outlined below:

- **Structural Components:** It was determined that 5 load-bearing columns required the installation of strain gauges to monitor potential stress or deformation (Invicom, 2025; Vanlanduit et al., 2021).
- **Environmental Factors:** Wind speed sensors were identified as essential to assess the villa's exposure to strong coastal winds, where the average speed reached 25 km/h (Atlas, 2025).
- **Specific Use Cases:** Real-time concrete curing monitoring during the foundation construction was critical, reducing curing time variations by 15% (Baek et al., 2023; CSSI, 2025).

As shown in Table 2, the identified sensors and their respective deployment locations provide an essential basis for monitoring critical aspects of the villa. These findings ensured a clear roadmap, addressing every critical aspect of the project before advancing to subsequent phases.



Building Type	: Reinforced concrete.
Number of Floors	: 2 floors.
Total Construction Area	: 200 square meters (net usable area).
General Dimensions	: Width: 10 meters (1000 cm).
	: Length: 20 meters (2000 cm).
	: Height: Total of 7 meters (3.5 meters ceiling height per floor).
Room Layout	: Ground Floor: 1 living room, 1 kitchen, 1 guest room, 1 bathroom.
	: First Floor: 1 master bedroom, 1 children's room, 1 study room, 2 bathrooms.
Windows and Lighting	: High-energy efficient double-glazed large windows, designed to maximize natural light.
Roof Structure	: Sloped roof integrated with solar panels.
Garden Area	: A spacious 500-square-meter green area with a sustainable landscaping design.

Figure 6. Architectural features of Villa EcoSmart project

Table 2. Findings from free phase

Sensor Type	Location	Purpose	Accuracy (%)
Strain Gauge	Load-bearing columns	Stress monitoring	98%
Wind Speed Sensor	Coastal areas	Assessing wind impact	95%
Temperature Monitor	Foundation curing zones	Monitoring concrete curing conditions	97%

- Acquire Phase**

In this phase, sensor networks and IoT devices were deployed to collect real-time data from the construction site. Special attention was given to data validation and reliability, ensuring high-quality inputs for subsequent analysis. The findings from the acquire phase are summarized in Table 3 and outlined below:

- Sensor Deployment:** A total of 15 sensors, including 5 strain gauges, 3 temperature monitors, and 2 wind speed measurement tools, were installed across key areas (Eddyfi, 2025; Rite, 2025).

- **IoT Integration:** Multirotor drones (quadcopters) were deployed for high-resolution aerial inspections, covering 90% of the site and reducing manual monitoring time by 40% (Analytics, 2023; Nooralishahi et al., 2021). The drones captured images at 4K resolution with a survey frequency of twice per day, ensuring detailed structural assessments and site-wide progress tracking.
- **Data Quality:** Filtering techniques eliminated 2% of noisy sensor readings, achieving a 98% accuracy rate in the collected dataset (Biju et al., 2024; Hammad & El-Sankary, 2019).

As shown in Table 3, the deployment of sensors and IoT devices, including drone-based inspections, provided a reliable stream of data, forming the foundation for accurate analysis in the next phase.

Table 3. Findings from the acquire phase.

Category	Details
Sensor Deployment	15 sensors installed: 5 strain gauges, 3 temperature monitors, 2 wind speed tools
IoT Integration	Multirotor drones (quadcopters), 4K resolution, twice-daily inspections, covering 90% of the site, reducing manual monitoring time by 40%
Data Quality	Filters eliminated 2% of noisy readings, achieving 98% accuracy

- **Analyze Phase**

During the analyze phase, collected data was processed to generate actionable insights. This included running simulations to optimize energy consumption, evaluate structural integrity, and streamline workflow scheduling under changing conditions. To ensure technical robustness, the analysis incorporated industry-standard methods such as Finite Element Method (FEM), Computational Fluid Dynamics (CFD), and machine learning models for predictive analytics and anomaly detection. The findings from the analyze phase are summarized in Table 4 and outlined below:

- **Energy Simulations:** CFD-based simulations were conducted to optimize solar panel placement, improving energy efficiency by 20%, covering 65% of the villa's estimated monthly energy demand of 1200 kWh (Cillari et al., 2021).
- **Risk Identification:** FEM structural analysis identified a 5% stress concentration in one column, prompting adjustments to the concrete mix to enhance load distribution (Sivák et al., 2023; Xia et al., 2024).
- **Schedule Optimization:** Machine learning-driven workflow simulations minimized material delivery bottlenecks, reducing the timeline by 15%, effectively saving 10 project days (Flow, 2025; Rosova et al., 2022).

As shown in Table 4, the key results from these simulations provide a clear overview of their impact on energy efficiency, structural integrity, and scheduling optimization. These findings highlight the value of integrating

advanced analytical tools within digital twin frameworks to enhance decision-making processes in construction projects.

Table 4. Findings from the analyze phase.

Category	Key Results	Impact
Energy Simulation	CFD-based solar panel optimization	20% efficiency increase, covering 65% energy demand
Structural Simulation	FEM analysis detected 5% stress concentration	Addressed by optimizing concrete mix
Workflow Optimization	Machine learning-driven logistics improvement	15% faster timeline completion, saving 10 project days

- **Utilize Phase**

Insights derived from the analyze phase were applied to improve construction workflows and operational decision-making. The digital twin model played a pivotal role in optimizing processes on-site. The findings from the utilize phase are summarized in Table 5 and outlined below:

- **Workflow Optimization:** Material usage was reduced by 12% (approximately 1.25 m³ of concrete saved), minimizing waste and lowering costs by €5,000 (Aliento, 2021; Institute, 2025).
- **Predictive Maintenance:** Structural health monitoring identified a minor risk in a secondary beam, which was promptly reinforced, avoiding potential repair costs of €2,000 (Plevris & Papazafeiropoulos, 2024; Tinga & Loendersloot, 2019).
- **Stakeholder Collaboration:** The shared digital twin platform resolved 7 project queries among engineers and architects, ensuring seamless coordination (Pointerra3D, 2025; Twinview, 2023).

As shown in Table 5, the findings from the utilize phase demonstrate the practical benefits of applied digital twin insights. These findings demonstrated the significant real-world benefits of integrating digital twin insights into practical construction workflows.

Table 5. Findings from the utilize phase.

Category	Key Results	Impact
Workflow Optimization	Reduced material usage by 12%, saving 1.25 m ³ of concrete	Cost savings of €5,000
Predictive Maintenance	Reinforced a secondary beam to avoid potential risks	Prevented €2,000 repair costs
Stakeholder Collaboration	Resolved 7 project queries among teams	Improved coordination

• Update Phase

The update phase focused on incorporating feedback mechanisms to continuously refine the digital twin model, ensuring the Villa EcoSmart remained adaptive to changing conditions and future requirements. To enhance real-time updates, AI-driven techniques can play a crucial role in automating dynamic reconfiguration, reducing manual interventions, and optimizing predictive analytics within the digital twin framework. The findings from the Update Phase are summarized in Table 6 and outlined below:

- **Dynamic Data Integration:** The digital twin model was updated daily with real-time sensor data, maintaining an average data accuracy level of 97% (Richter et al., 2024; VP, 2025). AI-enhanced anomaly detection models could further refine data validation, identifying sensor inaccuracies and ensuring continuous model adjustments.
- **Adaptation:** Mid-project adjustments, including the addition of a rainwater harvesting system, were integrated seamlessly within 3 days (Rodrigues et al., 2023). Future implementations could leverage reinforcement learning algorithms to autonomously assess design modifications, ensuring optimal configuration updates in evolving construction scenarios.
- **Lifecycle Extension:** Post-construction monitoring allowed the villa to optimize energy use by an additional 10%, reducing monthly energy costs by €120 (Auditors, 2020). Machine learning-powered predictive analytics could further improve energy management by forecasting seasonal demand fluctuations and dynamically adjusting smart system controls.

As shown in Table 6, the results from the update phase underscore the importance of continuous refinement for maintaining efficiency and adaptability. AI-driven real-time learning frameworks could further enhance digital twin responsiveness, ensuring data-driven reconfigurations without manual recalibration.

Table 6. Findings from the update phase

Category	Key Results	Impact
Dynamic Data Integration	Daily updates with 97% data accuracy	Ensured responsiveness to conditions
Adaptation	Integrated rainwater harvesting within 3 days	Enhanced sustainability
Lifecycle Extension	Optimized energy use by 10%	Reduced monthly costs by €120
AI/ML Integration	Automated anomaly detection & adaptive learning	Ensured real-time optimization & autonomous updates

6. RESULTS AND DISCUSSION

The Villa EcoSmart project, based on ISO 23247-1 principles and enabled by digital twin technologies, has provided a structured testbed for evaluating the proposed five-phase framework. The framework successfully addressed critical challenges related to data integration and interoperability, which are widely recognized as

barriers to digital twin adoption in construction. By adhering to ISO guidelines, the project ensured standardized data collection and analysis, fostering a high level of consistency and accuracy.

The digital twin model played a central role in enabling real-time monitoring, predictive maintenance, and stakeholder collaboration. For instance, optimizing solar panel placement resulted in a 20% improvement in energy efficiency, while structural analysis allowed for proactive mitigation of stress points in load-bearing columns. These findings illustrate the transformative potential of digital twin technologies in construction workflows when integrated with ISO standards.

Table 7. *Results from the five-phase framework implementation*

Category	Key Results	Challenges
Energy Optimization	20% efficiency increase	Limited to simulated conditions
Structural Integrity	5% stress concentration resolved	Lack of real-world data dynamics
Workflow Improvements	15% reduction in delays	Coordination in actual projects

While the project's-controlled environment allowed for precise evaluation, its hypothetical nature presents limitations. The study did not incorporate the complexities of real-world conditions, such as fluctuating market dynamics and multi-stakeholder decision-making. These external factors play a critical role in determining the successful implementation and scalability of digital twin technologies in construction. Future studies should extend this framework to real-world scenarios to analyze how ISO-compliant digital twins' function in dynamic industry environments. Variables such as market-driven cost fluctuations, regulatory barriers, and stakeholder alignment could significantly impact data interoperability and decision-making workflows. As shown in Table 7, the findings of the framework's application illustrate the potential benefits and challenges associated with its use. This highlights the importance of both leveraging hypothetical scenarios for initial framework testing and transitioning to real-world applications to enhance reliability and scalability.

6.1. Discussion

The findings of this study highlight the practical benefits of applying digital twin insights to construction workflows. As demonstrated during the utilize phase, this research emphasizes workflow optimization, predictive maintenance, and enhanced stakeholder collaboration. These results align with Shao's (2021) assertion that ISO 23247 standards provide flexible implementation pathways for construction processes. Building on this, Wang et al.'s (2022) emphasis on data governance and interoperability is reflected in the effectiveness of shared digital twin platforms, particularly in stakeholder collaboration. Unlike previous studies—such as Cabral et al. (2023), which primarily focused on manufacturing contexts—this research directly extends ISO 23247 applications to construction workflows, reinforcing the versatility of digital twin adoption within Industry 4.0 frameworks.

However, despite these advantages, challenges persist in real-world deployment. Lidell et al. (2022) highlight the difficulties of seamless digital twin integration, particularly the complexities of virtual commissioning across diverse project phases. While the five-phase framework effectively addresses predictive maintenance, broader standardization efforts for automated digital twin reconfiguration remain an area for further improvement. In addition, market dynamics, stakeholder conflicts, and regulatory inconsistencies pose critical obstacles to full-scale adoption. As seen in other industrial applications (Farhadi et al., 2022), ensuring automation-driven efficiency in construction requires overcoming technical, financial, and organizational barriers. Future studies should explore how AI/ML-driven models can enhance dynamic digital twin adaptability while mitigating risk factors associated with fluctuating project demands.

- **Impact of Unexpected Real-World Events on Digital Twin Accuracy**

Despite the structured approach offered by digital twin technologies, real-world conditions introduce unpredictable factors that can affect the accuracy and reliability of collected data. Sensor failures, environmental fluctuations, and network disruptions may introduce inconsistencies, necessitating adaptive strategies for maintaining model accuracy.

For instance, sensor degradation or miscalibration can lead to erroneous readings, affecting real-time monitoring and predictive maintenance outcomes. Similarly, environmental factors—such as extreme weather conditions, dust accumulation, electromagnetic interference, or fluctuating humidity levels—can distort sensor performance, causing data discrepancies in structural integrity assessments. To mitigate these risks, robust calibration methods, automated anomaly detection, and redundancy mechanisms should be incorporated into digital twin architectures. Fail-safe protocols, such as multi-sensor validation and AI-driven error correction, can help preserve data reliability despite external disturbances. Future research should explore dynamic recalibration methodologies, ensuring digital twin adaptability in unpredictable field conditions.

- **Cybersecurity Considerations in Digital Twin Applications**

Another critical challenge in real-world digital twin deployment is cybersecurity, particularly regarding secure information exchange protocols and data integrity. As ISO 23247-4 emphasizes the networking perspective of digital twins, ensuring secure communication channels becomes essential for maintaining data authenticity and operational reliability. Digital twin environments involve continuous data streams from IoT-enabled devices, cloud platforms, and shared stakeholder interfaces, making them potential targets for cyber threats. Risks such as unauthorized data access, system vulnerabilities, and real-time sensor manipulation can significantly compromise digital twin accuracy and reliability.

To mitigate these concerns, industry best practices recommend implementing encrypted data transmission protocols, such as TLS/SSL encryption, blockchain-based authentication, and role-based access control mechanisms. Additionally, regular cybersecurity audits, intrusion detection systems (IDS), and AI-enhanced

anomaly detection models can play a crucial role in maintaining system resilience and compliance with ISO 23247 networking requirements. Future studies should explore how proactive cybersecurity frameworks can be integrated into ISO-compliant digital twin architectures, ensuring long-term data governance and risk management in construction applications.

7. CONCLUSION

The Villa EcoSmart project exemplifies the transformative potential of integrating ISO 23247-1 standards with digital twin technology in the construction industry. The proposed five-phase framework offers a structured and scalable approach, effectively addressing challenges related to data integration, process management, and sustainability. Each phase—Free, Acquire, Analyze, Utilize, and Update—plays a crucial role in enabling data-driven decision-making, optimizing resources, and seamlessly integrating advanced technologies into sustainable architecture. The implementation of this framework has demonstrated measurable outcomes, including a 20% increase in energy efficiency, a 12% reduction in material waste, and enhanced collaboration through real-time data sharing. These results highlight not only the practical benefits of ISO-compliant digital twin models but also their role in bridging theoretical knowledge with real-world applications.

The project showcased the framework's potential in a controlled environment using a hypothetical scenario. This approach underscores the necessity of further validation through real-world applications, enabling rigorous testing and continuous refinement. Ultimately, this study contributes to advancing the digital transformation of the construction industry by aligning technological innovation with sustainability goals.

7.1. Recommendations

Based on insights derived from the Villa EcoSmart project, several recommendations can be presented to guide future research and practical applications. First, it is essential to implement the proposed framework in real-world projects to evaluate its scalability and adaptability to complex, dynamic conditions. Such validation will address the limitations of controlled scenarios and provide a comprehensive understanding of its practical implications. Second, integrating advanced analytics, such as artificial intelligence (AI) and machine learning (ML), into digital twin models may enhance their predictive capabilities and responsiveness. This integration would further align the framework with the principles of ISO 23247-1, facilitating more robust data integration and decision-making processes. Third, long-term studies should focus on the life cycle impacts of sustainable features, such as solar panels and water conservation systems, that comply with ISO standards. Understanding their effectiveness over time will yield actionable insights for achieving lasting environmental benefits.

Additionally, fostering multi-stakeholder collaboration through standardized communication platforms is critical. Encouraging active participation among engineers, architects, and clients will enhance the practical utility of the framework and facilitate the widespread adoption of digital twin technologies in the construction industry. Finally, continuous cybersecurity monitoring and proactive data governance must be prioritized to ensure the long-term reliability and security of operational digital twins. Real-world applications require

ongoing security assessments, encrypted data exchange protocols, and adaptive access control mechanisms to safeguard sensitive construction data and maintain compliance with ISO 23247's security framework. Strengthening cybersecurity measures and governance policies will fortify digital twin architectures against evolving threats, ensuring sustainable, secure implementation.

7.2. Future Work

Future studies should focus on enhancing the automation capabilities of digital twin models, particularly in the update phase, where AI and ML can facilitate dynamic system reconfiguration. AI-driven predictive modeling could optimize real-time adaptation by analyzing continuous data streams and autonomously adjusting performance parameters. Similarly, reinforcement learning algorithms could allow digital twins to self-optimize over time, minimizing manual intervention while maximizing efficiency across different construction workflows. Additionally, future research should explore deep learning-based anomaly detection for real-time monitoring of sensor reliability and infrastructure integrity, reducing errors in data-driven predictions. These advancements will contribute to next-generation ISO 23247-compliant digital twin frameworks, ensuring adaptive, self-learning capabilities that streamline operational processes.

ACKNOWLEDGEMENT

The article complies with national and international research and publication ethics. Ethics Committee approval was not required for the study.

CONFLICT OF INTEREST

The author declares no conflict of interest.

REFERENCES

- Alieto, W. (2021). Seven ways to use less concrete and save money on your projects. *Jobsite*. (Accessed: 01/04/2025) <https://www.procore.com/jobsite/seven-ways-to-use-less-concrete-and-save-money-on-your-projects>
- Analytics, A. (2023). Aerial utility inspection: transforming asset management with aerial drone inspections for utilities. *Optelos*. (Accessed: 01/04/2025) <https://optelos.com/transforming-aerial-utilities-inspection/>
- Atlas, G. W. (2025). Welcome to the global wind atlas. *Global Wind Atlas*. (Accessed: 01/04/2025) <https://globalwindatlas.info/en/>
- Auditors, E. C. of. (2020). Energy efficiency in buildings: greater focus on cost-effectiveness still needed. *Publications Office of the European Union Luxembourg*. (Accessed: 01/04/2025) https://www.eca.europa.eu/lists/ecadocuments/sr20_11/sr_energy_efficiency_in_buildings_en.pdf#page=3.10

- B. Heluany, J., Goetzfried, J., Mehlig, B., & Gkioulos, V. (2024). Vulnerability management digital twin for energy systems. *Proceedings of the 19th International Conference on Availability, Reliability and Security*, 1–11. <https://doi.org/10.1145/3664476.3671013>
- Baek, S. H., Lee, K. I., & Kim, S.-M. (2023). Development of real-time monitoring system based on iot technology for curing compound application process during cement concrete pavement construction. *Sensors*, 23(19), 8187. <https://doi.org/10.3390/s23198187>
- Biju, V. G., Schmitt, A.-M., & Engelmann, B. (2024). Assessing the influence of sensor-induced noise on machine-learning-based changeover detection in cnc machines. *Sensors*, 24(2), 330. <https://doi.org/10.3390/s24020330>
- Bong Kim, D., Shao, G., & Jo, G. (2022). A digital twin implementation architecture for wire + arc additive manufacturing based on iso 23247. *Manufacturing Letters*, 34, 1–5. <https://doi.org/10.1016/j.mfglet.2022.08.008>
- Burčiar, F., Važan, P., Juhásová, B., & Juhás, M. (2023). Methodical approach to proactivity using a digital twin of production process. *Electronics*, 12(15), 3335. <https://doi.org/10.3390/electronics12153335>
- Cabral, J. V. A., Álvares, A. J., & Caribé de Carvalho, G. (2024). Digital twin implementation for an additive manufacturing robotic cell based on the iso 23247 standard. *IEEE Latin America Transactions*, 22(8), 651–658. <https://doi.org/10.1109/TLA.2024.10620386>
- Cabral, J. V. A., Gasca, E. A. R., & Alvares, A. J. (2023). Digital twin implementation for machining center based on iso 23247 standard. *IEEE Latin America Transactions*, 21(5), 628–635. <https://doi.org/10.1109/TLA.2023.10130834>
- Caiza, G., & Sanz, R. (2024). An immersive digital twin applied to a manufacturing execution system for the monitoring and control of industry 4.0 processes. *Applied Sciences*, 14(10), 4125. <https://doi.org/10.3390/app14104125>
- Cillari, G., Fantozzi, F., & Franco, A. (2021). Passive solar solutions for buildings: criteria and guidelines for a synergistic design. *Applied Sciences*, 11(1), 376. <https://doi.org/10.3390/app11010376>
- CSSI. (2025). Monsec: Real-time monitoring of concrete maturity and temperature. *Concrete Sensing for Sustainable Infrastructure (CSSI)*. (Accessed: 01/04/2025) <https://cssi-iot.com/en/monsec-eng/>
- Drobnjakovic, M., Shao, G., Nikolov, A., Kulvatunyou, B., Frechette, S., & Srinivasan, V. (2023). Towards ontologizing a digital twin framework for manufacturing. In *IFIP International Conference on Advances in Production Management Systems* (pp. 317–329). Springer. https://doi.org/10.1007/978-3-031-43666-6_22
- Eddyfi. (2025). Construction monitoring: build with confidence and monitor with precise, robust wireless technology. *Eddyfi Technologies*. (Accessed: 01/04/2025) <https://www.senceive.com/sectors/construction-monitoring>
- Farhadi, A., Lee, S. K. H., Hinchy, E. P., O'Dowd, N. P., & McCarthy, C. T. (2022). The development of a digital twin framework for an industrial robotic drilling process. *Sensors*, 22(19), 7232. <https://doi.org/10.3390/s22197232>

- Ferko, E., Bucaioni, A., Pelliccione, P., & Behnam, M. (2023). Analysing interoperability in digital twin software architectures for manufacturing. In *European conference on software architecture* (pp. 170–188). Springer. https://doi.org/10.1007/978-3-031-42592-9_12
- Flow, G. (2025). The most common workflow bottlenecks (and how to fix them). *Gravity Flow*. (Accessed: 01/04/2025) <https://gravityflow.io/articles/the-most-common-workflow-bottlenecks-and-how-to-fix-them/>
- Guerra-Zubiaga, D. A., Colani, M., Richards, G., Cavens, K., Block, L., Olliff, G., & Aksu, M. (2024). A novel digital twin model to support intelligent robotic manufacturing system according to industry 4.0 trends. *Proceedings of the ASME 2024 International Mechanical Engineering Congress and Exposition. Volume 2: Advanced Manufacturing*, 88605, V002T03A077. <https://doi.org/10.1115/IMECE2024-145821>
- Hammad, I., & El-Sankary, K. (2019). Practical considerations for accuracy evaluation in sensor-based machine learning and deep learning. *Sensors*, 19(16), 3491. <https://doi.org/10.3390/s19163491>
- Huang, H., Ji, T., & Xu, X. (2024). An adaptable digital twin model for manufacturing. *Manufacturing Letters*, 41, 1163–1169. <https://doi.org/10.1016/j.mfglet.2024.09.142>
- Institute, K. (2025). Innovative strategies for cost savings in construction. *Kaizen Institute*. (Accessed: 01/04/2025) <https://kaizen.com/insights/cost-savings-strategies-construction/>
- Invicom. (2025). Strain gauge application. *Invicom Test & Measurement*. (Accessed: 01/04/2025) <https://invicom.com/applications/strain-gauge-application/>
- ISO/TC184/SC4. (1984). ISO/TC 184/SC 4 industrial data technical sub-committee. *ISO - International Organization for Standardization*. (Accessed: 01/04/2025) <https://www.iso.org/committee/54158.html>
- ISO/TC184. (1983). ISO/TC 184 automation systems and integration technical committee. *ISO - International Organization for Standardization*. (Accessed: 01/04/2025) <https://www.iso.org/committee/54110.html>
- ISO23247-1. (2021). ISO 23247-1:2021 automation systems and integration — digital twin framework for manufacturing part 1: overview and general principles. *ISO - International Organization for Standardization*. (Accessed: 01/04/2025) <https://www.iso.org/standard/75066.html>
- ISO23247-2. (2021). ISO 23247-2:2021 automation systems and integration — digital twin framework for manufacturing part 2: reference architecture. *ISO - International Organization for Standardization*. (Accessed: 01/04/2025) <https://www.iso.org/standard/78743.html>
- ISO23247-3. (2021). ISO 23247-3:2021 automation systems and integration — digital twin framework for manufacturing part 3: digital representation of manufacturing elements. *ISO - International Organization for Standardization*. (Accessed: 01/04/2025) <https://www.iso.org/standard/78744.html>
- ISO23247-4. (2021). ISO 23247-4:2021 automation systems and integration — digital twin framework for manufacturing part 4: information exchange. *ISO - International Organization for Standardization*. (Accessed: 01/04/2025) <https://www.iso.org/standard/78745.html>

- Jacoby, M., Volz, F., Weißenbacher, C., Stojanovic, L., & Usländer, T. (2021). An approach for industrie 4.0-compliant and data-sovereign digital twins. *At - Automatisierungstechnik*, 69(12), 1051–1061. <https://doi.org/10.1515/auto-2021-0074>
- Ji, T., Polzer, J., & Xu, X. (2023). Cognitive digital twin framework for smart manufacturing. *2023 IEEE 19th International Conference on Automation Science and Engineering (CASE)*, 1–6. <https://doi.org/10.1109/CASE56687.2023.10260436>
- Jones, D., Snider, C., Nassehi, A., Yon, J., & Hicks, B. (2020). Characterizing the digital twin: a systematic literature review. *CIRP Journal of Manufacturing Science and Technology*, 29, 36–52. <https://doi.org/10.1016/j.cirpj.2020.02.002>
- Jung, Y., Lee, H., Jin, Z., Kim, S., & Kim, K. J. (2023). Revamping facility asset management (fam) for smart manufacturing: variables for theory and implementation. *Proceedings of International Structural Engineering and Construction*, 10(1), FAM08-1. [https://doi.org/10.14455/ISEC.2023.10\(1\).FAM-08](https://doi.org/10.14455/ISEC.2023.10(1).FAM-08)
- Kang, M.-S., Lee, D.-H., Bajestani, M. S., Kim, D. B., & Noh, S. Do. (2024). Edge computing-based digital twin framework based on iso 23247 for enhancing data processing capabilities. *Machines*, 13(1), 19. <https://doi.org/10.3390/machines13010019>
- Kibira, D., Shao, G., Venketesh, R., & Triebe, M. J. (2024). Building a digital twin of a cnc machine tool. *2024 Winter Simulation Conference (WSC)*, 2915–2926. <https://doi.org/10.1109/WSC63780.2024.10838819>
- Kibira, D., Shao, G., & Weiss, B. A. (2021). Buiding a digital twin for robot workcell prognostics and health management. *2021 Winter Simulation Conference (WSC)*, 1–12. <https://doi.org/10.1109/WSC52266.2021.9715501>
- Kumaş, E. (2024). International standards for digital twins. *Acta Infologica*, 8(2), 0–0. <https://doi.org/10.26650/acin.1395045>
- Lidell, A., Ericson, S., & Ng, A. H. C. (2022). The current and future challenges for virtual commissioning and digital twins of production lines. In *SPS2022* (pp. 508–519). IOS Press. <https://doi.org/10.3233/ATDE220169>
- López De Ipiña, J. M., Aznar, G., Lopez, A., Olite, J., Koivisto, J., Bartolini, G., & Costa, A. (2021). Digital twins applied to the implementation of safe-by-design strategies in nano-processes for the reduction of airborne emission and occupational exposure to nano-forms. *Journal of Physics: Conference Series*, 1953(1), 012010. <https://doi.org/10.1088/1742-6596/1953/1/012010>
- Madubuike, O. C., Anumba, C. J., & Khallaf, R. (2022). A review of digital twin applications in construction. *Journal of Information Technology in Construction*, 27, 145–172. <https://doi.org/10.36680/j.itcon.2022.008>
- Marinkovic, M., Galka, S., & Meißner, S. (2023). Digital twins for internal transport systems: use cases, functions, and system architecture. *Proceedings of the 56th Hawaii International Conference on System Sciences*, 1195–1204. <https://doi.org/10.24251/HICSS.2023.147>

- Minh, T. B., Do, X. P., Nguyen, Q. H., Nguyen, K. H. V., & Phan, T. T. T. (2024). A digital twin implementation framework for a collaborative robot based on iso 23247. In *The International Conference on Sustainable Energy Technologies* (pp. 747–755). Springer. https://doi.org/10.1007/978-981-97-1868-9_74
- Noga, M., Juhás, M., & Gulán, M. (2022). Hybrid virtual commissioning of a robotic manipulator with machine vision using a single controller. *Sensors*, 22(4), 1621. <https://doi.org/10.3390/s22041621>
- Nooralishahi, P., Ibarra-Castanedo, C., Deane, S., López, F., Pant, S., Genest, M., Avdelidis, N. P., & Maldague, X. P. V. (2021). Drone-based non-destructive inspection of industrial sites: a review and case studies. *Drones*, 5(4), 106. <https://doi.org/10.3390/drones5040106>
- Patton, M. Q. (1980). Qualitative evaluation methods. Sage Publications.
- Plevris, V., & Papazafeiropoulos, G. (2024). AI in structural health monitoring for infrastructure maintenance and safety. *Infrastructures*, 9(12), 225. <https://doi.org/10.3390/infrastructures9120225>
- Pointerra3D. (2025). Collaborative design: how cloud-based digital twins facilitate coordination and communication among aec stakeholders. *Pointerra3D*. (Accessed: 01/04/2025) <https://www.pointerra.com/blog/collaborative-design-how-cloud-based-digital-twins-facilitate-coordination-and-communication-among-aec-stakeholders/>
- Pratap, A., Vi, T.-K., Lee, Y. W., Sardana, N., Hsiung, P.-A., & Kao, Y.-C. (2024). A dt framework integrating human and artificial intelligence for power consumption prediction in cnc machining. *The International Journal of Advanced Manufacturing Technology*, 135(1–2), 915–938. <https://doi.org/10.1007/s00170-024-14477-9>
- Richter, R., Knospe, F., Trapp, M., & Döllner, J. (2024). Dynamic digital twins: challenges, perspectives and practical implementation from a city's perspective (T. H. Kolbe, A. Donaubauer, & C. Beil (eds.); pp. 671–691). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-43699-4_41
- Rite, E. (2025). High-resolution environmental monitoring sensors in civil engineering. *Encardio Rite*. (Accessed: 01/04/2025) <https://www.encardio.com/blog/environmental-monitoring-sensors-civil-engineering>
- Rodrigues, A. M., Formiga, K. T. M., & Milograna, J. (2023). Integrated systems for rainwater harvesting and greywater reuse: a systematic review of urban water management strategies. *Water Supply*, 23(10), 4112–4125. <https://doi.org/10.2166/ws.2023.240>
- Rosova, A., Behun, M., Khouri, S., Cehlar, M., Ferencz, V., & Sofranko, M. (2022). Case study: the simulation modeling to improve the efficiency and performance of production process. *Wireless Networks*, 28(2), 863–872. <https://doi.org/10.1007/s11276-020-02341-z>
- Sacks, R., Brilakis, I., Pikas, E., Xie, H. S., & Girolami, M. (2020). Construction with digital twin information systems. *Data-Centric Engineering*, 1, e14. <https://doi.org/10.1017/dce.2020.16>
- Schwandt, T. A. (1996). Qualitative data analysis: An expanded sourcebook. *Evaluation and Program Planning*, 19(1), 106–107. [https://doi.org/10.1016/0149-7189\(96\)88232-2](https://doi.org/10.1016/0149-7189(96)88232-2)

- Semeraro, C., Lezoche, M., Panetto, H., & Dassisti, M. (2021). Digital twin paradigm: a systematic literature review. *Computers in Industry*, 130, 103469. <https://doi.org/10.1016/j.compind.2021.103469>
- Shao, G. (2021). Use case scenarios for digital twin implementation based on iso 23247. In *National institute of standards: Gaithersburg, MD, USA*. <https://doi.org/10.6028/NIST.AMS.400-2>
- Shao, G. (2024). Manufacturing digital twin standards. *Proceedings of the ACM/IEEE 27th International Conference on Model Driven Engineering Languages and Systems*, 370–377. <https://doi.org/10.1145/3652620.3688250>
- Shao, G., Frechette, S., & Srinivasan, V. (2023). An analysis of the new iso 23247 series of standards on digital twin framework for manufacturing. *Volume 2: Manufacturing Equipment and Automation; Manufacturing Processes; Manufacturing Systems; Nano/Micro/Meso Manufacturing; Quality and Reliability*, 87240, V002T07A001. <https://doi.org/10.1115/MSEC2023-101127>
- Shao, G., Hightower, J., & Schindel, W. (2023). Credibility consideration for digital twins in manufacturing. *Manufacturing Letters*, 35, 24–28. <https://doi.org/10.1016/j.mfglet.2022.11.009>
- Sivák, P., Delyová, I., & Bocko, J. (2023). Comparison of stress concentration factors obtained by different methods. *Applied Sciences*, 13(24), 13328. <https://doi.org/10.3390/app132413328>
- Song, J., & Le Gall, F. (2023). Digital twin standards, open source, and best practices. In *The Digital Twin* (pp. 497–530). Springer International Publishing. https://doi.org/10.1007/978-3-031-21343-4_18
- Spaney, P., Becker, S., Ströbel, R., Fleischer, J., Zenhari, S., Möhring, H.-C., Splettstößer, A.-K., & Wortmann, A. (2023). A model-driven digital twin for manufacturing process adaptation. *2023 ACM/IEEE International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C)*, 465–469. <https://doi.org/10.1109/MODELS-C59198.2023.00081>
- Tao, F., Xiao, B., Qi, Q., Cheng, J., & Ji, P. (2022). Digital twin modeling. *Journal of Manufacturing Systems*, 64, 372–389. <https://doi.org/10.1016/j.jmsy.2022.06.015>
- Tinga, T., & Loendersloot, R. (2019). Physical model-based prognostics and health monitoring to enable predictive maintenance. In E. Lughofer & M. Sayed-Mouchaweh (Eds.), *Predictive Maintenance in Dynamic Systems* (pp. 313–353). Springer International Publishing. https://doi.org/10.1007/978-3-030-05645-2_11
- Twinview. (2023). Digital twins and bim: a synergy for improved stakeholder management. *Twinview*. (Accessed: 01/04/2025) <https://www.twinview.com/insights/digital-twins-and-bim-a-synergy-for-improved-stakeholder-management>
- VanDerHorn, E., & Mahadevan, S. (2021). Digital twin: generalization, characterization and implementation. *Decision Support Systems*, 145, 113524. <https://doi.org/10.1016/j.dss.2021.113524>
- Vanlanduit, S., Sorgente, M., Zadeh, A. R., Güemes, A., & Faisal, N. (2021). Strain monitoring. In M. G. R. Sause & E. Jasiūnienė (Eds.), *Structural Health Monitoring Damage Detection Systems for Aerospace* (pp. 219–241). Springer International Publishing. https://doi.org/10.1007/978-3-030-72192-3_8
- VP, A. (2025). Importance of real-time data integration in digital twins. *Toobler*. (Accessed: 01/04/2025) <https://www.toobler.com/blog/real-time-integration-in-digital-twins>

- Wagg, D. J., Worden, K., Barthorpe, R. J., & Gardner, P. (2020). Digital twins: state-of-the-art and future directions for modeling and simulation in engineering dynamics applications. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part B: Mechanical Engineering*, 6(3), 30901. <https://doi.org/10.1115/1.4046739>
- Wallner, B., Zwölfer, B., Trautner, T., & Bleicher, F. (2023). Digital twin development and operation of a flexible manufacturing cell using iso 23247. *Procedia CIRP*, 120, 1149–1154. <https://doi.org/10.1016/j.procir.2023.09.140>
- Wang, K., Wang, Y., Li, Y., Fan, X., Xiao, S., & Hu, L. (2022). A review of the technology standards for enabling digital twin. *Digital Twin*, 2(1), 4. <https://doi.org/10.12688/digitaltwin.17549.2>
- Wickberg, P., Fattouh, A., Afshar, S., & Bohlin, M. (2023). Adopting a digital twin framework for autonomous machine operation at construction sites. *2023 7th CAA International Conference on Vehicular Control and Intelligence (CVCI)*, 1–6. <https://doi.org/10.1109/CVCI59596.2023.10397254>
- Xia, B., Zheng, G., Zhou, H., Yu, X., Zhao, J., & Diao, Y. (2024). Stability analysis and optimization of concrete column-supported embankments in soft soil. *Acta Geotechnica*, 19(5), 2515–2531. <https://doi.org/10.1007/s11440-024-02302-2>
- Yoo, S. K., Sun, K. J., & Kim, S. H. (2024). Digital twin standardization: trends and future prospects. *Electronics and Telecommunications Trends*, 39(3), 79–86. <https://doi.org/10.22648/ETRI.2024.J.390308>



A Novel Technique for Criteria Weighting in Multi-Criteria Decision-Making: Tanimoto Contrast Approach (TCA)

Furkan Fahri ALTINTAŞ^{1*}

¹ Aydın Provincial Gendarmerie Command, Aydın, Türkiye

Keywords	Abstract
Tanimoto Similarity	This study introduces the Tanimoto Contrast Approach (TCA), a novel objective method for determining criterion weights in Multi-Criteria Decision-Making (MCDM) problems. Built on the internal–external dispersion logic of the CRITIC method, TCA replaces Pearson correlation with Tanimoto similarity to capture both linear and non-linear relationships, enabling a more comprehensive evaluation of inter-criterion contrasts and similarities. The method was tested using the 2024 Global Innovation Index data from selected seven countries. Sensitivity analysis revealed that TCA maintains ranking stability under varying conditions, while comparative analysis showed strong correlation with ENTROPY, SVP, and MEREC methods, confirming its reliability and credibility. In addition, simulation analysis based on ten different decision matrix scenarios demonstrated that TCA produces high average variance and consistent, homogeneous weight distributions evidence of its robustness and stability. TCA’s advantages include distribution free applicability, insensitivity to zero or negative values, scale independence, and effectiveness with large datasets. Moreover, its comparative performance against widely used objective weighting methods such as ENTROPY, CRITIC, SD, SVP, MEREC, and LOPCOW has been thoroughly discussed. In conclusion, TCA offers contrast-based, decision-maker-independent weighting framework that generates meaningful, balanced, and sensitive results. Its integration into MCDM applications provides a valuable contribution to the advancement of objective weighting techniques.
Tanimoto Contrast	
MCDM	
Criteria Weight	

Cite

Altıntaş, F.F., (2025). A Novel Technique for Criteria Weighting in Multi-Criteria Decision-Making: Tanimoto Contrast Approach (TCA). *GU J Sci, Part A, 12(2)*, 445-478. doi:10.54287/guj.1673755

Author ID (ORCID Number)	Article Process
0000-0002-0161-5862	Furkan Fahri ALTINTAŞ
	Submission Date 10.04.2025
	Revision Date 13.05.2025
	Accepted Date 21.05.2025
	Published Date 30.06.2025

1. INTRODUCTION

Criterion weighting is a critical component in Multi-Criteria Decision Making (MCDM), as it defines the relative importance of each criterion, effectively structuring decision-making processes (Ecer, 2020; Thakkar, 2021). Objective weighting methods, rooted in data-driven analysis, mitigate subjectivity and enhance consistency, contrasting with subjective approaches prone to bias (Baş, 2021; Zardari et al., 2014). Thus, advancing innovative objective weighting techniques is vital to accurately model criterion interdependencies and improve decision-making precision (Bircan, 2020; Demir et al., 2021).

In this context, this study introduces a novel analytical framework, the Tanimoto Contrast Approach (TCA), which determines criterion weights by scrutinizing spatial contrasts among criteria and leveraging Tanimoto similarity. The assessment of contrast in TCA adopts a methodology analogous to that of the CRITIC method,

*Corresponding Author, e-mail: furkanfahrialtintas@yahoo.com

albeit with notable divergences. CRITIC is utilized as the foundational framework owing to its established reliability in quantifying criterion contrasts (Ecer, 2020). Furthermore, Tanimoto similarity is integrated into the approach due to its efficacy in accurately and consistently capturing the relationships between variables (Sharma & Lal, 2011; Chung et al., 2019). By incorporating Tanimoto similarity, the proposed methodology facilitates the development of diverse algorithms and models grounded in both distances (contrasts) and similarities (Surendran et al., 2025).

The primary motivation behind the proposed method is to assess sensitivity through sensitivity analysis at an ideal level, to evaluate credible and reliable results through comparative analysis, and to measure robustness and stability through simulation analysis. Another aim of the study is to highlight the advantages of the proposed method. Furthermore, there has been no study in the existing MCDM literature that uses Tanimoto similarity in determining criterion weights. In this context, the proposed method integrates Tanimoto similarity based on the CRITIC technique and is the first study to apply this approach in criterion weighting. This study makes a significant contribution to the MCDM literature, particularly in the areas of criterion weighting and Tanimoto similarity. In the materials and methods section of the study, various objective criterion weighting methods are first explained, followed by a detailed description of how Tanimoto similarity is applied and the advantages of the proposed method. Finally, the findings are discussed, and the results of the study are summarized and concluded.

2. MATERIAL AND METHOD

2.1. Some Objective Criteria Weighting Methods and Their Characteristics

Subjective weight coefficients are frequently determined by the decision-makers' personal experiences and judgments, thus reflecting inherent individual biases. Consequently, these coefficients can vary substantially among different decision-makers (Baş, 2021). Typically, these weights are assigned based on expert opinions. However, relying exclusively on expert evaluations may introduce biases and inaccuracies into the decision-making process (Bircan, 2020). Conversely, objective methods disregard the subjective inconsistencies and uncertainties present in decision-makers' inputs. These methods employ mathematical models and utilize data derived from the decision matrix to ascertain the criterion weights. Essentially, objective weighting techniques prioritize the inherent structural properties of the data during the evaluation process (Demir et al., 2021).

In the realm of MCDM, a variety of objective weighting methodologies are employed, including CRITIC (Criteria Importance Through Inter-criteria Correlation) (Dhara et al., 2022; Momena et al., 2024), ENTROPY (Eligüzcel & Eligüzcel, 2024; Wang et al., 2023), SVP (Statistical Variance Procedure) (Tayalı & Timor, 2017; Nasser et al., 2019), SD (Standard Deviation) (Baydaş, et al., 2024; Mukhametzzyanov, 2021), MEREC (Method Based on Removal Effects of Criteria) (Yazdi et al., 2025; Sehgal et al., 2025), and LOPCOW (Logarithmic Percentage Change-driven Objective Weighting) (Durdu, 2025; Yasin et al., 2025). The CRITIC method, notably, operates on the principle of leveraging the inherent information within a system. A fundamental tenet of CRITIC is that a criterion's significance is directly proportional to its variability or

distinctiveness in relation to other criteria (Diakoulaki et al., 1995). Consequently, CRITIC places a strong emphasis on the interrelationships among criteria. The methodology involves analyzing correlations between criteria to identify any inconsistencies or contradictions (Ulutaş & Topal, 2020). These are subsequently addressed by applying weights derived from the standard deviation, thereby facilitating the computation of appropriate weights for each criterion (Ecer, 2020). The process commences with the construction of a decision matrix, followed by the normalization of its constituent values (Aydemir, 2025). Upon completion of the normalization, the correlations between the criteria are evaluated to determine their respective weights (Sleem et al., 2023).

The ENTROPY methodology offers a structured and impartial framework for ascertaining the relative significance of criteria within decision-making processes. This technique commences with the development of a decision matrix, succeeded by the calculation of normalized values (Ayçin, 2019). Subsequently, the ENTROPY metric is employed to evaluate the degree of indeterminacy or irregularity associated with each criterion, thereby quantifying the informational content it contributes (Dinçer, 2019). This approach utilizes the standardized values and the derived ENTROPY metric. The said metric is determined by the ratio of the negative sum of the product of each criterion's standardized value and the natural logarithm of that standardized value to the natural logarithm of the number of alternatives (Ulutaş & Topal, 2020). This approach augments objectivity in weight assignment, empowering decision-makers to conduct more equitable and evidence-based assessments (Baş, 2021).

The SD technique determines criterion weights by quantifying the dispersion of each criterion's values relative to its arithmetic mean (Öztel & Alp, 2020). Initially, the decision matrix undergoes normalization to achieve data standardization. Subsequently, the standard deviation for each criterion is calculated, and these values are then utilized to derive the corresponding weights (Uludağ & Doğan, 2021).

Conversely, the SVP methodology establishes criterion weights by computing the variance of each criterion based on the data within the decision matrix (Dinçer, 2019). Given the criteria exhibiting higher variance are assigned greater weights, reflecting their amplified influence on the overall decision-making framework (Öztel & Alp, 2020).

The LOPCOW employs a multifaceted approach to ascertain optimal weights, while simultaneously striving to mitigate disparities between the most and least influential criteria (Keleş, 2023). This methodology also accounts for the interrelationships among criteria. Initially, a decision matrix is constructed, and its values are subsequently normalized. Following this, the mean squared value, expressed as a percentage of the standard deviation for each criterion, is computed to minimize the impact of variations in data scale. This process culminates in the calculation of weight coefficients for each criterion (Ecer & Pamucar, 2022).

The MEREC initiates its process, analogous to other weighting methodologies, with the creation and normalization of a decision matrix. Subsequently, the aggregate performance indices of decision alternatives

are calculated utilizing a framework grounded in natural logarithmic functions. Following this, these performance indices are refined by factoring in the contribution of each decision alternative, with iterative recalculations performed employing the natural logarithm (Keshavarz-Ghorabae, et al., 2021). In the concluding stage of the methodology, criterion weight coefficients are derived by quantifying the removal effect of each criterion, specifically through the summation of absolute deviations. Furthermore, as the influence of a criterion on decision alternatives escalates, its corresponding weight coefficient also increases, thereby ensuring a more representative and reflective weighting paradigm (Popović et al., 2022).

2.2. Tanimoto Similarity

Tanimoto similarity is a statistical measure that numerically expresses the similarity between two sets. It is commonly used to assess the overlap and differences between sets, especially in binary data, where the presence or absence of a feature is represented by 1 or 0, respectively (Lipkus, 1999). This measure plays a critical role in various fields such as information retrieval (Zainudin & Nurjana, 2018; Thiel et al., 2014), recommendation systems (Selvi & Sivasankar, 2018; Vivek et al., 2018), bioinformatics (Mulia et al., 2018), and data mining (Anastasiu & Karypis, 2017; Paulose et al., 2018).

Tanimoto similarity can be calculated in two forms: set-based and vector-based. In the set-based Tanimoto similarity, the similarity between two sets is calculated by dividing the number of common elements (the intersection count) by the difference between the union count of the two sets and the intersection count, as shown in Equation 1. In the vector-based Tanimoto similarity, the similarity between two sequences or sets is measured by the ratio of the sum of the products of corresponding quantities in different sets to the difference between the sum of the squares of the individual quantities in each set, as demonstrated in Equation 2 (Bajusz et al., 2015).

$P: \{p_1, p_2, p_3 \dots p_n\}$: First cluster, $Q: \{q_1, q_2, q_3 \dots q_n\}$: Second cluster, $P \cap Q$: The intersection of sets P and Q , $|P|$: The cardinality of set P , $|Q|$: The cardinality of set Q .

$$T(P, Q) = \frac{P \cap Q}{|P| + |Q| - P \cap Q} \quad (1)$$

$$T(P, Q) = \frac{P \cdot Q}{\|P\|^2 + \|Q\|^2 - P \cdot Q} = \frac{\sum_{i=1}^n (P_i * Q_i)}{\left(\sqrt{\sum_{i=1}^n (P_i)^2}\right)^2 + \left(\sqrt{\sum_{i=1}^n (Q_i)^2}\right)^2 - \sum_{i=1}^n (P_i * Q_i)} \quad (2)$$

Tanimoto similarity metric offers several advantages over other similarity methods. First, Tanimoto similarity allows for the evaluation of common components between data points not only in terms of directional similarity but also considering magnitude differences (Zhang et al., 2015). For example, while cosine similarity measures the similarity between two vectors based on the angle between them (Mana & Sasipraba, 2021). Tanimoto similarity provides a more comprehensive comparison by considering both the intersection and the

union of sets (Surendran et al., 2025). This characteristic ensures more accurate and reliable measurements, particularly in large datasets.

Another significant advantage of Tanimoto similarity is its ability to work with both binary data and vectors containing continuous variables (Bero et al., 2018). While methods like cosine similarity typically perform better with continuous vectors, and Jaccard similarity is mainly useful for set-based analysis (Zahrotun, 2016), Tanimoto similarity combines the strengths of both methods, making it applicable to both categorical and numerical data. In contrast to cosine similarity, which focuses only on the direction of vectors and disregards magnitude differences, Tanimoto similarity accounts for both direction and magnitude, offering a more holistic similarity analysis. This is especially beneficial when working with weighted vectors, as Tanimoto similarity allows for more accurate evaluations (Kryszkiewicz, 2013). Compared to Jaccard similarity, Tanimoto similarity is particularly stronger when working with continuous variables (Sankara et al., 2011). While Jaccard measures the proportion of common elements between sets, relying solely on presence/absence information, it can fall short in vector-based calculations (Kryszkiewicz, 2013). Tanimoto similarity, however, is compatible with numerical data and works effectively with both binary and continuous variables, giving it a broader range of applicability.

Euclidean distance, which measures the direct distance between two data points, is sensitive to different measurement scales (Dokmanic et al., 2015) and may not provide meaningful results in high-dimensional datasets. Tanimoto similarity, however, focuses on the density of shared components between datasets or vectors rather than direct distance, making it a more reliable comparison method in high-dimensional data (Lasek & Mei, 2019). It can produce more meaningful results, especially in large data analysis and compound similarity assessments, when compared to Euclidean distance.

Mahalanobis distance, which takes correlations between variables into account when calculating similarity (Iglesias & Kastner, 2013) requires high computational costs and may not be efficient with large datasets. Tanimoto similarity, on the other hand, offers a more scalable approach as it is based directly on the intersection and union of components, providing a faster and more effective similarity measure, especially in large-scale machine learning and bioinformatics applications (Baldi & Benz, 2008).

Minkowski distance, a general form of Euclidean distance that can be customized with different parameters (Xu et al., 2019), still compares vector magnitudes directly and does not consider the intersection-union relationship. Tanimoto similarity, by considering both the distance between data points and the degree of overlap, provides more reliable results in object-based similarity analysis. Additionally, Tanimoto similarity offers scalability advantages when applied to large datasets, enabling efficient computations. The intersection and union-based calculation method ensures that similarity analysis can be performed without accuracy loss, even in datasets with numerous variables. Furthermore, Tanimoto similarity is more robust to noise. While other similarity metrics may show significant differences when there are minor deviations or missing data in a

dataset, Tanimoto similarity provides a more stable measure, making it a valuable tool in biological data analysis (Mellor et al., 2019), chemical compound comparisons (Ma et al., 2011) and document similarity calculations. In conclusion, Tanimoto similarity stands out as a more flexible, comprehensive, and reliable method compared to other similarity metrics. Its ability to work with both binary and continuous data, scalability, robustness to noise, and precise measurement based on intersection-union ratios make it an effective analytical tool across various scientific fields.

A review of the literature reveals numerous studies focusing on Tanimoto similarity. Mohebbi et al. (2022) utilized Tanimoto similarity in the QSAR modeling of a ligand-based pharmacophore derived from Hepatitis B virus surface antigen inhibitors. Ying et al. (2021) applied Tanimoto similarity in improving the chemical structure elucidation process using Morgan Fingerprints. Kryszkiewicz (2021) employed Tanimoto similarity to identify similarity neighborhoods of real-valued vectors through triangle inequalities and length boundaries. Shan et al. (2025) evaluated ionic liquid toxicity using machine learning and structural similarity methods with Tanimoto similarity. Feldmann and Bajorath (2022) used Tanimoto similarity to calculate exact Shapley values for support vector machines. Yoon and Lee (2025) leveraged Tanimoto similarity in silico screening to discover new compounds for focal adhesion kinase activation using virtual screening, AI-based prediction, and molecular dynamics. Ahmad et al. (2025) applied Tanimoto similarity in developing a data mining and machine learning-based model for optimal materials for perovskite and organic solar cells. Yoshizawa et al. (2025) utilized Tanimoto similarity in creating a data-driven generative strategy to avoid reward manipulation in multi-target molecular design. Nowatzky et al. (2025) applied Tanimoto similarity for local neighborhood-based prediction of compound mass spectrometric data from a single fragmentation event. Feldmann et al. (2025) used Tanimoto similarity in the uncertainty analysis of neural fingerprint-based models.

2.3. Proposed Method (Tanimoto Contrast Approach-TCA)

In the determination of criteria weights, the distinctiveness and contrasts among criteria are fundamental in defining their inherent characteristics. The proposed method shares a conceptual similarity with the CRITIC method in its underlying logic. The CRITIC approach posits that a criterion's importance is directly proportional to its degree of contrast. Consequently, a criterion's significance increases with greater standard deviation and lower correlation with other criteria, indicating a higher degree of contrast (Ecer, 2020).

The proposed method emphasizes the uniqueness and oppositional nature of criteria, highlighting their distinguishing attributes. It employs Tanimoto similarity values to quantify the degree of contrast between criteria. Tanimoto similarity assesses the numerical proximity of criteria within a vector space. According to this principle, smaller differences within a criterion's numerical series imply closer proximity to other criteria in the vector space, resulting in a lower degree of contrast series (Bajusz et al., 2015). Conversely, the CRITIC method utilizes Pearson correlation coefficients, which rely on the proportional relationship between criteria's numerical series (Diakoulaki et al., 1995). For example, if a criterion with a sequence of relatively small values

exhibits a similar progression to another criterion with larger values, a significant and positive correlation is inferred.

The CRITIC method determines each criterion's weight by multiplying its internal and external distributions. The internal distribution, representing a criterion's variability, is quantified by its standard deviation. The external distribution, reflecting a criterion's divergence from others, is measured by summing the differences between one and the Pearson correlation coefficients between the given criterion and all others (Diakoulaki et al., 1995). The fundamental formulation of the CRITIC weighting mechanism is expressed in Equations 3, 4, and 5 (Ayçin, 2019; Dinçer, 2019).

r_{ij} : Correlation coefficient between the j – th and j – th criteria

σ_j : Standard deviation of the j – th criterion ($j = 1, 2, \dots, n$)

w_j : Weight of the j – th evaluation criterion ($j = 1, 2, \dots, n$)

p_{jk} : It indicates the correlation coefficient between any criterion j and criterion k .

$$p_{jk} = \frac{\sum_{i=1}^m (r_{ij} - \bar{r}_j) \cdot (r_{ik} - \bar{r}_k)}{\sqrt{\sum_{i=1}^m (r_{ij} - \bar{r}_j)^2 \cdot (r_{ik} - \bar{r}_k)^2}} \quad j, k=1, 2, \dots, n \quad (3)$$

$$\sigma_j = \sqrt{\frac{\sum_{i=1}^m (r_{ij} - \bar{r}_j)^2}{m-1}} \quad (4)$$

$$C_j = \sigma_j \cdot \sum_{i=1}^m (1 - r_{ij}) \quad j=1, 2, \dots, m \quad (5)$$

As shown in Equation 3, term σ_j represents the internal contrast of each criterion, reflecting the dissimilarity among its own values, while term $\sum_{i=1}^m (1 - r_{ij})$ represents the external contrast, indicating the dissimilarity of each criterion relative to all others. The product of σ_j and $\sum_{i=1}^m (1 - r_{ij})$ determines the spatial dispersion of each criterion, thereby establishing its weight (Diakoulaki et al., 1995). The proposed method employs a weighting mechanism conceptually aligned with the CRITIC approach. Specifically, internal contrast in the proposed method is measured by σ_j , and external contrast, representing dissimilarity with other criteria, is quantified by $\sum_{i=1}^m (1 - r_{ij})$ (Diakoulaki et al., 1995), consistent with the CRITIC method's logic. Therefore, in the proposed method, each criterion's weight is calculated by multiplying its internal distributions (σ_j), with its external dispersion $[\sum_{i=1}^m (1 - r_{ij})]$, derived from the dissimilarity or distance based on Tanimoto similarity values (i.e., the external contrast among the criteria's own value distributions). The model illustrating the implementation steps of the proposed method within this scope is presented in Figure 1, while the corresponding mathematical procedures are meticulously detailed below.

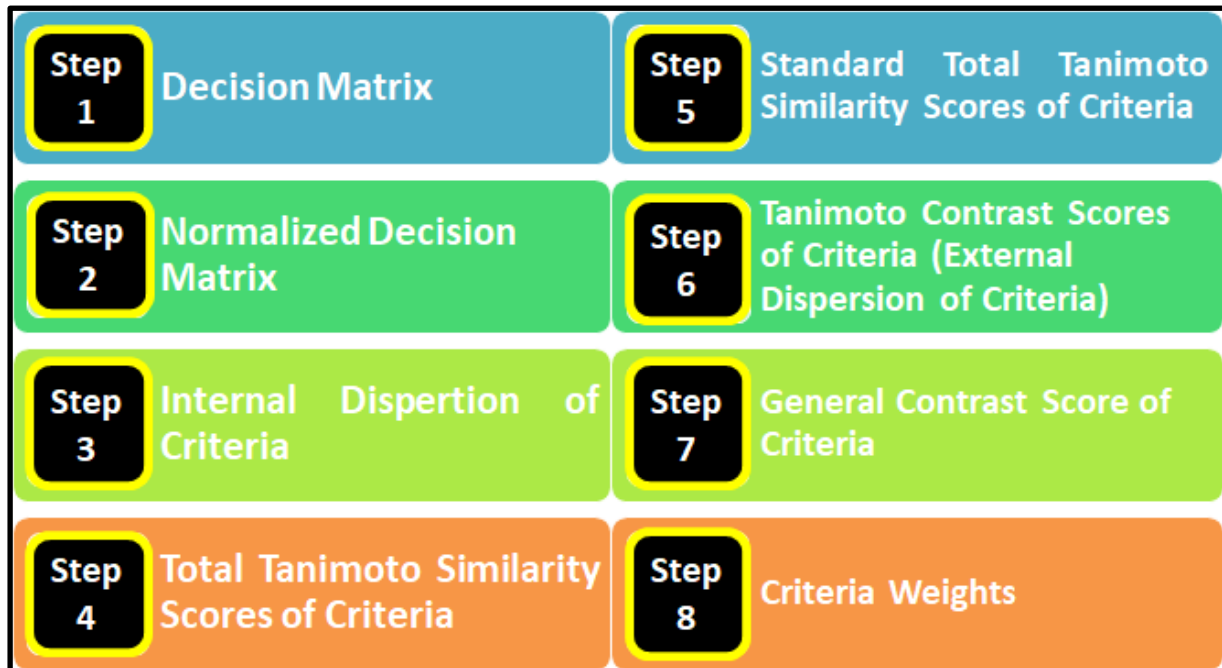


Figure 1. TCA Model

Step 1. Obtaining Decision Matrix (X)

In the first step of the proposed method, the decision matrix is constructed using Equation 6, similar to the CRITIC method.

$$DM = [d_{ij}]_{m \times n} = \begin{matrix} & \begin{matrix} ALT_1 & ALT_2 & \dots & ALT_m \end{matrix} \\ \begin{matrix} CRT_1 \\ CRT_2 \\ \dots \\ CRT_n \end{matrix} & \begin{bmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \vdots & \vdots & \dots & \vdots \\ d_{m1} & d_{m2} & \dots & d_{mn} \end{bmatrix} \end{matrix} \quad (6)$$

Step 2: Calculation of Normalisation Values of Decision Matrix (d_{ij}^*)

In the second step of the method, the normalization of the decision matrix values is achieved using Equation 7 for benefit-oriented criteria and Equation 8 for cost-oriented criteria. Subsequently, the normalized decision matrix is constructed using Equation 9, based on the direction of the criteria.

For benefit-oriented criteria:

$$d_{ij}^* = \frac{\min(d_{ij})}{d_{ij}} \quad (7)$$

For cost-oriented criteria:

$$d_{ij}^* = \frac{d_{ij}}{\max(d_{ij})} \quad (8)$$

Normalized matrix:

$$DM^* = [d_{ij}]_{m \times n}^* = \begin{matrix} & \begin{matrix} ALT \\ ALT_1 \\ ALT_2 \\ \vdots \\ ALT_m \end{matrix} & \begin{bmatrix} CRT_1 & CRT_2 & \dots & CRT_n \\ d_{11}^* & d_{12}^* & \dots & d_{1n}^* \\ d_{21}^* & d_{22}^* & \dots & d_{2n}^* \\ \vdots & \vdots & \dots & \vdots \\ d_{m1}^* & d_{m2}^* & \dots & d_{mn}^* \end{bmatrix} \end{matrix} \quad (9)$$

Step 3: Calculation of Internal Distribution $(ID)_j$ of Criteria

In this step, the standard deviation $(SD)_j$ values of each criterion are calculated using Equation 10 to measure the distribution of values within each criterion.

$$(ID)_j = (SD)_j = \sqrt{\frac{\sum_{i=1}^m (d_{ij}^* - \bar{d}_{.j})^2}{m-1}} \quad (10)$$

Step 4: Calculation of Total Tanimoto Similarity Scores of Criteria $(TTSS)_j$

In the third step, the Tanimoto Similarity Scores $((TSS)_{TAN_{CRT_j \rightarrow CRT_j}})$ of each criterion with respect to the other criteria are first calculated using the equations provided below. These calculations are performed based on the normalized values and are determined using Equations 11, 12, 13, 14, and 15.

$$CRT \in \{CRT_1, CRT_2, CRT_3, \dots, CRT_n\} \quad (11)$$

(1) for CRT_1

$$(TAN_{CRT_1 \rightarrow CRT_2}) = \frac{CRT_1 \cdot CRT_2}{\|CRT_1\| \cdot \|CRT_2\| - (CRT_1 \cdot CRT_2)} \quad (12)$$

$$\vdots$$

$$(TAN_{CRT_1 \rightarrow CRT_n}) = \frac{CRT_1 \cdot CRT_n}{\|CRT_1\| \cdot \|CRT_n\| - (CRT_1 \cdot CRT_n)} \quad (13)$$

(n) for CRT_n

$$(TAN_{CRT_n \rightarrow CRT_1}) = \frac{CRT_n \cdot CRT_1}{\|CRT_n\| \cdot \|CRT_1\| - (CRT_n \cdot CRT_1)} \quad (14)$$

$$\vdots$$

$$(TAN_{CRT_n \rightarrow CRT_{n-1}}) = \frac{CRT_n \cdot CRT_{n-1}}{\|CRT_n\| \cdot \|CRT_{n-1}\| - (CRT_n \cdot CRT_{n-1})} \quad (15)$$

Secondly, the Total Tanimoto Similarity Scores of Criteria $(TTSS)_j$ for each criterion are calculated using Equations 16 and 17.

(1) for CRT_1

$$(TTSS)_{CRT_1} = \sum_{j=1}^{(m-1)} [(TAN_{CRT_1 \rightarrow CRT_2}) + \dots + (TAN_{CRT_1 \rightarrow CRT_n})] \quad (16)$$

$\vdots$ $\vdots$ $\vdots$
 $\vdots$ $\vdots$ $\vdots$
 $\vdots$ $\vdots$ $\vdots$

(n) for CRT_n

$$(TTSS)_{CRT_n} = \sum_{j=1}^{(m-1)} [(TAN_{CRT_n \rightarrow CRT_1}) + \dots + (TAN_{CRT_n \rightarrow CRT_{n-1}})] \quad (17)$$

Step 5: Calculation of Standard Total Tanimoto Similarity Scores of Criteria $(STTSS)_j$

In this step, the $(TTSS)_j$ values of each criterion are standardized to the [0,1] range using Equation 18.

$$(STTSS)_j = \frac{(TTSS)_j}{\sum_{j=1}^n (TTSS)_j} \quad (18)$$

Step 6: Calculation of Tanimoto Contrast Scores of Criteria $(TCS)_j$ (External Dispersion of Criteria: ED_j)

In this step, the $(STTSS)_j$ of the criteria are converted into distance (contrast) values using the method proposed by Todeschini et al. (2007) and formulated in Equation 19. Accordingly, the relationship between similarity and distance is explained in Equation 20.

$$\text{Similarity } (S) = \frac{1}{1 + \text{Distance } (Contrast)} = \text{Distance } (Contrast) = \frac{1 - S}{S} \quad (19)$$

$$(TCS)_j = \frac{1 - (STTSS)_j}{(STTSS)_j} \quad (20)$$

The method presented in Equation 20 is not confined solely to the [0, 1] interval for similarity and distance measures but is also applicable to all positive similarity and distance metrics beyond this range (Bajusz et al., 2015). This inherent flexibility enables the proposed method to effectively transform similarity and distance values into a standardized [0, 1] scale, establishing it as a robust and comprehensive metric for similarity assessment. In the existing literature, conventional similarity and distance measures are typically restricted to the [0, 1] interval, with transformations predominantly achieved by subtracting the obtained value from 1 (Podani, 2000). However, this conventional approach can introduce metric inconsistencies, particularly in similarity assessments, thereby potentially compromising measurement accuracy. In this context, the

transformation framework outlined in Equation 20 transcends the limitations of traditional methods based solely on the subtraction of 1. The proposed method not only addresses measurements within the $[0, 1]$ range but also effectively accommodates a broader spectrum of positive values. Consequently, it provides a more inclusive and integrative perspective on the similarity-distance transformation process. Therefore, in comparison to conventional approaches, the proposed method not only encompasses a wider range of measurements but also facilitates a more accurate and coherent evaluation of the similarity-distance relationship. In this respect, the transformation methodology articulated in Equation 20 not only distinguishes itself from existing methods in the literature but also establishes a more comprehensive and theoretically sound foundation for similarity measurement.

Step 7: Calculation of General Contrast Score of Criteria (GCS_j)

In this step, the Standard Deviation $(SD)_j$ value of each criterion, measured using Equation 10, is multiplied by the TCS_j value presented in Equation 21 to calculate the General Contrast Score (GCS_j) or External Dispersion values for each criterion. Similar to the weighting logic in the CRITIC method, in Equation 21, SD represents the standard deviation of the self-normalized values of each criterion (i.e., the internal distribution of each criterion's own values), while GCS_j , again as in the CRITIC method, indicates the contrast intensity $\sum_{i=1}^m (1-r_{ij})$ of each criterion with respect to the other criteria. Within this scope, similar to the CRITIC method, the overall contrast value of the criteria is measured by multiplying the SD value of each criterion with the total contrast value of that criterion in relation to the other criteria.

$$GCS_j = (SD)_j * (TCS_j) \quad (21)$$

Step 8: Calculation of Criteria Weightings (w_j)

In the final step of the method, the weights of the criteria are calculated by determining the ratio of each criterion's GCS_j value to the total GCS_j value, as given in Equation 22.

$$w_j = \frac{GCS_j}{\sum_{j=1}^m GCS_j} \quad (22)$$

The proposed criterion weighting method offers an objective, data-driven, and mathematically grounded approach, providing significant advantages in decision-making processes. Its primary strength lies in the ability to determine criterion weights independently of the decision-maker's subjective judgments. As a result, the method can generate consistent and reliable outcomes across various applications. The proposed approach optimizes the weighting process by considering both similarities and dissimilarities among criteria. In this context, the Tanimoto similarity plays a pivotal role by measuring the degree of similarity and contrast between criteria. This prevents the excessive weighting of criteria with high uniqueness or dissimilarity, thereby

ensuring a more balanced distribution of weights in terms of informational value. This balanced weighting contributes to a more meaningful and interpretable structure within the decision-making framework. The advantage provided by the Tanimoto similarity enables the method to produce more sensitive and robust results, particularly in MCDM problems. Furthermore, the scale-independence of the method allows for the normalization of criteria with different units of measurement, facilitating comparability and enhancing reliability across diverse data structures. This contributes to a more robust and accurate decision-making process. In addition, the high computational efficiency of the proposed method enables rapid and effective implementation, even on large datasets. Since the proposed method is conceptually based on the weighting logic of the CRITIC method, it has the potential to produce highly correlated results with other widely used objective weighting methods in the literature. This reinforces the reliability and scientific validity of the method. Moreover, its unique structure, which takes inter-criteria relationships into account, offers a novel perspective compared to existing methods and contributes meaningfully to the academic literature. In conclusion, when evaluated within a scientific and systematic framework, the proposed criterion weighting method can be regarded as a valuable approach that enhances objectivity, consistency, and accuracy in decision-making processes. By utilizing mathematical metrics such as Tanimoto similarity to analyze inter-criterion relationships, the method is considered to yield more realistic and reliable results in decision support systems.

Although the proposed criterion weighting method offers several advantages, it also presents certain limitations that may affect its applicability or pose challenges under specific conditions. First, the computational process of the method is relatively complex and requires more mathematical operations compared to traditional approaches. The integration of Tanimoto similarity and other associated calculations can lead to high computational costs, particularly when dealing with complex-scale datasets. This may render the method more time-consuming to implement and necessitate additional computational resources in big data environments. Second, the performance of the method can vary depending on the normalization techniques employed and the inherent characteristics of the dataset. For instance, in cases where an extremely high degree of similarity exists among the criteria, the discriminative power of Tanimoto similarity may diminish. As a result, the method may struggle to accurately reflect the relative importance of such criteria, potentially affecting the reliability of the weighting process. This limitation is particularly relevant in datasets where high correlations exist among multiple criteria, which can reduce the sensitivity of the method.

The proposed TCA method offers several advantages when compared to the widely used CRITIC method. One major limitation of the CRITIC approach lies in its reliance on the Pearson correlation coefficient to capture the contrast among criteria (Ayçin, 2019). The application of this coefficient is constrained by its underlying assumption of normally distributed data (Kilmen, 2015). If the dataset does not meet this assumption, the resulting correlation values may be misleading or incorrect (Kalaycı, 2014). In contrast, the Tanimoto Similarity employed in the TCA method is not subject to such restrictions. It can be effectively applied to

datasets regardless of whether they follow a normal distribution, thus providing broader applicability. Secondly, while the Pearson correlation coefficient only measures linear relationships between variables ignoring potential nonlinear patterns (Karagöz, 2014; Özdamar, 2018). However, Tanimoto Similarity offers a more generalized similarity measure that accounts for both linear and nonlinear structures. This characteristic makes the TCA method more flexible and suitable for a wider range of applications, particularly in complex MCDM contexts. However, it is worth noting that the CRITIC method may yield more reliable results in cases where the relationships between criteria are strongly linear. In such scenarios, the use of TCA might introduce a risk of less accurate outcomes due to its sensitivity to complex or nonlinear patterns, whereas CRITIC maintains robustness due to its strong linear correlation-based structure (Diakoulaki et al., 1995). Moreover, the proposed TCA method shares conceptual similarities with the CRITIC approach in terms of its weighting logic. Both methods consider the contrast within individual criteria (internal dispersion) and the contrast between different criteria (external dispersion) in calculating the final criterion weights. This parallel highlights the methodological compatibility of TCA with established objective weighting approaches, while also emphasizing its unique advantages in terms of distributional flexibility and nonlinear relationship handling.

The ENTROPY method is inherently sensitive to negative and zero values due to its reliance on logarithmic calculations. To mitigate this issue, Zhang et al. (2014) proposed the use of Z-Score normalization to convert the dataset into positive values before applying the method. However, this pre-processing step introduces additional computational costs. In contrast, the proposed TCA method is not affected by zero or negative values, thus eliminating the need for such pre-processing and enhancing its computational efficiency. From a methodological perspective, the ENTROPY method evaluates each criterion based solely on its own entropy values, without considering the interrelationships or contrasts with other criteria (Ayçin, 2019). This lack of external interaction may limit its representational comprehensiveness. On the other hand, the TCA method incorporates both internal dispersion (measured via standard deviation) and external contrast between criteria (captured through Tanimoto Similarity). As a result, the TCA method offers a broader analytical scope compared to the ENTROPY method. Similarly, the SD and SVP methods also calculate criterion weights using only the intrinsic statistical properties of each individual criterion standard deviation for SD and variance for SVP (Demir et al., 2021). These approaches, while straightforward, do not account for inter-criterion contrasts. In this regard, the TCA method demonstrates a more comprehensive structure by integrating both intra-criterion and inter-criterion variability. Nevertheless, ENTROPY, SD, and SVP methods maintain an advantage in terms of computational simplicity (Ecer, 2020). Their ease of implementation makes them more practical for certain applications, and the risk of error during the weighting process tends to be lower compared to more complex approaches like TCA. Thus, while TCA offers greater analytical depth, traditional methods may be preferable in contexts where simplicity and efficiency are prioritized.

The LOPCOW method determines criterion weights based on logarithmic computations applied separately to each criterion, aiming to analyze the differences among them (Keleş, 2023). In contrast, the proposed TCA

method considers both the similarity and contrast relationships between criteria, enabling a more comprehensive evaluation of their mutual interactions. However, due to its logarithmic formulation, the LOPCOW method is particularly sensitive to subtle differences among criteria (Ecer & Pamucar, 2022). As a result, it can more effectively distinguish between criteria with closely spaced values compared to the TCA method. This sensitivity allows LOPCOW to offer a finer differentiation in weight assignment when the criteria exhibit minimal variation. Therefore, while TCA provides a broader analysis by incorporating inter-criterion relationships, LOPCOW may outperform it in scenarios where a more granular distinction between closely related criteria is essential (Ecer & Pamucar, 2022).

The MEREC method determines criterion weights by analyzing the impact caused by the removal of each individual criterion. In other words, it considers the deviation in overall performance resulting from the exclusion of a specific criterion and computes this impact using a logarithmic function (Keshavarz-Ghorabae et al., 2021). In contrast, the proposed TCA method does not rely on the removal effect but instead adopts a broader framework by integrating similarity (Tanimoto similarity) and contrast (Total Contrast Score) analyses to assess inter-criterion relationships. Furthermore, because MEREC evaluates the information loss due to the removal of a criterion, it tends to focus on linear dependencies. On the other hand, TCA captures both similarity and contrast simultaneously, enabling it to better model non-linear and more complex relationships among criteria. Additionally, similar to the ENTROPY method, MEREC is sensitive to zero and negative values due to its reliance on logarithmic calculations (Keshavarz-Ghorabae et al., 2021). Nevertheless, MEREC offers a powerful mechanism for identifying critical criteria. If the removal of a criterion leads to significant changes in the system's performance, MEREC effectively highlights its importance [32]. By contrast, since TCA focuses on general similarity and contrast values, it may not emphasize the importance of such critical criteria as clearly. Finally, because MEREC directly demonstrates the change in the system when a criterion is removed, it provides a more intuitive and interpretable structure for decision-makers [32]. In contrast, the similarity and contrast-based approach of TCA may present interpretative challenges, as the underlying computations are less straightforward and may require additional explanation to fully understand the resulting weights.

2.4. Data Set

This study proposes a method based on Tanimoto similarity to evaluate the weights of criteria within the scope of decision-making problems. The dataset consists of the 2024 Global Innovation Index (GII) criterion values for seven selected countries with varying performance levels. These countries were chosen to ensure that the criterion values do not dominate the performance outcomes and that the differences among alternatives are not excessively large. Accordingly, there are no dominant values for any criterion within the dataset. In this context, the proposed method aims to reveal an ideal differentiation in the weights of the criteria. For clarity, the abbreviations of the countries and the GII criteria are explained in Table 1.

Table 1. Data Set

GII Criteria	Abbreviations
Institutions	CRT1
Human Capital and Research	CRT2
Infrastructure	CRT3
Market Sophistication	CRT4
Business Sophistication	CRT5
Knowledge and Technology Outputs	CRT6
Creative Outputs	CRT7
Countries/Alternatives (A)	Abbreviations
Hungary	ALT1
Türkiye	ALT2
Bulgaria	ALT3
India	ALT4
Poland	ALT5
Thailand	ALT6
Latvia	ALT7

3. RESULTS AND DISCUSSION

3.1. Computational Analysis

In the first step of the proposed method, the decision matrix was constructed using Equation 6, and in the second step, the normalized decision matrix was generated based on Equations 7 and 9. Accordingly, the calculated values are presented in Table 2.

In the third stage of the method, according to Equation 10, the standard deviations of each criterion based on their normalized values were calculated. Accordingly, the internal dispersion $(ID)_j$ of each criterion defined as the contrast among the values within the same criterion was determined through the measurement of these standard deviation values. In the fourth stage of the method, under the first case, the Tanimoto Similarity Scores $\left((TSS)_{TAN_{CRT_j \rightarrow CRT_j}}\right)$ between each criterion and the others were calculated using Equations 11, 12, 13, 14, and 15. In the second case, the Total Tanimoto Similarity Scores $(TTSS)_j$ of each criterion were computed using Equations 16 and 17. In the fifth stage, the $TTSS_j$ values of each criterion were standardized using Equation 18, resulting in the Standardized Tanimoto Similarity Scores $(STTSS)_j$. In the sixth stage, the Tanimoto Contrast Scores $(TCS)_j$ or External Dispersion values $(ED)_j$ of each criterion were calculated using

Equations 19 and 20. Subsequently, the General Contrast Scores $(GCS)_j$ for each criterion were computed using Equation 21. Finally, in the last stage of the method, the final criterion weights $(w)_j$ were derived using Equation 22. The calculated values for all these steps are presented in Table 3.

Table 2. Decision and Normalized Decision Matrix

Decision Matrix							
G7	CRT1	CRT2	CRT3	CRT4	CRT5	CRT6	CRT7
ALT1	52,2	42,9	51	34,1	46,3	35,6	32,1
ALT2	33,3	40	50,2	43,4	31,1	28,6	48,3
ALT3	41,8	32,3	54,4	37,7	32,1	31,7	42,9
ALT4	51,5	34,8	39	52,3	28,1	38,8	32,1
ALT5	44,9	42,6	45,8	33,6	38	28	38,1
ALT6	44,8	30,7	45,8	50,6	35,4	29,8	34,9
ALT7	57,9	39,2	51,3	36,6	35,9	24,2	32,8
Normalized Decision Matrix							
G7	CRT1	CRT2	CRT3	CRT4	CRT5	CRT6	CRT7
ALT1	0,638	0,716	0,765	0,985	0,607	0,680	1,000
ALT2	1,000	0,768	0,777	0,774	0,904	0,846	0,665
ALT3	0,797	0,950	0,717	0,891	0,875	0,763	0,748
ALT4	0,647	0,882	1,000	0,642	1,000	0,624	1,000
ALT5	0,742	0,721	0,852	1,000	0,739	0,864	0,843
ALT6	0,743	1,000	0,852	0,664	0,794	0,812	0,920
ALT7	0,575	0,783	0,760	0,918	0,783	1,000	0,979

According to the values presented in the table, the CRT4 criterion was identified as the most significant one, with a weight of 0.334. This was followed by CRT1 (0.127), CRT7 (0.123), CRT5 (0.114), CRT6 (0.112), CRT2 (0.101), and CRT3 (0.090), respectively. The incorporation of three distinct metrics (*ID*, *ED*, and *GCS*) in the determination of criterion weights demonstrates that the TCA method offers a balanced approach in the weighting process. In particular, the fact that CRT4 received the highest weight indicates that this criterion plays a pivotal role in the decision-making process within the TSA framework. Conversely, CRT3, having the lowest weight of 0.090, is shown to be relatively less influential in the evaluation. Within this context, the ranking of criteria based on the weights obtained through the TSA method is as follows: CRT4 > CRT1 > CRT7 > CRT5 > CRT6 > CRT2 > CRT3. This ranking, derived from computations performed on the decision matrix, provides a significant reference point for decision-makers in assessing the relative importance of the criteria involved. The mathematical calculation of the weight value of the CRT1 criterion, based on the proposed method, is presented below.

Table 3. $SD_j = ID_j, TTSS_j, STTSS_j, TCS_j = ED_j, GCS_j$, and w_j Scores of Criteria

	ID	ED			GCS_j	w_j	Rank
CRT	$SD_j = ID_j$	$TTSS_j$	$STTSS_j$	$TCS_j = ED_j$			
CRT1	0,140	5,501	0,145	5,888	0,822	0,127	2
CRT2	0,113	5,614	0,148	5,750	0,653	0,101	6
CRT3	0,094	5,262	0,139	6,202	0,585	0,090	7
CRT4	0,328	4,994	0,132	6,589	2,161	0,334	1
CRT5	0,127	5,554	0,147	5,823	0,737	0,114	4
CRT6	0,125	5,559	0,147	5,817	0,725	0,112	5
CRT7	0,132	5,412	0,143	6,002	0,795	0,123	3
Sum		37,895	Sum		6,479	----	-----

Step 2. Normalized Matrix (D^*)

Equation 7:

$$\frac{33,3}{52,2} = 0,638$$

Step 3. Internal distribution (SD Score) of Criteria $[(SD)_{CTR1} = (ID)_{CTR1}]$

Equation 10:

$$(SD)_{CTR1} = (ID)_{CTR1} = \text{Equation 4} = \frac{[(0,638 - 0,734)^2 + \dots + (0,575 - 0,734)^2]}{7 - 1} = 0,140$$

Step 4: Calculation of Total Tanimoto Similarity Scores of CTR1 $(TTSS)_{CTR1}$ Condition 1: Tanimoto Similarity Scores of CTR1 with Other Criteria $(TAN_{CTR1 \rightarrow CTRj})$

Equation 11 – 15:

$$TAN_{CTR1 \rightarrow CTR2} = \frac{[(0,638 * 0,716) + \dots + (0,575 * 0,783)]}{[(0,638)^2 + \dots + (0,575)^2] + [(0,716)^2 + \dots + (0,783)^2] - [(0,638 * 0,716) + \dots + (0,575 * 0,783)]} = 0,945$$

$$TAN_{CTR1 \rightarrow CTR3} = \frac{[(0,638 * 0,765) + \dots + (0,575 * 0,760)]}{[(0,638)^2 + \dots + (0,575)^2] + [(0,765)^2 + \dots + (0,760)^2] - [(0,638 * 0,765) + \dots + (0,575 * 0,760)]} = 0,943$$

$$TAN_{CTR1 \rightarrow CTR4} = \frac{[(0,638 * 0,985) + \dots + (0,575 * 0,918)]}{[(0,638)^2 + \dots + (0,575)^2] + [(0,985)^2 + \dots + (0,918)^2] - [(0,638 * 0,985) + \dots + (0,575 * 0,918)]} = 0,824$$

$$TAN_{CTR1 \rightarrow CTR5} = \frac{[(0,638*0,607)+\dots+(0,575*0,783)]}{[(0,638)^2+\dots+(0,575)^2]+[(0,607)^2+\dots+(0,783)^2]-[(0,638*0,607)+\dots+(0,575*0,783)]} = 0,958$$

$$TAN_{CTR1 \rightarrow CTR6} = \frac{[(0,638*0,680)+\dots+(0,575*1)]}{[(0,638)^2+\dots+(0,575)^2]+[(0,680)^2+\dots+(1)^2]-[(0,638*0,680)+\dots+(0,575*1)]} = 0,948$$

$$TAN_{CTR1 \rightarrow CTR7} = \frac{[(0,638*1)+\dots+(0,575*0,979)]}{[(0,638)^2+\dots+(0,575)^2]+[(1)^2+\dots+(0,979)^2]-[(0,638*1)+\dots+(0,575*0,979)]} = 0,885$$

Condition 2: Total Tanimoto Similarity Scores of CTR1 ($TTSS_{(CTR1)}$)

Equation 16 – 17:

$$(TTSS_{(CTR1)}) = 0,945 + 0,943 + 0,824 + 0,958 + 0,948 + 0,885 = 5,501$$

Step 5: Calculation of Standard Total Tanimoto Similarity Scores of CRT1 ($STTSS_{CRT1}$)

Equation 18:

$$(STTSS)_{CRT1} = \frac{5,501}{5,501 + (TTSS_{(CTR2)}) + (TTSS_{(CTR3)}) + \dots (TTSS_{(CTRn)})} = 0,145$$

Step 6: Calculation of Tanimoto Contrast Score (TCS_{CRT1}) of CRT1 (External Dispersion of CRT1: ED_{CRT1})

Equation 19 – 20:

$$(ED_{CRT1} = TCS_{CRT1}) = \frac{(1 - 0,145)}{0,145} = 5,888$$

Step 7. General Contrast Score of CTR1 Criteria (GCS_{CTR1})

Equation 21:

$$GCS_{CTR1} = 0,140 * 5,888 = 0,822$$

Step 8. Weight Score of CTR1 (w_{CTR1})

Equation 22:

$$w_{CTR1} = \frac{0,822}{6,479} = 0,127$$

3.2. Sensitivity Analysis

Evaluating the resilience of MCDM methodologies often involves the introduction of supplementary criteria or the exclusion of suboptimal options from the initial dataset. Under such circumstances, a robust MCDM

framework should exhibit stability, ensuring that the ordinal arrangement of criteria remains relatively consistent (Demir & Arslan, 2022). To scrutinize this particular facet, a sensitivity analysis was executed, commencing with the criteria identified as possessing the lowest relative importance according to the developed methodology. The resultant country rankings derived from this analytical procedure are tabulated in Table 4, and a corresponding visual representation is depicted in Figure 2.

Table 4. Rank Reversal

Criteria	S0	S1	S2	S3	S4	S5
GII7	7	0	0	0	0	0
GII5	6	6	0	0	0	0
GII1	5	5	4	0	0	0
GII2	4	4	5	4	0	0
GII4	3	3	3	3	3	0
GII6	2	2	2	2	2	2
GII3	1	1	1	1	1	1

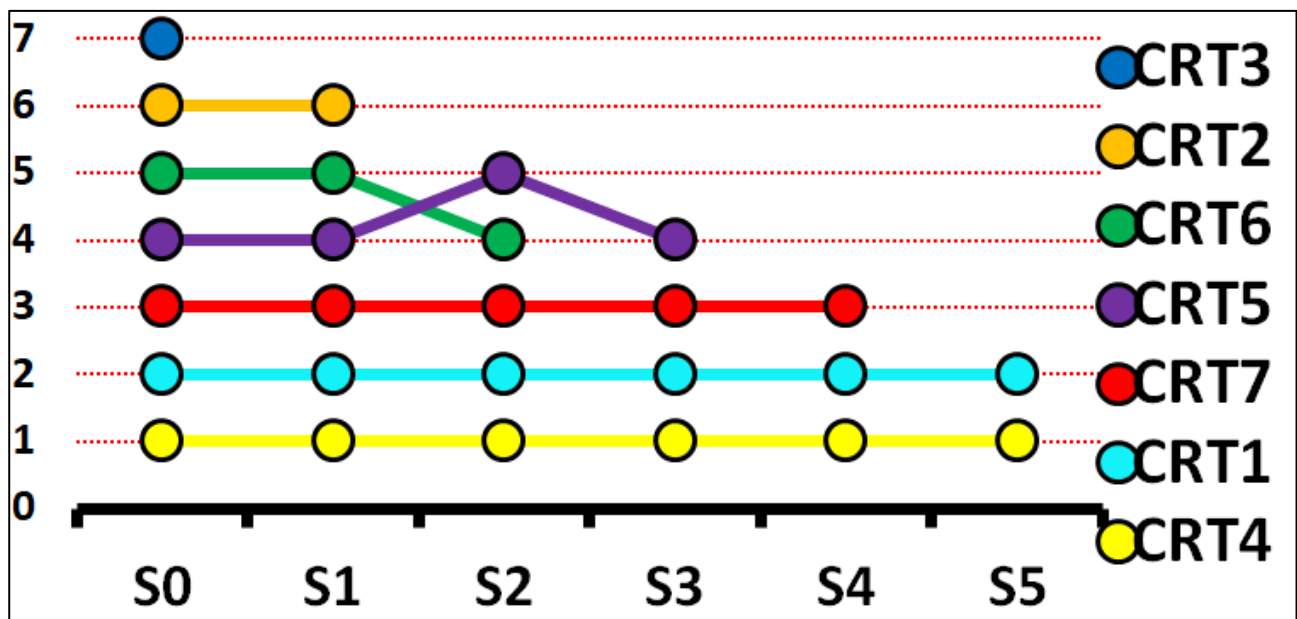


Figure 2. Rank Reversal Chart

When Table 4 and Figure 2 are evaluated together, it is observed that the ranking of criteria remains largely stable under certain scenarios. Specifically, in scenarios S1 (when CRT3 is removed), S3 (when CRT3, CRT2, and CRT6 are removed), S4 (when CRT3, CRT2, CRT6, and CRT5 are removed), and S5 (when CRT3, CRT2, CRT6, CRT5, and CRT7 are removed), the rankings of the criteria remain unchanged. This indicates that the proposed method provides a stable structure in terms of overall ranking. However, in scenario S2 (when CRT3

and CRT2 are removed), certain changes occur. In this scenario, the CRT5 criterion moves from the 4th to the 5th position, while the CRT6 criterion rises from the 5th to the 4th position. This change can be attributed to the relatively small difference in the weight values of these two criteria, as shown in Table 4. Therefore, the fact that the sensitivity analysis yields a different result only in scenario S2, while the overall ranking remains largely preserved, supports the robustness of the method in terms of ranking stability. In this context, it is concluded that the proposed TCA method exhibits an optimal level of sensitivity. The results of the sensitivity analysis demonstrate that the method shows only minor variations when specific criteria are removed, while the overall ranking structure remains largely intact. This finding underscores the robustness of the method and reinforces its reliability in decision-making processes.

3.3. Comparative Analysis

This comparative analysis scrutinizes the interrelationships and relative positions of the proposed methodology against established techniques for weight determination. The objective is to validate the proposed method's efficacy, reliability, and consistency with existing approaches, while demonstrating a robust and statistically significant correlation with diverse weighting methodologies (Keshavarz-Ghorabae et al., 2021). Consequently, the initial stage of this comparative assessment involved the computation of criterion weights using ENTROPY, CRITIC, SD, SVP, LOPCOW, and MEREC methods, all of which are commonly utilized in MCDM research. The resulting weight values for GII criteria, along with their corresponding rankings, as derived from these weighting methods, are presented in Table 5 and Figure 3.

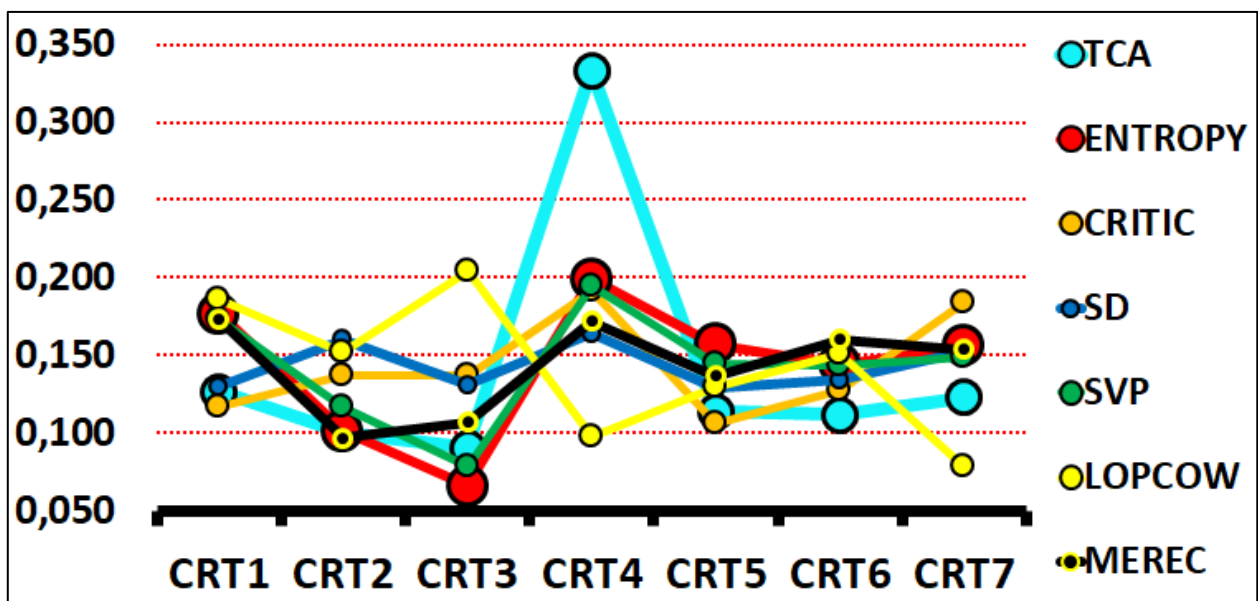
The proposed TCA method's rankings, based on the Table 3 and Table 5, show varying degrees of consistency with different methods across various criteria. Notably, there is a complete alignment between the TCA and SVP rankings. This suggests that both TCA and SVP adopt a similar approach in determining the criteria weights.

Additionally, the rankings derived from the TCA method exhibit a high degree of consistency with those of the ENTROPY method. This indicates that the TCA method functions in parallel with the ENTROPY method's criterion weighting and ranking approach, producing generally similar results. In contrast, the rankings of the TCA method also exhibit a similar trend when compared to the MEREC method. However, when examining other methods, the TCA rankings show more distinct differences, particularly with CRITIC, SD, and LOPCOW. Specifically, the rankings of these methods diverge noticeably from those of TCA.

Both CRITIC and SD methods assign higher weights to certain criteria, while TCA ranks these criteria lower. In conclusion, the TCA method produces rankings that are more consistent with SVP, ENTROPY, and MEREC methods, while it shows noticeable differences in rankings compared to CRITIC, SD, and LOPCOW. These discrepancies are likely attributable to the different weighting techniques used by each method and the distinct importance assigned to the criteria.

Table 5. Weight Scores According to Methods

Criteria	ENTROPY	Rank	CRITIC	Rank	SD	Rank
CRT1	0,177	2	0,117	6	0,130	6
CRT2	0,101	6	0,137	4	0,160	2
CRT3	0,066	7	0,137	3	0,131	5
CRT4	0,199	1	0,191	1	0,164	1
CRT5	0,157	3	0,106	7	0,129	7
CRT6	0,144	5	0,127	5	0,134	4
CRT7	0,156	4	0,184	2	0,153	3
Criteria	SVP	Rank	LOPCOW	Rank	MEREC	Rank
CRT1	0,173	2	0,186	2	0,173	1
CRT2	0,117	6	0,152	3	0,097	7
CRT3	0,078	7	0,204	1	0,107	6
CRT4	0,195	1	0,098	6	0,172	2
CRT5	0,145	4	0,130	5	0,137	5
CRT6	0,143	5	0,151	4	0,160	3
CRT7	0,149	3	0,078	7	0,154	4

*Figure 3. Position of Methods*

A careful examination of Figure 3 reveals that the proposed TCA method generally exhibits a fluctuating pattern that aligns with other weighting methods (ENTROPY, SVP, and MEREC), with the exception of

LOPCOW, CRITIC, and SD. Specifically, the trends of increasing and decreasing criterion values show a clear parallelism between the TCA method and the other methods. This observation supports the hypothesis that, with the exception of LOPCOW, CRITIC, and SD, the TCA method demonstrates a significant, positive, and high correlation with the other methods. In this context, the correlation values between the TCA method and the other methods are presented in detail in Table 6. This correlation analysis is expected to enhance the understanding of the validity of the proposed method and its relationship with other widely used methods.

Table 6. *rho Correlation Scores-1*

Methods	TCA	ENTROPY	CRITIC	SD	SVP	LOPCOW	MEREC
TCA	1						
ENTROPY	0,964**	1					
CRITIC	0,198**	0,018	1				
SD	0,179**	0,036	0,883**	1			
SVP	0,998**	0,964**	0,198*	0,179	1		
LOPCOW	-0,571*	-0,500*	-0,396*	-0,393*	-0,571*	1	
MEREC	0,821**	0,786**	-0,018	-0,036	0,821**	-0,214	1

p**<.01, p*<.05

Upon examining Table 6, this analysis reveals that the TCA method exhibits a significant, high, and positive correlation with the ENTROPY, SVP, and MEREC methods (0.964, 0.998, and 0.821, respectively). This supports the observation in Figure 1 that the fluctuation patterns of the TCA method closely overlap with those of these methods. However, the correlation values between TCA and the CRITIC and SD methods are relatively lower, calculated at 0.198 and 0.179, respectively. These results indicate that the TCA method shows only a limited degree of similarity with these methods. On the other hand, a significant negative correlation (-0.571*) was found between the TCA method and the LOPCOW method. This finding, as seen in Figure 3, confirms that the general trends of the LOPCOW method significantly differ from those of the other methods and are not in parallel with the TCA method. Therefore, this correlation analysis demonstrates that the TCA method generally maintains a meaningful relationship with the ENTROPY, SVP, and MEREC methods, while exhibiting an inverse interaction with the LOPCOW method. This contributes to a better understanding of the compatibility of the proposed method with other commonly used methods. Consequently, the comparative analysis leads to the conclusion that the TCA method is reliable and credible, especially due to its high and significant correlations with the ENTROPY, SVP, and MEREC methods.

3.4. Simulation Analysis

To assess the reliability of the proposed method's findings, a simulation analysis was conducted wherein different values were assigned to the decision matrices to generate various scenarios. It is anticipated that as the number of scenarios increases, the method's results will progressively diverge from those of other criterion

weighting methods. In the subsequent phase, the average variance values derived from the proposed method should exceed the variance values calculated by one or more comparative objective weighting techniques. This outcome would affirm the proposed method's capacity to better differentiate the relative importance of the criteria. Ultimately, the variances of the criterion weights measured across the scenarios are expected to demonstrate homogeneity (Keshavarz-Ghorabae et al., 2021).

In this endeavor, ten distinct scenarios, represented as decision matrices, were initially formulated and partitioned into two discrete groups (First Group: Scenario1, Scenario2, Scenario3, Second Group: Scenario4, Scenario5, Scenario6, Scenario7, Scenario8, Scenario9, Scenario10). Subsequently, the correlation coefficients between the TCA method and other weighting methods were calculated for these scenarios. The results of these correlational analyses are presented in Table 7 and Figure 4.

When Figure 4 and Table 7 are evaluated together, it is observed that the correlation values of the TCA method under different scenarios gradually decrease in comparison to the other methods. Specifically, the initially high correlation values with the SVP, ENTROPY, and MEREC methods show a decline as the number of scenarios increases. In contrast, the correlation with the MEREC method remains generally positive, and although it decreases over time, it stabilizes at a certain level. This indicates that while SVP, ENTROPY, and MEREC provide relatively similar results with the TCA method initially, they begin to diverge as the scenarios increase. The correlation with the CRITIC and SD methods, on the other hand, remains consistently low from the outset. Moreover, a negative correlation is observed between the TCA method and the LOPCOW method. In conclusion, the correlations of the TCA method with other objective weighting methods decrease as the scenarios diversify. Particularly, the negative correlation with the LOPCOW method and the positive but decreasing correlation with SVP, ENTROPY, and MEREC suggest that the TCA method adopts a more characteristic and distinctive structure as the number of scenarios increases.

Table 7. rho Correlation Scores-2

Scenarios	ENTROPY	CRITIC	SD	SVP	LOPCOW	MEREC
Scenario1	0,988**	0,212	0,200	0,999**	-0,555*	0,855**
Scenario2	0,993**	0,244	0,209	0,999**	-0,533*	0,875**
Scenario3	0,921**	0,191	0,176	0,995**	-0,558*	0,809**
Scenario4	0,974**	0,182	0,165	0,998**	-0,544*	0,863**
Scenario5	0,908**	0,169	0,151	0,986**	-0,533*	0,843**
Scenario6	0,888**	0,155	0,136	0,974**	-0,505*	0,831**
Scenario7	0,873**	0,144	0,128	0,951**	-0,488*	0,811**
Scenario8	0,859**	0,132	0,125	0,933**	-0,451*	0,802**
Scenario9	0,837**	0,127	0,103	0,921**	-0,439*	0,788**
Scenario10	0,823**	0,118	0,089	0,909**	-0,421*	0,763**

$p^{**}<.01, p^{*}<.05$

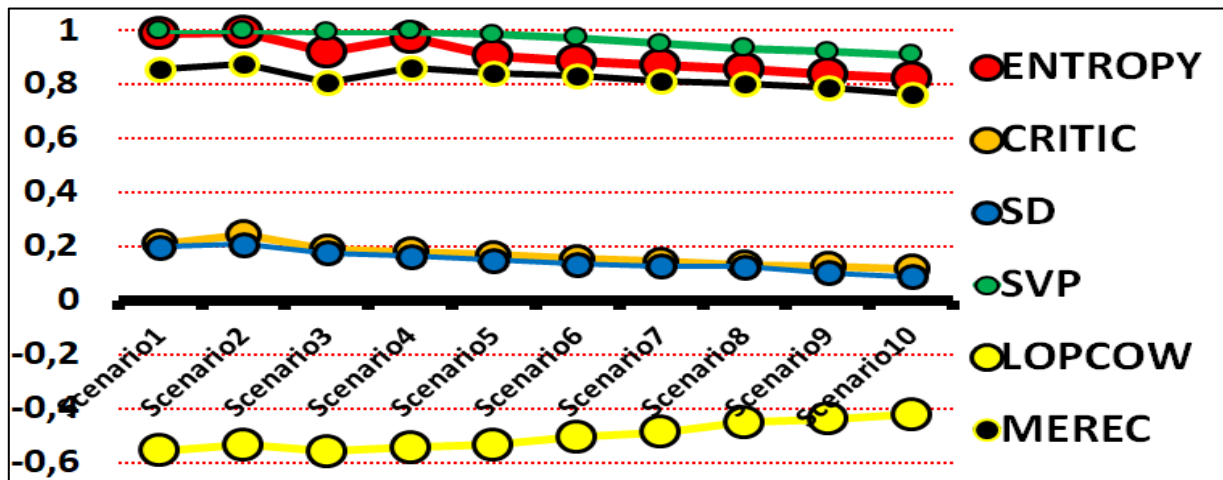


Figure 4. Correlation Positions

In the second phase of the simulation analysis, the variance value for each method was calculated under the respective scenarios, and the average variance values for each method were also determined. The relevant calculated values are presented in Table 8.

When the variance values provided in Table 8 are examined, it is observed that the average variance value of the TCA method is 0.00621, which is higher than the average variance values of all other weighting methods. This finding indicates that the TCA method has a greater capacity to distinguish between differences among the criteria. A high variance value for a weighting method implies that the method assigns weights with a broader distribution, thereby allowing for a clearer differentiation between the criteria. The fact that the variance values obtained by the TCA method are higher than those of other methods such as ENTROPY, CRITIC, SD, SVP, LOPCOW, and MEREC demonstrates that this method can determine the relative importance levels of the criteria more sharply. In this context, the TCA method can be considered as an effective method for identifying the characteristic features of the criteria and revealing the differences between them, as much as other weighting methods. Particularly in analyses where the distinguishing power of the criteria needs to be enhanced, the use of the TCA method may be an appropriate approach.

In the concluding phase of the simulation analysis, the uniformity of variance exhibited by the criterion weights within the TCA methodology was appraised across the spectrum of Levene's test technique. This analytical approach furnishes a visual instrument for evaluating the consistency of variance. The graphical depiction is structured around three pivotal elements: the aggregate mean ADM (ANOM for variances based on Levene), serving as the central benchmark, and the upper decision limit (UDL) and lower decision limit (LDL), which define the boundaries of acceptable variation. Should the variance of any specific group or cluster transcend these decision limits, it signifies a statistically significant departure from the overall mean ADM, thereby indicating heterogeneity of variance. Conversely, if the variance of all clusters remain confined within the UDL and LDL thresholds, it validates the homogeneity of variance. Therefore, the homogeneity of the

variances under the scenarios strengthens the robustness and stability of the method. (Keshavarz-Ghorabae et al., 2021). Figure 5 illustrates the graphical outcomes of this ADM analysis.

Table 8. Variance Scores of Methods in scope of Scenarios

Criteria	TCA	ENTROPY	CRITIC	SD	SVP	LOPCOW	MEREC
Scenario1	0,00640	0,00180	0,00099	0,00025	0,00125	0,00189	0,00093
Scenario2	0,00690	0,00190	0,00093	0,00029	0,00129	0,00187	0,00089
Scenario3	0,00550	0,00156	0,00088	0,00016	0,00116	0,00171	0,00077
Scenario4	0,00580	0,00169	0,00079	0,00015	0,00118	0,00168	0,00071
Scenario5	0,00590	0,00181	0,00071	0,00017	0,00119	0,00166	0,00069
Scenario6	0,00530	0,00145	0,00069	0,00013	0,00132	0,00163	0,00078
Scenario7	0,00490	0,00249	0,00074	0,00013	0,00118	0,00165	0,00071
Scenario8	0,00740	0,00178	0,00129	0,00033	0,00135	0,00176	0,00096
Scenario9	0,00710	0,00197	0,00101	0,00031	0,00129	0,00188	0,00091
Scenario10	0,00690	0,00135	0,00122	0,00039	0,00131	0,00191	0,00093
Mean	0,00621	0,00178	0,00093	0,00023	0,00125	0,00180	0,00083

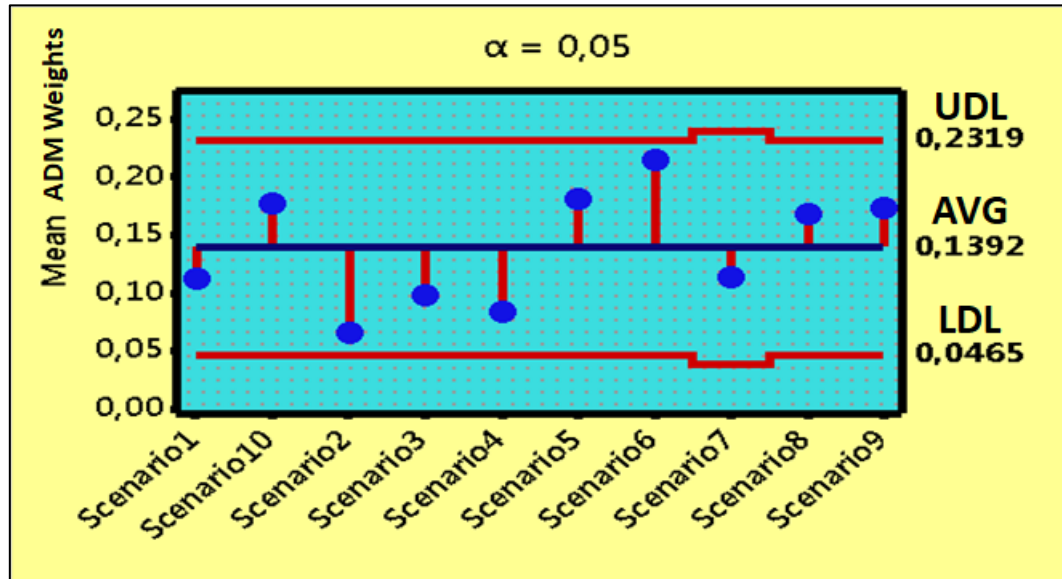


Figure 5. ADM Chart

Figure 5 illustrates the distribution of ADM weights calculated under different scenarios. In the graph, the mean ADM weights for each scenario are represented by blue dots, while the upper and lower confidence limits are indicated by red lines. Notably, the ADM weights for all scenarios remain below the Upper Decision Limit (UDL = 0.2319) and above the Lower Decision Limit (LDL = 0.0465). This observation indicates that the computed weights exhibit a stable distribution within the predefined decision boundaries. Additionally, it is evident that the weight values are predominantly concentrated around the Average (AVG = 0.1392) and do

not display significant fluctuations. These findings demonstrate that the weighting process obtained through the TCA method does not produce extreme outliers and maintains a generally reliable weighting structure. The fact that the weight values remain consistently within the specified decision boundaries suggests that the method generates stable results across different scenarios and adheres to statistically acceptable variance levels. Accordingly, Levene's test was conducted to assess the homogeneity of variances across the scenarios. The results of the Levene's test are presented in Table 9.

Table 9. *Levene Statistic*

Levene Statistic	df1	df2	Sig.
0,163	2	10	0,241
p* < .05			

Upon examining Table 8, it is observed that the p-value (0.241) is greater than the significance level of 0.05 ($p > 0.05$). This finding confirms that the variance of criterion weights across the scenarios does not exhibit a statistically significant difference, thereby affirming the homogeneity of variances. The homogeneity of variances indicates that the weights calculated by the method do not exhibit substantial variations across different scenarios, ensuring a certain level of stability. Consequently, the simulation analysis concludes that the proposed TCA method demonstrates a robust and stable weighting structure. These findings validate that the method is not only applicable to specific datasets or particular cases but is also reliable for various data structures and scenarios, providing objective weight assignments. Notably, the method consistently produces meaningful results even when the dataset contains distinct similarities and differences. In conclusion, when the findings from both Figure 5 and Table 9 are evaluated together, it is evident that the proposed TCA method ensures statistical homogeneity, reliability, and consistency in determining criterion weights. The weights obtained under different scenarios remained within the defined decision boundaries and did not exhibit excessive fluctuations in terms of variability. The variance homogeneity confirmed by Levene's test supports the applicability of the method across diverse datasets and enhances its scientific credibility.

3.5. Discussion

In the sensitivity analysis of the TCA method, it was observed that changes in alternative rankings were minimal and did not lead to significant deviations in performance values. This finding highlights the method's capacity to produce reliable outcomes in decision-making processes. The success of the TCA method lies in its structure, which incorporates the similarities between criteria, and its balanced weighting process based on Tanimoto similarity. This allows the influence of highly similar criteria to be minimized, bringing forward criteria with higher informational value and resulting in a more stable model. Comparative analyses have assessed the similarities and differences between TCA and other methods such as CRITIC, ENTROPY, SD, SVP, LOPCOW, and MEREC. TCA showed strong positive correlations with ENTROPY, SVP, and MEREC, further supporting its reliability and robustness. In simulation analyses, although correlation values between TCA and other objective methods declined over time, the unique structural characteristics of TCA became

more pronounced. The method effectively differentiates criterion weights through high average variance and maintains homogeneity across scenarios. These results demonstrate that TCA is a strong and stable method. Hence, the findings from sensitivity, comparative, and simulation analyses collectively confirm the applicability of the TCA method within the scope of MCDM.

The most prominent advantage of the TCA method lies in its ability to determine criterion weights independently of subjective judgments. This feature enables the production of consistent and reliable results across different applications. Furthermore, by utilizing Tanimoto similarity, the method quantifies the similarities between criteria, thereby preventing excessive weighting and promoting a more balanced weighting and decision-making process. The method's scale independence also allows criteria with different measurement units to be normalized and compared effectively. Additionally, TCA demonstrates high computational efficiency, enabling fast and effective application even with large datasets. However, the proposed criterion weighting method also has certain limitations. First, due to the inclusion of Tanimoto similarity and other computational steps, the method's calculation process can be complex and potentially costly for large datasets. Second, the performance of the method is dependent on the normalization techniques used and the structure of the dataset. Specifically, when criteria exhibit high degrees of similarity, the method's discriminative power may decrease, leading to potential accuracy issues in the weighting process.

Compared to the CRITIC method, the TCA approach does not require the assumption of normal distribution within the dataset and is capable of accounting for non-linear relationships. This characteristic allows TCA to offer a broader applicability, particularly within MCDM processes. While TCA considers contrasts between criteria similar to CRITIC's weighting mechanism, it differs in its theoretical foundations. Considering the Pearson correlation coefficient and its relation to linear regression (Kalaycı, 2014), CRITIC can be deemed more compatible with datasets that exhibit linear structures. On the other hand, the ENTROPY method is sensitive to negative and zero values (Ayçin, 2019), a limitation that does not affect TCA. TCA not only takes into account the individual values of the criteria but also considers the similarities between them. This enables TCA to provide a more comprehensive weighting approach compared to ENTROPY. The SD and SVP methods focus solely on the dispersion of criterion values (Demir et al., 2021), whereas TCA offers a more holistic evaluation by incorporating both similarity and contrast dimensions. Nonetheless, ENTROPY, SD, and SVP provide practical advantages due to their computational simplicity (Demir et al., 2021), while the more complex structure of the TCA method may increase the potential for calculation errors. The LOPCOW method analyzes differences between criteria (Ecer & Pamucar, 2022), while TCA captures both similarities and contrasts, allowing for a more extensive assessment of inter-criteria interactions. However, LOPCOW may be more sensitive to subtle differences between criteria (Ecer & Pamucar, 2022). MEREC, which evaluates the impact of removing each criterion from the system, offers a distinct approach (Keleş, 2023), while TCA combines similarity and contrast assessments to better accommodate non-linear relationships. Although MEREC may be more effective in identifying critical criteria (Keshavarz-Ghorabae et al., 2021), TCA

provides decision-makers with a deeper analytical framework. Additionally, since MEREC directly illustrates the system changes when a criterion is removed, it offers a more intuitive and interpretable structure for decision-makers (Keshavarz-Ghorabae et al., 2021). In contrast, the similarity and contrast based analysis required by TCA may make it more challenging to clearly explain how results are derived.

The proposed TCA method offers an original contribution to the literature on objective weighting techniques. By utilizing spatial contrasts and Tanimoto similarity to determine criterion weights, TCA introduces a novel perspective distinct from existing methods. This approach enables a more comprehensive evaluation, particularly in complex decision-making scenarios, by modeling different dimensions of inter-criteria relationships. Moreover, by adapting Tanimoto similarity originally used in biology and chemistry to the MCDM domain, the method facilitates interdisciplinary knowledge transfer and incorporates spatial information embedded within the dataset into the evaluation process. In addition, as the proposed method is entirely data-driven, it enhances objectivity and consistency in decision-making processes, offering new opportunities for solving more complex problems. In summary, the proposed TCA method contributes to the literature by introducing a new objective weighting approach, offering a unique perspective on inter-criteria relationships, integrating a novel concept like Tanimoto similarity into the MCDM domain, and utilizing spatial data characteristics for enhanced evaluation. These contributions have the potential to open new avenues for modeling and solving real-world, complex decision-making problems. Furthermore, the TCA method significantly enriches the weighting literature within the scope of MCDM by incorporating a statistical concept Tanimoto similarity into a numerical weighting framework. This interdisciplinary integration strengthens the literature by combining numerical and statistical methodologies. Thus, the TCA method introduces a novel perspective to weighting processes and offers a multidisciplinary approach for addressing more complex and realistic problems, contributing meaningfully to both numerical methods and statistical literature.

4. CONCLUSION

As a result, the TCA method emerges as a scientifically grounded and practically applicable approach for objective criterion weighting in MCDM processes. By integrating the internal variation based on standard deviation with the external contrast structure derived from Tanimoto similarity, the method enables a more detailed and precise evaluation of the contribution of each criterion.

Its resilience to data changes, insensitivity to rank reversal, strong correlation with other objective weighting methods, and consistency observed in simulation environments further reinforce the methodological robustness of the approach. Furthermore, the comparative advantages of TCA over widely used methods in the literature, such as ENTROPY, CRITIC, SD, SVP, LOPCOW, and MEREC, are noteworthy in terms of interpretability, stability, and comprehensiveness. Therefore, the TCA method can be recommended as an effective tool for decision-makers seeking robust and reliable weighting mechanisms. In conclusion, TCA offers a comprehensive approach among objective weighting methods, particularly excelling in the modeling

success of inter-criterion similarities. The method provides decision-makers with a strong alternative, both in terms of its theoretical foundations and practical application outcomes. In this regard, it can be stated that TCA fills a significant gap in the multi-criteria decision-making literature and opens up new perspectives for future research.

Regarding the proposed method, the current study has been conducted on a sample decision problem that includes a limited number of criteria and alternatives. In future research, it would be valuable to test the performance of the TCA method on more complex and multidimensional decision-making problems across various sectors such as sustainability, healthcare, logistics, and public administration. In the logistics sector, the TCA method could be employed to address critical decision-making problems such as distribution center location selection, supply chain optimization, and routing challenges. Given the necessity to evaluate a multitude of alternatives while balancing conflicting criteria, the TCA method offers decision-makers a more objective and comprehensive analytical framework capable of capturing nuanced distinctions among alternatives. Similarly, in the realm of public policy, the TCA method holds significant promise for evaluating the effectiveness of policy alternatives, conducting risk assessments, and determining optimal resource allocation strategies. Considering the inherently multi-criteria nature of public administration processes, the TCA method's ability to accommodate a broad range of measurement scales provides decision-makers with more granular and reliable insights. In the healthcare sector, the TCA method can serve as a robust tool for assessing healthcare service quality, comparing hospital performance, and optimizing healthcare delivery systems. This methodological approach is particularly valuable in reconciling conflicting criteria such as patient satisfaction, treatment duration, and cost-effectiveness, thereby offering a more holistic perspective in complex healthcare decision-making scenarios. Therefore, future research should focus on systematically exploring the practical applications of the TCA method across diverse sectors, utilizing complex, multi-dimensional decision-making contexts to rigorously validate its effectiveness and further extend its applicability. Such investigations would not only substantiate the method's practical relevance but also contribute significantly to the broader body of decision-making literature by positioning the TCA method as a comprehensive and versatile decision-support framework. Secondly, subjecting the weights obtained by the TCA method to sensitivity analysis would be beneficial for understanding the stability of the proposed method under different data structures and increased scenario variations. Additionally, the applicability of the TCA method in decision-making environments characterized by uncertainty such as those involving fuzzy logic and grey system theory could be explored to assess its robustness in such contexts. Finally, the integration of TCA derived weights with different MCDM ranking algorithms (e.g., TOPSIS, VIKOR, MAUT, WASPAS, etc.) can be investigated. This would enable the development of more holistic decision support systems where both the weighting and ranking processes are jointly optimized.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Ahmad, N., Ibrahim, M. A., Sayed, S. R., Shah, S. A., Tahir, M. H., & Zou, Y. (2025). Data mining and machine learning based search for optimal materials for perovskite and organic solar cells. *Solar Energy*, 267, 105-113. <https://doi.org/10.1016/j.solener.2024.113223>
- Anastasiu, D. C., & Karypis, G. (2017). Efficient identification of Tanimoto nearest neighbors. *Int J Data Sci Anal*, 4, 153–172. <https://doi.org/10.1007/s41060-017-0064-z>
- Ayçin, E. (2019). *Çok Kriterli Karar Verme*. Ankara: Nobel Yayın.
- Aydemir, M. F. (2025). Evaluation of foreign direct investment attractiveness of BRICS-T countries: The critic-lopcow based aras approach. *Politik Ekonomik Kuram*, 9(1), 372-392. <https://doi.org/10.30586/pek.1613421>
- Bajusz, D., Rácz, A., & Héberger, K. (2015). Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations? *Journal of Cheminformatics*, 7(20), 1-13. <https://doi.org/10.1186/s13321-015-0069-3>
- Baldi, P., & Benz, R. W. (2008). BLASTing small molecules—statistics and extreme statistics of chemical similarity scores. *BIOINFORMATICS*, 24, 357-365. <https://doi.org/10.1093/bioinformatics/btn187>
- Baş, F. (2021). *Çok Kriterli Karar Verme Yöntemlerinde Kriter Ağırlıklarının Belirlenmesi*. Ankara: Nobel Bilimsel.
- Baydaş, M., Elma, O. E., & Stević, Ž. (2024). Proposal of an innovative MCDA evaluation methodology: knowledge discovery through rank reversal, standard deviation, and relationship with stock return. *Financial Innovation*, 10(4), 1-35. <https://doi.org/10.1186/s40854-023-00526-x>
- Bero, S. A., Muda, A. K., Choo, Y. H., Muda, N. A., & Pratama, S. F. (2018). Similarity measure for molecular structure: A brief review. *IOP Conf. Series: Journal of Physics: Conf. Series*(892), 1-9. <https://doi.org/10.1088/1742-6596/892/1/012015>
- Bircan, H. (2020). *Çok Kriterleri Karar Verme Problemlerinde Kriter Ağırlıklandırma Yöntemleri*. Ankara: Nobel Akademik.
- Chung, N. C., Miasojedow, B., Startek, M., & Gambin, A. (2019). Jaccard/Tanimoto similarity test and estimation methods for biological presence-absence data. *BMC Bioinformatics*, 20(15), 1-11. <https://doi.org/10.1186/s12859-019-3118-5>
- Demir, G., & Arslan, R. (2022). Sensitivity analysis in multi-criteria decision-making problems. *Ankara Hacı Bayram Veli Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 24(3), 1025-1056.
- Demir, G., Özyalçın, A. T., & Bircan, H. (2021). *Çok kriterli karar verme yöntemleri ve ÇKKV yazılımı ile problem çözümü*. Ankara: Nobel.
- Dhara, A., Kaur, G., Kishan, P. M., Majumder, A., & Yadav, R. (2022). An efficient decision support system for selecting very light business jet using critic-topsis method. *Aircraft Engineering and Aerospace Technology*, 94(3), 458-472. <https://doi.org/10.1108/AEAT-04-2021-013>
- Diakoulaki, D., Mavrotas, G., & Papayannakis, L. (1995). Determining objective weights in multiple criteria problems: The critic method. *Computers & Operations Research*, 22(7), 763-770.

- Dinçer, S. E. (2019). *Çok Kriterli Karar Alma*. Ankara: Gece Akademi.
- Dokmanic, I., Parhizkar, R., Ranieri, J., & Vetterli, M. (2015). Euclidean distance matrices: Essential theory, algorithms, and applications. *IEEE Signal Processing Magazine*, 32(6), 12-30. <https://doi.org/10.1109/MSP.2015.2398954>
- Durdu, D. (2025). Evaluating financial performance with spc-lopcow-marcos hybrid methodology: A Case study for firms listed in BIST sustainability index. *Knowledge and Decision Systems with Applications*, 1, 92-111. <http://www.doi.org/10.59543/kadsa.v1i.13879>
- Ecer, F. (2020). *Çok Kriterli Karar Verme*. Ankara: Seçkin Yayıncılık.
- Ecer, F., & Pamucar, D. (2022). A novel lopcow-dobi multi-criteria sustainability performance assessment methodology: An application in developing country banking sector. *Omega*, 1-35. <https://doi.org/10.1016/j.omega.2022.102690>
- Eligüz el, İ. M., & Eligüz el, N. (2024). Evaluation of sustainable transportation in 25 European countries using gra and entropy mabac. *Journal of Transportation and Logistics*, 9(2), 245-259. <https://doi.org/10.26650/JTL.2024.1437521>
- Feldmann, C. W., Sieg, J., & Mathea, M. (2025). Analysis of uncertainty of neural fingerprint-based models. *Faraday Discussions*, 256, 551-567. <http://dx.doi.org/10.1039/D4FD00095A>
- Feldmann, C., & Bajorath, J. (2022). Calculation of exact Shapley values for support vector machines with Tanimoto kernel enables model interpretation. *iScience*, 25, 1-13. <https://doi.org/10.1016/j.isci.2022.105023>
- Iglesias, F., & Kastner, W. (2013). Analysis of similarity measures in times series clustering for the discovery of building energy patterns. *Energies*, 6, 579-597. <https://doi.org/10.3390/en6020579>
- Kalaycı, Ş. (2014). *SPSS Uygulamalı Çok Değişkenli İstatistik Teknikleri*. Ankara: Anı Yayın Dağıtım.
- Karagöz, Y. (2014). *SPSS 21.1 Uygulamalı İstatistik Tıp, Eczacılık, Dış Hekimliği ve Sağlık Bilimleri İçin*. Ankara: Nobel Akademik Yayıncılık.
- Keleş, N. (2023). *Uygulamalarla Klasik ve Güncel Karar Verme Yöntemleri*. Ankara: Nobel Bilimsel.
- Keshavarz-Ghorabae, M., Amiri, M., Zavadskas, E. K., Turskis, Z., & Antucheviciene, J. (2021). Determination of objective weights using a new method based on the removal effects of criteria (MEREC). *Symmetry*, 13(525), 1-21. <https://doi.org/10.3390/sym13040525>
- Kilmen, S. (2015). *Eğitim Araştırmacıları için SPSS uygulamalı istatistik*. Ankara: Edge Akademi.
- Kryszkiewicz, M. (2013). Using non-zero dimensions for the cosine and tanimoto similarity search among real valued vectors. *Fundamenta Informaticae*, 127(1-4), 307-323. <https://doi.org/10.3233/FI-2013-911>
- Kryszkiewicz, M. (2021). Determining Tanimoto Similarity Neighborhoods of Real-Valued Vectors by Means of the Triangle Inequality and Bounds on Lengths. In: S. C. Ramanna (Eds.), *Rough Sets. IJCRS Lecture Notes in Computer Science* (pp. 8–34). Singapore Springer.
- Lasek, P., & Mei, Z. (2019). Clustering and visualization of a high-dimensional diabetes dataset. *Procedia Computer Science*, 159, 2179–2188. <https://doi.org/10.1016/j.procs.2019.09.392>

- Lipkus, A. H. (1999). A proof of the triangle inequality for the tanimoto distance. *Journal of Mathematical Chemistry*, 26, 263–265. <https://doi.org/10.1023/A:1019154432472>
- Ma, C., Wang, L., & Xie, X.-Q. (2011). GPU accelerated chemical similarity calculation for compound library comparison. *Journal of Chemical Information and Modeling*, 51(7), 1521–1527. <https://doi.org/10.1021/ci1004948>
- Mana, S. C., & Sasipraba, T. (2021). Research on cosine similarity and pearson correlation based. *International Conference on Mathematical Sciences*(1770), 1-7. <https://doi.org/10.1088/1742-6596/1770/1/012014>
- Mellor, C. L., Marchese, R. L., Robinson, Benigni, R., Ebbrell, D., Enoch, S. J., . . . Cronin, M. (2019). Molecular fingerprint derived similarity measures for toxicological read across: Recommendations for optimal use. *Regulatory Toxicology and Pharmacology*, 101, 121-134.
- Mohebbi, A., Ebrahimi, M., Askari, F. S., Shaddel, R., Mirarab, A., & Oladnabi, M. (2022). QSAR modeling of a ligand-based pharmacophore derived from hepatitis B virus surface antigen inhibitors. *ACTA MICROBIOLOGICA BULGARICA*, 38(3), 197-206.
- Momena, A. F., Gazi, K. H., Rahaman, M., Sobczak, A., Salahshour, S., Mondal, S. P., & Ghosh, A. (2024). Ranking and challenges of supply chain companies using MCDM methodology. *Logistics*, 8, 1-32. <https://doi.org/10.3390/logistics8030087>
- Mukhametzhanov, I. Z. (2021). Specific character of objective methods for determining weights of criteria in mcdm problems: Entropy, critic, sd. decision making. *Applications in Management and Engineering*, 4(2), 76-105.
- Mulia, I., Kusuma, W. A., & Afendi, F. M. (2018). Algorithm for predicting compound protein interaction using tanimoto similarity and klekota-roth fingerprint. *TELKOMNIKA*, 16(4), 1785-1792. <https://doi.org/10.12928/TELKOMNIKA.v16i4.5916>
- Nasser, A. A., Alkhulaidi, A., Ali, M. N., Hankal, M., & Al-olofe, M. (2019). A study on the impact of multiple methods of the data normalization on the result of saw, wed and topsis ordering in healthcare multi-attributte decision making systems based on ew, entropy, critic and svp weighting approaches. *Indian Journal of Science and Technology*, 12(4), 1-21. <https://doi.org/10.17485/ijst/2019/v12i4/140756>
- Nowatzky, Y., Russo, F. F., Lisec, J., Kister, A., Reinert, K., Muth, T., & Benner, P. (2025). FIORA: Local neighborhood-based prediction of compound mass spectra from single fragmentation events. *Nature Communications*, 16, 1-17. <https://doi.org/10.1038/s41467-025-57422->
- Özdamar, K. (2018). *Eğitim, Sağlık ve Sosyal Bilimler için SPSS Uygulamalı Temel İstatistik*. Bursa: Nisan Yayınevi.
- Özgel, A., & Alp, İ. (2020). *Çok Kriterli Karar Verme Seçiminde Yeni Bir Yaklaşım*. İstanbul: Kriter Yayıncılık.
- Paulose, R., Jegatheesan, K., & Balakrishnan, G. S. (2018). A big data approach with artificial neural network and molecular similarity for chemical data mining and endocrine disruption prediction. *Indian Journal of Pharmacology*, 50(4), 169-176. https://doi.org/10.4103/ijp.IJP_304_17
- Podani, J. (2000). Introduction to the exploration of multivariate biological data. Leiden: Backhuys Publishers.

- Popović, G., Pucar, D., & Smarandache, F. (2022). MEREC-COBRA approach in e-commerce development strategy selection. *Journal of Process Management and New Technologies*, 10(3-4), 66-74.
- Sankara, R. A., Durga, B. S., Sobha, R. T., Raju, S. B., & Narahari, S. G. (2011). Study of Diversity and Similarity of Large Chemical Databases Using Tanimoto Measure. In: K. R. Venugopal & L. M. Patnaik (Eds.). *Communications in Computer and Information Science* (pp. 1-12. https://doi.org/10.1007/978-3-642-22786-8_5
- Sehgal, K., Kaur, H., Kaur, S., Singh, S., Channi, H. K., & Stević, Ž. (2025). Cost-Effective optimization of hybrid renewable energy system for micro, small, and medium enterprises: A decision-making framework integrating merec and marcos. *Opportunities and Challenges in Sustainability*, 4(1), 17-32. <https://doi.org/10.56578/ocs040102>
- Selvi, C., & Sivasankar, E. (2018). *A Novel Singularity Based Improved Tanimoto Similarity Measure for Effective Recommendation Using Collaborative Filtering*. In: Proceedings of the 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (NODIA, 2018)., (pp. 256-262).
- Shan, R., Zhang, R., Gao, Y., Wang, W., Zhu, W., Xin, L., . . . Cui, P. (2025). Evaluating ionic liquid toxicity with machine learning and structural similarity methods. *Green Chemical Engineering*, 6, 249–262. <https://doi.org/10.1016/j.gce.2024.08.008>
- Sharma, A., & Lal, S. P. (2011). Tanimoto based similarity measure for intrusion detection system. *Journal of Information Security*, 2, 195-201. <https://doi.org/10.4236/jis.2011.24019>
- Sleem, A., Mostafa, N., & Elhenawy, I. (2023). Neutrosophic critic mcdm methodology for ranking factors and needs of customers in product's target demographic in virtual reality metaverse. *Neutrosophic Systems with Applications*, 2, 55-66. <https://doi.org/10.61356/j.nswa.2023.10>
- Surendran, A., Zsigmond, K., López-Pérez, K., & Miranda-Quintana, R. A. (2025). Is Tanimoto a metric? *bioRxiv*, 1-12. <https://doi.org/10.1101/2025.02.18.638904>
- Tayalı, H. A., & Timor, M. (2017). Ranking with Statistical variance procedure based analytic hierarchy process. *Acta Infologica*, 1(1), 31-38.
- Thakkar, J. J. (2021). *Multi criteria decision making*. Springer Singapore.
- Thiel, P., Sach-Peltason, L., Ottmann, C., & Kohlbacher, O. (2014). Blocked inverted indices for exact clustering of large chemical spaces. *Chem. Inf. Model*, 54(9), 2395–2401.
- Todeschini, R., Ballabio, D., Consonni, V., Mauri, A., & Pavan, M. (2007). CAIMAN (Classification And Influence Matrix Analysis): A new approach to the classification based on leverage-scaled functions. *Chemometrics and Intelligent Laboratory Systems*, 87(1), 3-17. <https://doi.org/10.1016/j.chemolab.2005.11.001>
- Uludağ, A. S., & Doğan, H. (2021). *Üretim Yönetiminde Çok Kriterli Karar Verme*. Ankara: Nobel.
- Ulutaş, A., & Topal, A. (2020). *Bütünleştirilmiş Çok Kriterli Karar Verme Yöntemlerinin Üretim Sektörü Uygulamaları*. Ankara: Akademisyen Kitapevi.

- Vivek, M. B., Manju, N., & Vijay, M. B. (2018). *Machine Learning Based Food Recipe Recommendation System*. In: Guru, D. S., Vasudev, T., Chethan, H. K., & Kumar, Y.S, In: Proceedings of International Conference on Cognition and Recognition (SINGAPORE 2018), (pp. 11-19).
- Wang, C.-N., Nguyen, N.-A.-T., & Dang, T.-T. (2023). Sustainable evaluation of major third-party logistics providers: a framework of an mcdm-based entropy objective weighting method. *Mathematics*, 11, 1-27. <https://doi.org/10.3390/math11194203>
- Xu, H., Zeng, W., Zeng, X., & Yen, G. (2019). An evolutionary algorithm based on minkowski distance for many objective optimization. *IEEE Transactions on Cybernetics*, 49(11), 3968-3979. <https://doi.org/10.1109/TCYB.2018.2856208>
- Yasin, Y., Riaz, M., & Alsager, K. (2025). Synergy of machine learning and the einstein choquet integral with lopcowand fuzzy measures for sustainable solid waste management. *AIMS Mathematics*, 10(1), 460–498. <https://doi.org/10.3934/math.2025022>
- Yazdi, A. K., Tan, Y., Birau, R., Frank, D., & Pamučar, D. (2025). Sustainable solutions: using mcdm to choose the best location for green energy projects. *International Journal of Energy Sector Management*, 19(1), 146-180. <https://doi.org/10.1108/IJESM-01-2024-0005>
- Ying, C. Z., Leong, C. H., Alvin, L. W., & Yang, C. J. (2021). Improving the Workflow of Chemical Structure Elucidation with Morgan Fingerprints and the Tanimoto Coefficient. In: Guo, H, Ren, H & Kim, N (Eds.). *IRC-SET 2020* (pp. 13-24). Singapore Springer.
- Yoon, D., & Lee, H. (2025). In silico discovery of novel compounds for fak activation using virtual screening, AI based prediction, and molecular dynamics. *Computational Biology and Chemistry*, 117, <https://doi.org/10.1016/j.compbiolchem.2025.108420>
- Yoshizawa, T., Ishida, S., Sato, T., Ohta, M., Honma, T., & Terayama, K. (2025). A data-driven generative strategy to avoid reward hacking in multi-objective molecular design. *Nature Communications*, 16, 1-12. <https://doi.org/10.1038/s41467-025-57582-3>
- Zahrotun, L. (2016). Comparison jaccard similarity, cosine similarity and combined both of the data clustering with shared nearest neighbor method. *Computer Engineering and Applications*, 5(1), 11-18.
- Zainudin, S., & Nurjana, N. R. (2018). Comparison of similarity method to improve retrieval performance for chemical data. *Jurnal Teknologi Maklumat dan Multimedia Asia-Pasifik*, 7(1), 91-98. <https://doi.org/10.17576/apjitm-2018-0701-08>.
- Zardari, N. H., Ahmed, K., Shirazi, S. M., & Yusop, Z. B. (2014). *Weighting Methods and their Effects on Multi Criteria Decision Making Model Outcomes in Water Resources Management*. Springer Nature.
- Zhang, B., Vogt, M., Maggiora, G. M., & Bajorath, J. (2015). Design of chemical space networks using a Tanimoto similarity variant based upon maximum common substructures. *J Comput Aided Mol Des*, 29, 937–950. <https://doi.org/10.1007/s10822-015-9872-1>
- Zhang, X., Wang, C., Li, E., & Xu, C. (2014). Assessment model of ecoenvironmental vulnerability based on improved entropy weight method. *Scientific World Journal*, 2014, 1-7. <http://dx.doi.org/10.1155/2014/797814>



Gazi University

Journal of Science

PART A: ENGINEERING AND INNOVATION

<http://dergipark.org.tr/guj.1664093>

Comparison of Deep Learning Algorithms for Image Segmentation on Satellite Images

Huseyin ACEMLI¹ Nida KUMBASAR^{2*} ¹ Ozyegin University, İstanbul, Türkiye² TÜBİTAK, Informatics and Information Security Research Center (BİLGEM), Kocaeli, Türkiye

Keywords	Abstract
Remote Sensing High-Spatial Resolution Semantic Segmentation Deep Learning	Recent advancements in deep learning have significantly contributed to the development of high spatial resolution (HSR) land cover mapping. However, the distinct geographic patterns between urban and rural areas have limited the generalizability of deep learning algorithms across these domains. To address this challenge, separate datasets for rural and urban environments have been proposed in the literature, aiming to achieve more reliable results in real-world applications. In this study, we utilize the publicly available LoveDA HSR dataset for model and parameter comparison. Experiments were conducted on two distinct scenarios: rural and urban areas. The combination of the Adam optimizer, Dice loss function, and UNet++ architecture exhibited the highest performance in both datasets. A weighted average of this combination, based on the number of test samples, was calculated for both groups, yielding a final performance score of 62.14% in terms of mean Intersection over Union (IoU).
Cite	
Acemli, H., & Kumbasar, N. (2025). Comparison of Deep Learning Algorithms for Image Segmentation on Satellite Images. <i>GU J Sci, Part A, 12(2)</i> , 479-502. doi: 10.54287/guj.1664093	
Author ID (ORCID Number)	Article Process
0009-0004-1804-4834 Huseyin ACEMLI	Submission Date 25.03.2025
0000-0001-5497-4618 Nida KUMBASAR	Revision Date 17.04.2025
	Accepted Date 26.05.2025
	Published Date 30.06.2025

1. INTRODUCTION

Image analysis in remote sensing is becoming more and more widespread as artificial intelligence advances in the field of computer vision. In addition, with the help of various sensors, satellite vehicles and platforms, remote sensing technology is renewing itself day by day. Remote sensing image segmentation aims to segment images with semantic labels. Deep learning (DL) algorithms with a representation learning approach have been used for many segmentation problems including remote sensing in recent years and these algorithms have been observed to be very useful.

High spatial resolution (HSR) land-cover datasets are land-cover data created using high-resolution satellite imagery or aerial photography. HSR data allow the study of fine details and small areas, usually on images with a resolution of 1 meter or higher. Such data play an important role in various fields such as urban planning (Zou et al., 2024), environmental management (Zhang et al., 2023), agricultural traceability (Aksoy et al., 2023), climate change (Abunnasr & Mhawej, 2023), natural disaster assessment (Xia et al., 2023). HSR data can be collected from spaceborne sensors such as Sentinel 2, Landsat or airborne sensors such as LiDAR.

*Corresponding Author, e-mail: nida.kumbasar@tubitak.gov.tr

Spaceborne sensors are devices usually located on satellites and collect data by observing the Earth from space, while airborne sensors are sensors usually located on aircraft such as airplanes, helicopters or drones. Spaceborne sensors are in space, usually on orbiting satellite systems. Because they are well positioned on fixed platforms, spaceborne sensors have fewer problems with degradation than airborne sensors and offer continuous monitoring of large areas. Visiting a given area at regular intervals, these sensors can be affected by the state of the atmosphere, but are weather independent. On the downside, they sometimes provide coarse resolution data as an image can cover several hundred square kilometers.

In the literature, it has been emphasized that spatial resolution is more important than spectral resolution when extracting urban land cover information from remote sensing image (RSI) data (Neupane et al., 2021). It has been concluded that an image pixel with good resolution is more useful for analyzing RSI data than a larger number of spectral bands or a narrower wavelength range (Myint et al., 2011). The rapid accumulation and availability of high-resolution RSI and the advancement of DL methods have shifted HSR data from traditional pixel-based and object-based methods to DL-based semantic segmentation. Semantic segmentation is an approach based on assigning a class label to each image pixel in order to make the image highly interpretable. Land cover semantic segmentation in remote sensing aims to identify the type of land cover in each image pixel. The main objective of the remote sensing image segmentation process is to divide the image into a homogeneous set of segments according to certain criteria and map the separate regions to real world objects such as buildings, rivers, fields, roads, etc.(Long et al., 2015).

In traditional pixel-based approaches, pixel size may be insufficient to identify an object in an image. Although significant progress has been made with convolutional neural network (CNN)-based DL, the high intra-class variation and low sensor resolution of remote sensing images make segmentation difficult (Chan et al., 2021). To overcome these challenges, various strategies such as hierarchical feature structures (Tao et al., 2020), polymorphism (Peng et al., 2019) and fusion schemes (Yu et al., 2018) have been used to extend the pipeline. Kemker et al. (Kemker et al., 2018) improved performance with a hybrid architecture based on SharpMask and RefineNet for six-band multi-spectral image segmentation. Kampffmeyer et al. (Kampffmeyer et al., 2016) performed land cover mapping segmentation in urban areas by combining Monte Carlo dropout uncertainty maps with three CNN models. Sun and Wang (Sun & Wang, 2018) utilized a digital surface model based on geometry information to improve the segmentation results of HSR images with a fully convolutional network (FCN). Yuan et al. (Yuan et al., 2021) designed a multichannel fusion module for water body detection using RGB, NIR and SWIR bands of Sentinel-2. To cope with different image resolutions, they proposed that the multichannel water body detection network (MC-WBDN) is more robust to changes in light and weather conditions and can better distinguish small water bodies compared to other models. Baek et al. (Baek et al., 2024) implemented a modified U-Net architecture, called SiU-Net, with two separate inputs for Sentinel-2's RGB and NIR data. They also performed a comparison with DeepLabV3+ and U-Net to evaluate the performance of their model. Hossain et al. (Hossain & Chen, 2022) introduced a hybrid segmentation algorithm that integrates homogeneity and heterogeneity simultaneously to identify buildings. In that

algorithm, since no prior knowledge about the shape of building footprints is required, all building shapes are included in the analysis and a donut-filling technique is introduced to extract roof elements. The algorithm resulted in homogeneity within segments and heterogeneity between segments. Finally, the study segmented small and large buildings without using scale or object size parameters.

The development of convolutional networks with existing HSR land cover datasets such as Gaofen Image Dataset (Kampffmeyer et al., 2016), DeepGlobe (Demir et al., 2018), which contain pixel-wise information for remote sensing applications, has been encouraged. However, these datasets ignored the different styles between the geographical areas of urban and rural areas. Wang et al. (J. Wang et al., 2021) prepared the LoveDA dataset based on the knowledge that urban and rural land covers have large differences in class distributions, object scales and pixel spectra, and aimed to improve model generalizability separately. In the HSR images of different cities in China, they observed different patterns, especially in buildings, roads and wetlands. For example, they emphasized that buildings in rural areas are more irregularly arranged than in urban areas, and roads in rural areas are narrower than in urban areas. They noted that agricultural areas are large-scale and continuous in rural areas, while in urban areas they are found in the spaces between buildings. Wetlands are located in the form of small-scale ponds and ditches in rural areas and large-scale rivers and lakes in urban areas. In their study, they emphasized that UNet++ performs better than UNet due to its nested structure.

In this study, the LoveDA dataset was used to compare the most preferred models and parameters in the literature for urban and rural areas. UNet, UNet++, MANet and DeepLabv3+ were used as architectures; Dice Loss, Cross Entropy Loss and Weighted Cross Entropy Loss were used as loss functions; Adam and Stochastic Gradient Descent (SGD) were used as optimizers.

The rest of the paper is organized as follows: Section 2 describes the dataset, DL architectures, loss functions and optimizers used in the study. Section 3 presents the performance evaluation criteria, the experimental results and the comparison between the experimental groups. Section 4 discusses the limitations, with a comparison to the literature, and emphasizes the importance of the work. Finally, conclusions and future work are presented in Section 5.

2. MATERIAL AND METHOD

This section describes the dataset, DL approaches, loss functions and optimizers used in the study.

2.1. Dataset

In this study, we utilize the publicly available LoveDA dataset (J. Wang et al., 2021), introduced in the literature by Wang et al. in 2021. The LoveDA dataset was collected in July 2016 from 18 diverse urban and rural scenes in Nanjing, Changzhou, and Wuhan, covering a total area of 536.15 km². The dataset comprises 5,987 high spatial resolution (HSR) images, each with a size of 1024x1204 pixels and a spatial resolution of

0.3 meters. The original dataset contains 166,768 objects and is divided into training, validation, and test sets. However, since ground truth (GT) masks for the test data were unavailable for performance evaluation, this study focuses on the training and validation data. To this end, the training and validation dataset was shuffled and then randomly split into 80% for training and 20% for testing. As summarized in Table 1, a total of 4,191 RGB remote sensing image-mask pairs were used in this study.

Table 1. Image distribution in training and test sets

	Train	Test	Total
Rural Dataset	1886	472	2358
Urban Dataset	1466	367	1833
Total	3352	839	4191

Sample images and corresponding masks from rural and urban scenes, considering the segmentation task with a total of 8 classes, including the "No Data" class, are presented in Figure 1.

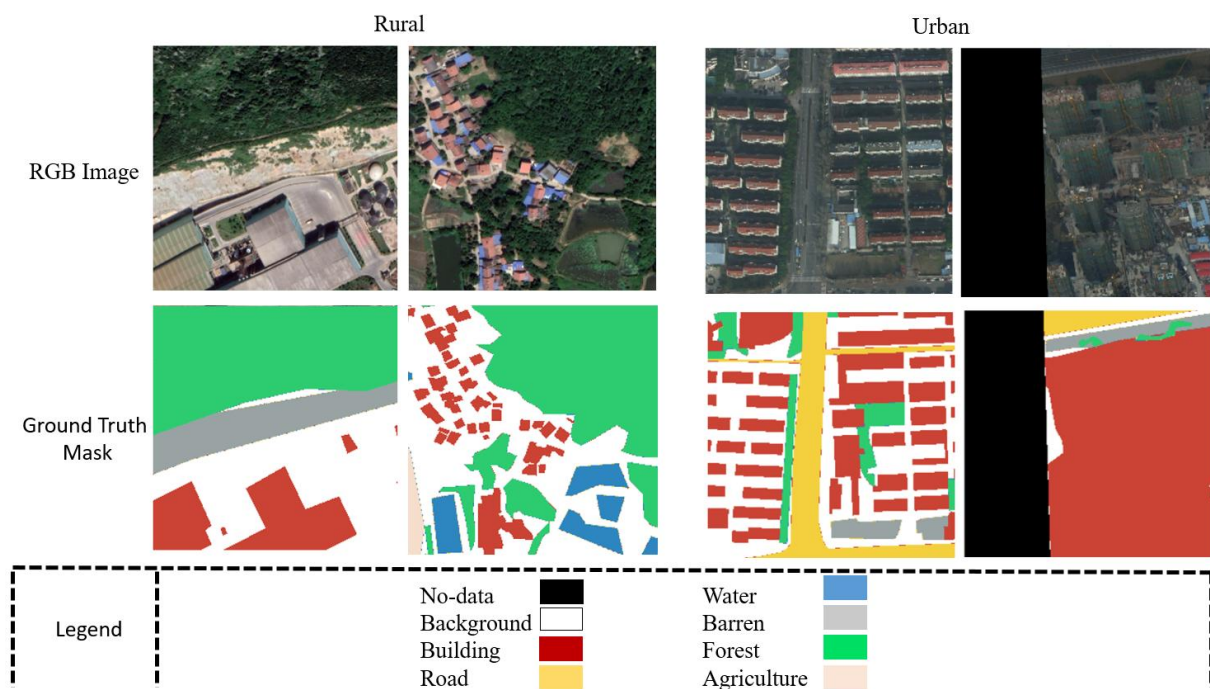


Figure 1. Example images and corresponding segmentation masks from the LoveDA Dataset

Wang et al. (J. Wang et al., 2021) emphasized that objects of the same category have completely different patterns in different scenes. In their dataset, urban and rural scenes have different class distributions. Urban scenes with high population density contain many artificial objects such as buildings and roads, while rural scenes contain more natural elements such as forests and water.

2.2. Method

The models, loss functions and optimization algorithms used for comparison in the study are briefly mentioned in this section.

2.2.1. Models

CNN architectures are able to distinguish land covers by classifying pixels thanks to their ability to learn spatial relationships in images. UNet, UNet++, MANet, and DeepLabv3+ DL models, which are frequently preferred in the literature for segmentation in HSR data, are briefly described in this section.

UNet:

The UNet architecture (Ronneberger et al., 2015) consists of an encoder block that coarsens the image resolution and a decoder block that increases the image resolution. The UNet architecture consists of convolution layers, pooling layers, feature fusion layer, upsampling layer and softmax layer. The structure of the encoder and decoder parts is symmetric with jump connections between them, which positively affects fine-grained segmentation. At the same time, UNet is able to preserve the feature maps with the same size as the original image.

UNet++:

UNet++ (Zhou et al., 2018) is developed by integrating multi-depth UNet(Ronneberger et al., 2015) models and linking all encoder and decoder blocks with the same resolution. The decoders in the UNet++ architecture combine multi-scale feature maps at the same and different image resolutions. Interwoven with jump links, this architecture exploits the UNet structure at different depths. Thus, low-level attributes are preserved in complex tasks. Furthermore, these jump links help to bridge the semantic gap between the attribute maps of encoders and decoders in the UNet model.

MANet:

Multiscale attention network, MANet (Fan et al., 2020), is a multiscale network that can capture local attributes with their global dependencies based on the attention mechanism. It consists of two blocks: position-wise attention block (PAB) and multiscale fusion attention block (MFAB). The detection of spatial and channel dependencies of the feature maps of the PAB and MFAB blocks is based on the self-attention mechanism. While the PAB is used to model attribute dependencies capturing spatial dependencies between pixels, the MFAB handles channel dependencies between any attribute map with multiscale semantic attribute fusion. In addition to high-level, MFAB also takes into account the channel dependencies of low-level attribute maps. Finally, the channel dependencies of high and low level attribute maps are fused to obtain multiscale semantic information and improve network performance.

DeepLabv3+:

DeepLabv3+ (Chen et al., 2018) is roughly based on adding a decoder module to DeepLabv3 (Chen et al., 2017) and extending DeepLabv3 to improve segmentation results along object boundaries. DL uses a spatial pyramid pooling module or an encoding-decoder structure for semantic segmentation. The spatial pyramid

pooling module can multi-scale encode incoming features at multiple rates with filters or pooling operations, while the encoder-decoder architecture can capture sharper object boundaries by gradually recovering spatial information. DeepLabv3+ combines the advantages of both approaches. By applying depth-separable convolution to both atrous spatial pyramid pooling and decoder modules, DeepLabv3+ achieves a faster and more powerful encoder-decoder network.

2.2.2. Losses

Loss functions play an important role in determining model performance. However, it is not possible to decide on a single universal loss function, especially for multi-class segmentation problems (Jadon, 2020). In this study, Dice Loss, Cross Entropy Loss and Weighted Cross Entropy Loss, which are widely used in segmentation applications, are compared.

Dice Loss

The Dice coefficient is a widely used metric for calculating the similarity between two images. It was later adapted as a loss function known as Dice Loss (Sudre et al., 2017). For multi-class problems, the calculation of Dice Loss over the i -th instance in the dataset is presented in Equation 1:

$$L_{Dice}(y, \hat{y}) = \frac{1}{C} \sum_{k=1}^C \left(1 - \frac{2 \sum_i (y_{i,k} \hat{y}_{i,k})}{\sum_i y_{i,k} + \sum_i \hat{y}_{i,k}} \right) \quad (1)$$

In Equation 1, y is the true value, $\hat{y}$ is the predicted result and C is the number of classes.

Cross Entropy (CE) Loss

Cross entropy (CE) is a measure of the difference between two probability distributions for a given random variable or set of events (Yi-de et al., 2004). CE loss is widely used especially in classification problems and is the most preferred loss function in segmentation, since segmentation is considered as pixel-level classification. The CEL calculation is presented in Equation 2 for the i -th sample in the dataset,

$$L_{CE}(y, \hat{y}) = - \sum_{k=1}^C y_{i,k} \log(\hat{y}_{i,k}) \quad (2)$$

where, y is the true value, $\hat{y}$ is the predicted result and C is the number of classes.

Weighted Cross Entropy (WCE) Loss

Weighted cross entropy (WCE) is a variant of binary cross entropy in which positive samples are weighted by coefficients (Pihur et al., 2007). WCE loss is calculated over the i -th sample in the dataset as in Equation 3,

$$L_{WCE}(y, \hat{y}) = - \sum_{k=1}^C w_k y_{i,k} \log(\hat{y}_{i,k}) \quad (3)$$

where y is the actual value, $\hat{y}$ is the predicted result and w_k is the weighting coefficient for the k -th class. In case of class imbalance, larger weights are assigned for rare classes.

2.2.3. Optimizer approaches

In DL algorithms, optimizers are used to obtain the best match between the actual values and the estimated outputs (Bottou, 2010). The difficulties in satellite imagery can be taken into account in the optimization of CNN-based segmentation, allowing the development of more accurate and generalizable models (Pan et al., 2020) (H. Wang et al., 2022). In this study, experiments were conducted with Adam and SGD approaches. An overview of these optimizers is presented in this section.

Adam

The Adam optimizer (Kingma & Ba, 2014) is an iterative optimization algorithm used to minimize the loss function during the training of neural networks. In Adam optimizer, faster convergence is guaranteed by bias in the early stages of training and the training process is stabilized. Adam helps faster convergence by adjusting the learning rate for each parameter and is therefore preferred for problems with sparse gradients or noisy data (Goodfellow, 2016).

Stochastic Gradient Descent (SGD)

Gradient descent (GD) is one of the basic optimization algorithms used to minimize the loss of the model. GD tries to minimize the loss function in the learning process. The model parameters are updated and the learning process is performed. SGD (Bottou, 2010) introduces randomness into the optimization process by randomly selecting a data point at each step instead of using all data points when calculating the gradient. The size of the batch size, the frequency of parameter updates and the convergence process are determined in the gradient descent (Kingma & Ba, 2014).

3. RESULTS

In this section, evaluation criteria and experimental results are presented. Resnet50 (He, 2016) was used as the backbone in all models and the input size was set to 256x256 to reduce computational complexity. In all scenarios pretrained Resnet50 weights were used and fine-tuned according to the datasets.

3.1. Evaluation criteria

In segmentation tasks, accuracy is a commonly used metric to evaluate the proportion of correctly classified pixels to the total number of pixels in the image. However, accuracy can be misleading in imbalanced datasets, where the model might classify the majority class correctly while neglecting smaller, less frequent regions. So such as IoU (Intersection over Union) to get a more comprehensive understanding of model performance, especially in cases of class imbalance.

IoU metric is commonly used to evaluate the performance of predicted regions in segmentation tasks (J. Wang et al., 2021). IoU is computed as the ratio of the intersection of the predicted and GT segmentations to the union of these areas. This metric provides a quantitative measure of how well the predicted masks align with the GT masks.

In this study, detailed results are presented on iou and accuracy is used as a support.

The formulas for Accuracy and IoU are presented in Equation 4-5, respectively.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$IoU = \frac{TP}{TP + FP + FN} \quad (5)$$

The definitions of True Positive (TP), False Positive (FP), and False Negative (FN) in Equation 4-5 are as follows: TP are pixels for which the model correctly predicts the target region. FP are pixels where the model positively predicts a region that does not actually exist. FN are pixels where the model incorrectly predicts a region with a positive true label.

3.2. Experimental results

In this section, the results of the experimental studies conducted separately for rural and urban categories are presented and compared using Intersection over Union (IoU) as the evaluation metric.

3.2.1. Rural Dataset

This section presents the results for the rural regions of the dataset. Figure 2 illustrates the IoU values of four models and three loss functions on the test data, using the Adam optimizer, while Figure 3 presents the IoU values obtained with the SGD optimizer.

Rural landscapes exhibit gradual transitions between land cover types such as forest, agriculture, and barren land. These transitions often lack clear geometric boundaries, requiring models to generalize broader spatial patterns rather than focus on sharply defined features. In this context, as illustrated in Figure2, UNet++ consistently outperforms the other models across all loss functions, particularly under Dice Loss (IoU = 0.6601). This suggests that UNet++'s nested architecture and dense skip connections are well-suited for modeling the multi-scale, continuous transitions often found in rural regions such as irregular forest edges or variable agricultural fields. Conversely, models like DeepLabv3+ and MANet, which rely on dilated convolutions and attention mechanisms respectively, may struggle with such gradual class boundaries, potentially overemphasizing texture or failing to capture context over large areas.

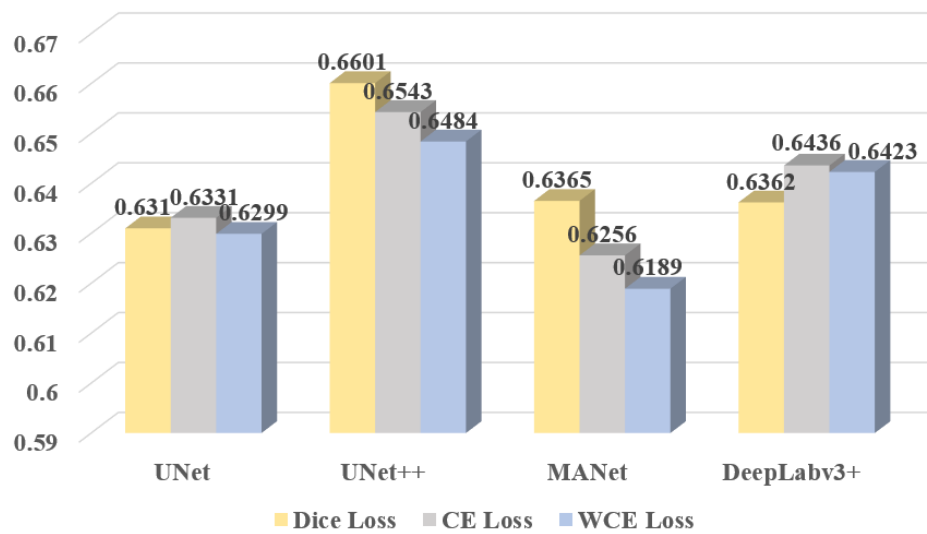


Figure 2. *IoU performance on the rural dataset using the Adam optimizer*

Dice Loss outperforms CE and WCE, particularly for UNet++ and UNet, likely due to its sensitivity to class imbalance and ability to focus on region-level overlap. This is important in rural areas where classes like barren or roads are often underrepresented and exhibit ambiguous visual features.

Overall, model architectures that allow for deep, context-aware feature fusion and fine-scale localization perform better in rural domains, where semantic boundaries are not as crisp, but broader regional patterns carry strong class cues.

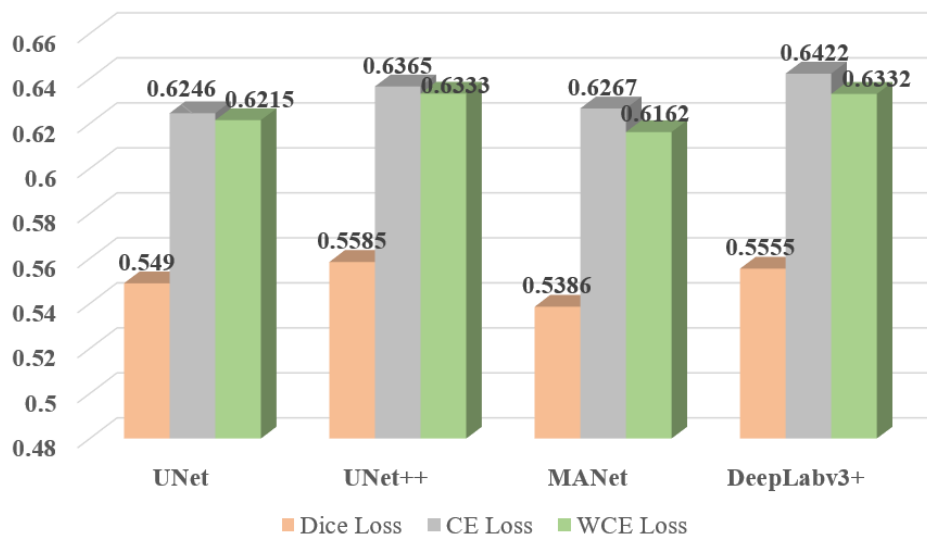


Figure 3. *IoU performance on the rural dataset using the SGD optimizer*

In rural area segmentation tasks, the performance of models trained with the SGD optimizer shows (Figure 3) notable variation depending on both architecture and loss function. Among the evaluated models, DeepLabv3+ achieved the highest IoU score of 0.6422 when trained with the standard CE loss, indicating its strength in handling broad and visually ambiguous rural land classes. UNet++ closely follows, particularly excelling under WCE and Dice Loss, which highlights its robustness in managing class imbalance and capturing multi-scale

contextual features. MANet, on the other hand, consistently underperforms across all loss types, suggesting that its attention mechanisms may be less effective in rural settings where class boundaries are less distinct and visual patterns are more heterogeneous.

While SGD provides reasonable performance, especially with CE-based training, it appears less capable of handling the nuanced spatial transitions and imbalanced class distributions characteristic of rural environments compared to adaptive optimizers like Adam. Rural landscapes pose unique challenges for segmentation due to their class ambiguity, spatial variability, and class imbalance. Our experiments reveal that optimizer choice has a significant impact on model performance in such settings. Specifically, Adam optimizer consistently yields higher IoU scores across all models and loss functions, with UNet++ showing the largest gain (10% improvement in Dice Loss). This suggests that adaptive optimization strategies are better suited to handling the complex, diffuse class boundaries and imbalanced distributions typical of rural regions. In contrast, SGD results in lower Dice-based IoU scores, possibly due to its fixed learning rate and sensitivity to sparse gradients, which are common in rural image segmentation tasks.

In all experiments; for Adam optimizer, we use 0.003 as learning rate and 0.0001 as weight decay while For SGD optimizer, we use 0.05 as learning rate and 0.9 as momentum.

When the aforementioned figures are evaluated, it is clear that the highest performance of 0.66 is obtained in the UNet++, Dice loss, Adam optimizer combination. The GT and predict masks of this combination on the test data are presented in Figure 4 along with the RGB images.

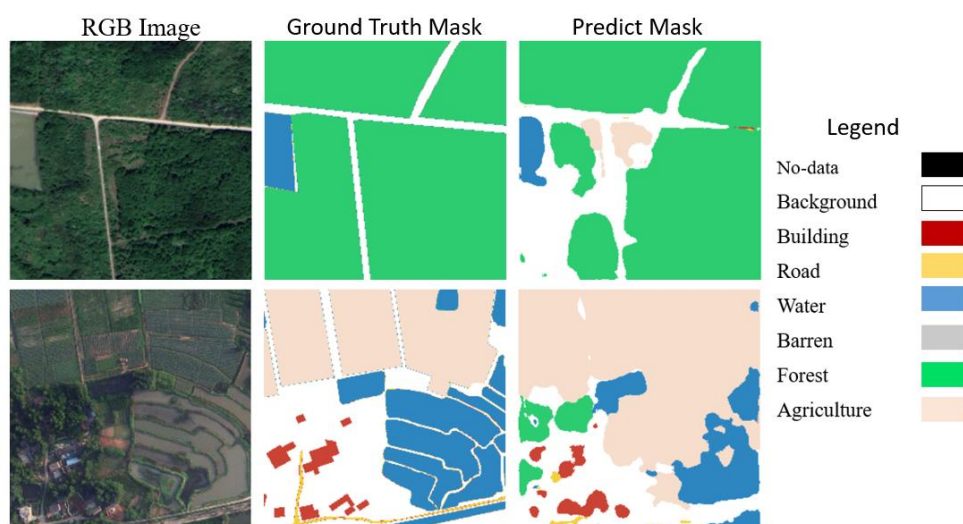


Figure 4. RGB images, GT, and predicted segmentations from the rural dataset

Figure 5 displays the confidence interval (CI) plots for both the IoU and accuracy metrics on the rural test set. These plots offer a statistical perspective on the model's performance, illustrating the range of values that can be expected for each metric with a specified level of confidence. To enhance visual interpretability, the graphs are summarized. Each point in the plot corresponds to a batch, with each batch representing 16 test samples.

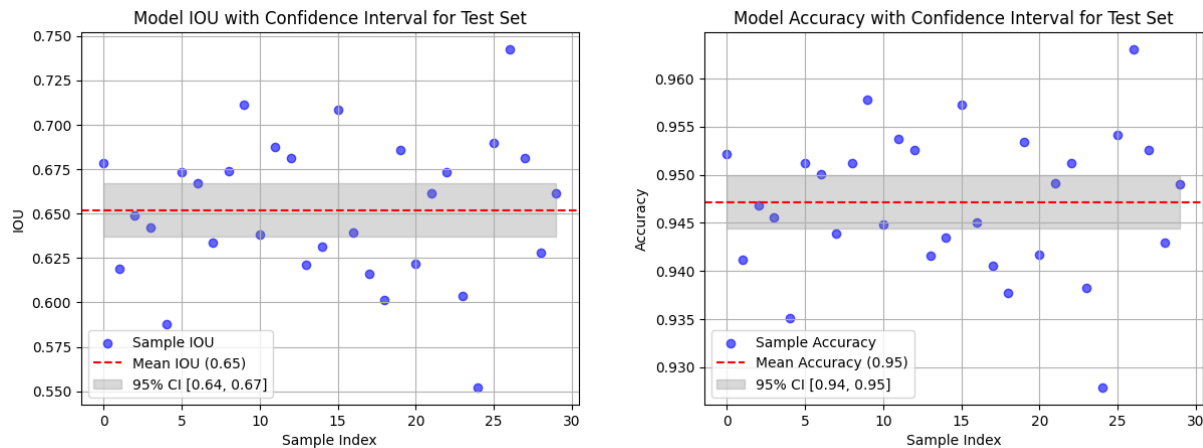


Figure 5. CI analysis using different metrics on the rural test set

The complexity matrix for analyzing the pixel-based correct and incorrect decisions of the proposed model is presented in Figure 6.

As shown in Figure 7, the class-based accuracy values indicate that the segmentation model performs at a relatively high level across most categories. Notably, the "no-data" class exhibits exceptional performance with an accuracy of 99.67%, highlighting the model's proficiency in handling missing or unlabeled areas. Furthermore, the "forest" class, with an accuracy of 83.37%, stands out as one of the best-performing classes, suggesting effective segmentation of forested regions. The "water" and "agriculture" classes also demonstrate solid results, with accuracy rates of 83.08% and 82.25%, respectively, reflecting the model's capability to accurately identify these land cover types. In contrast, the "barren" class shows the weakest performance, with an accuracy of only 60.66%, indicating that further improvements are needed in this area. The "building" (80.20%) and "road" (69.35%) classes exhibit moderate accuracy, revealing room for improvement in distinguishing these features from other land types. Overall, while the model achieves high accuracy for most classes, there is clear potential for enhancement, particularly for more challenging categories such as "barren".

Grad-CAM (Gradient-weighted Class Activation Mapping) visualizes which areas the model is focusing on. The areas marked with vibrant colors represent the regions that have the most influence on the model's decisions. This map helps us understand which features the model is concentrating on while making its predictions.

As shown in Figure 8 the rural test image, the Grad-CAM heatmaps show distinct and well-localized attention patterns for high-performing classes such as forest, water, and agriculture. The model's attention in these cases aligns closely with semantically meaningful regions for instance, forest areas activate dense regions with tree coverage, and agriculture areas correspond to regular, patch-like field textures.

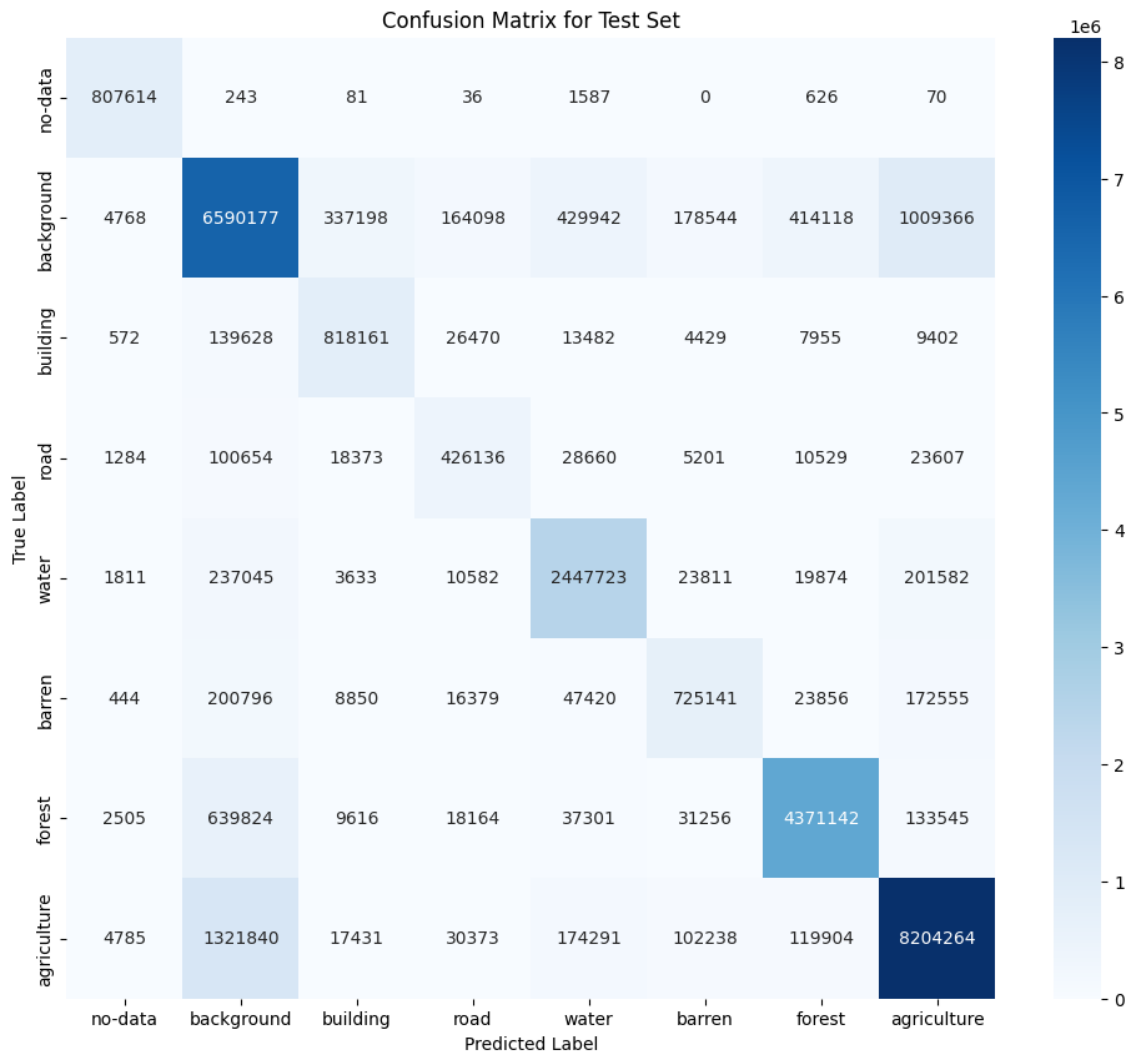


Figure 6. Confusion matrix of the rural test set

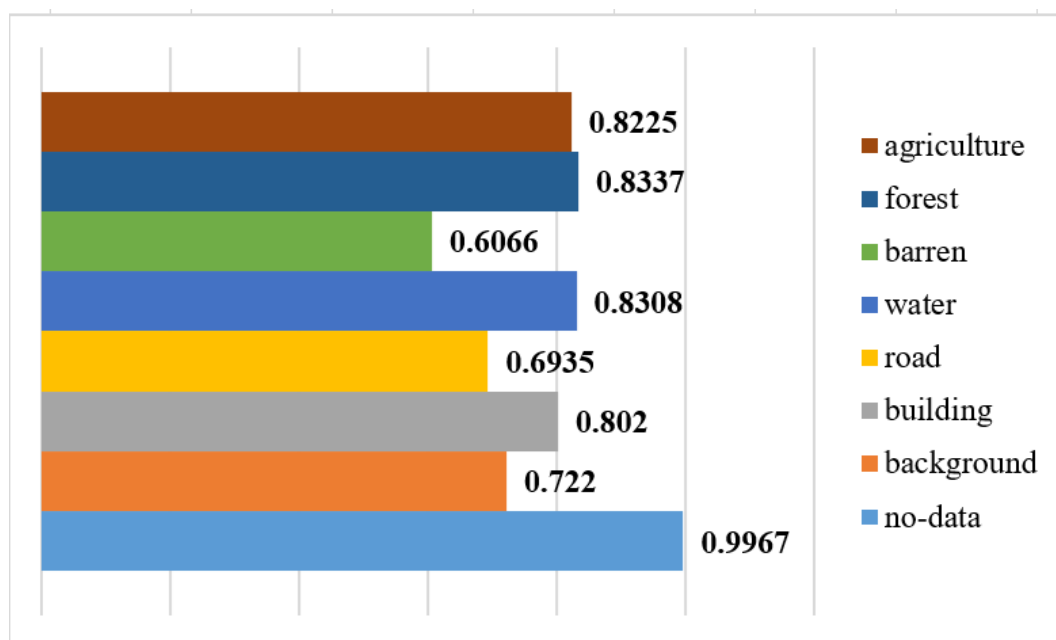


Figure 7. Class-wise Accuracy performance of the rural segmentation model

Classes with lower accuracy, such as barren and road, exhibit more scattered or ambiguous Grad-CAM activations. For example, road segments in rural areas may be less defined (e.g., dirt paths), leading to broader or less confident attention distributions in the visualizations. Similarly, barren regions often overlap visually with agricultural or background areas, making it harder for the model to consistently focus on them.

3.2.2. Urban Dataset

After rural dataset models are trained, it is obvious some models give worse than other results. Therefore, we omit some models in order to decrease time spent for training and GPU resources. According to result, we prefer only use Adam optimizer instead of Adam and SGD because most models having SGD as optimizer has imbalance class prediction or poor focal loss result (increasing gradually instead of decreasing). Also, SGD has slow converge compared to Adam. Also, we prefer not to use UNet because we have already used UNet++, and UNet++ is more complex and advanced version of UNet. Figure 9 shows the IoU values of three models and three loss approaches on the test data according to the Adam optimizer.

According to Figure 9, in the segmentation of urban regions, the models trained with the Adam optimizer display moderate performance, with IoU scores ranging between 0.5172 and 0.5717. Among the tested architectures, UNet++ achieves the highest IoU across all three loss functions, with its best performance observed under Dice Loss (0.5717). This result reflects the architecture's ability to capture multi-scale spatial features and preserve fine-grained details essential characteristics in urban settings where class boundaries such as roads, buildings, and water bodies are relatively well-defined. DeepLabv3+ and MANet follow closely, though their scores remain slightly lower, especially when trained with WCE loss, which may overcompensate for class imbalance in well-structured urban environments. The marginal differences in performance also suggest that while urban areas provide clearer visual cues for segmentation, accurately delineating narrow or adjacent classes (e.g., roads vs. buildings) remains challenging. Adam's adaptive learning capabilities contribute to stable convergence and generally consistent results across architectures, though the overall IoU levels indicate room for further enhancement, particularly through architectural improvements or targeted post-processing. As a result, the combination of UNet++, Dice loss, and Adam optimizer was the most successful model with 0.5717 IoU in the urban dataset as in the rural dataset.

The GT and predict masks of this combination on the test data are presented in Figure 10 along with the RGB images.

Figure 11 presents the CI plots for both the IoU and accuracy metrics on the urban test set. As in Figure 6, each point in the graph represents a batch, with each batch corresponding to 16 test samples.

The complexity matrix for analyzing the pixel-based correct and incorrect decisions of the proposed model is presented in Figure 12.

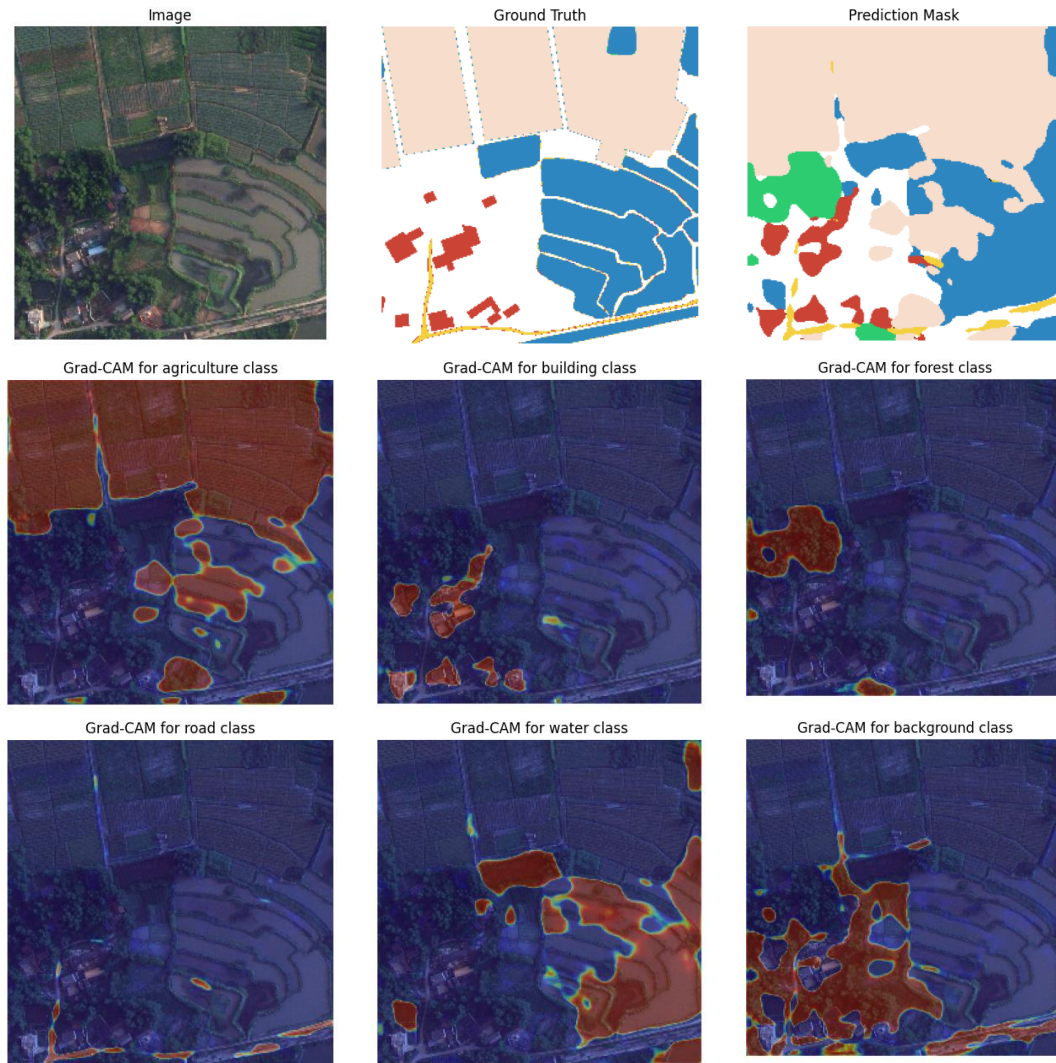


Figure 8. Grad-CAM visualization for a rural test image

As presented in Figure13, the per-class accuracy results for the urban region reveal considerable variation in model performance across land cover types. The model achieves relatively high accuracy for water (84.56%), road (76.83%), building (75.21%), and forest (75.13%), suggesting that these classes possess distinct spectral or structural characteristics that the model can effectively learn and distinguish. The no-data class (99.88%) is classified with near-perfect accuracy, likely due to its clearly separable visual traits from other categories. In contrast, accuracy is notably lower for the background class (64.75%) and dramatically low for agriculture (6.25%). The background class likely suffers from high intra-class variability and semantic overlap with adjacent classes, leading to increased confusion. The extremely low performance for agriculture may indicate several challenges: agricultural regions in urban environments are often fragmented, spectrally similar to other vegetative classes (e.g., forest or gardens), and underrepresented in the training dataset. This underperformance points to the difficulty of learning robust representations for agriculture in heterogeneous urban landscapes.

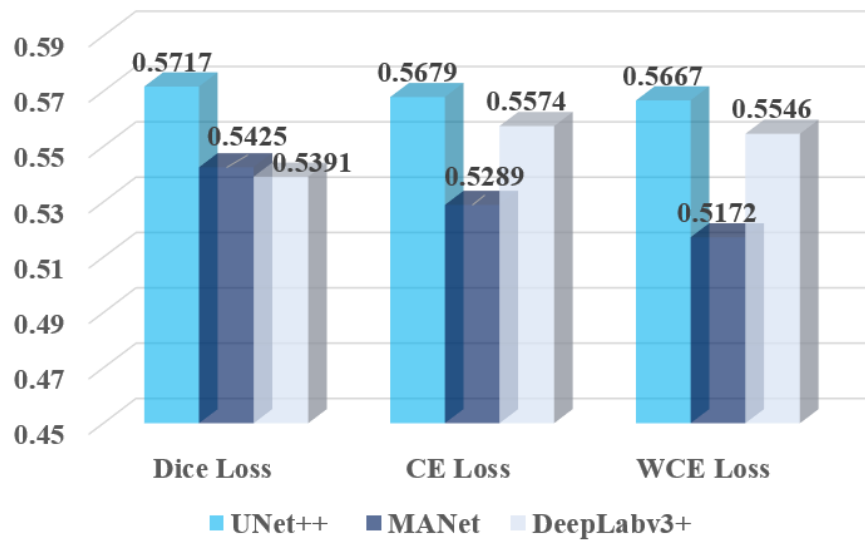


Figure 9. IoU performance on the urban dataset using the Adam optimizer

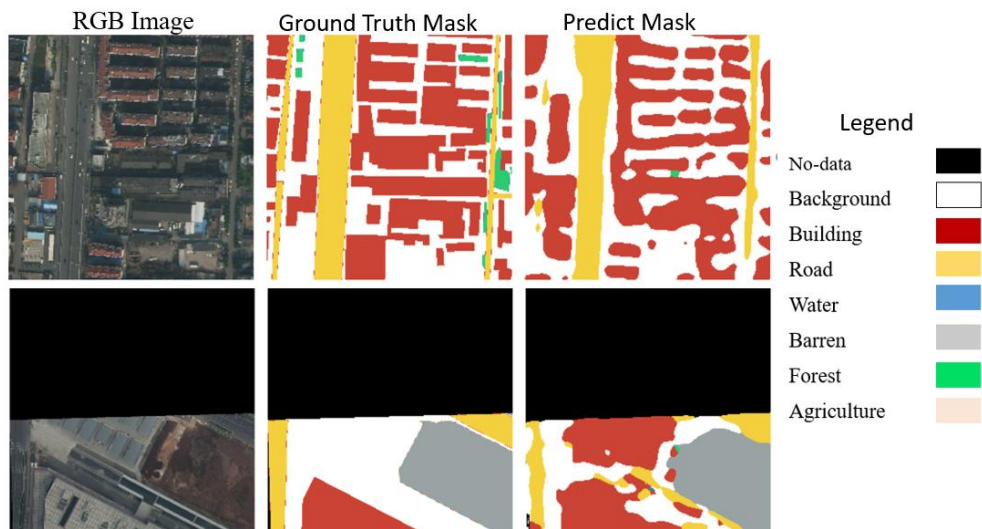


Figure 10. RGB images, GT, and predicted segmentations from the urban dataset

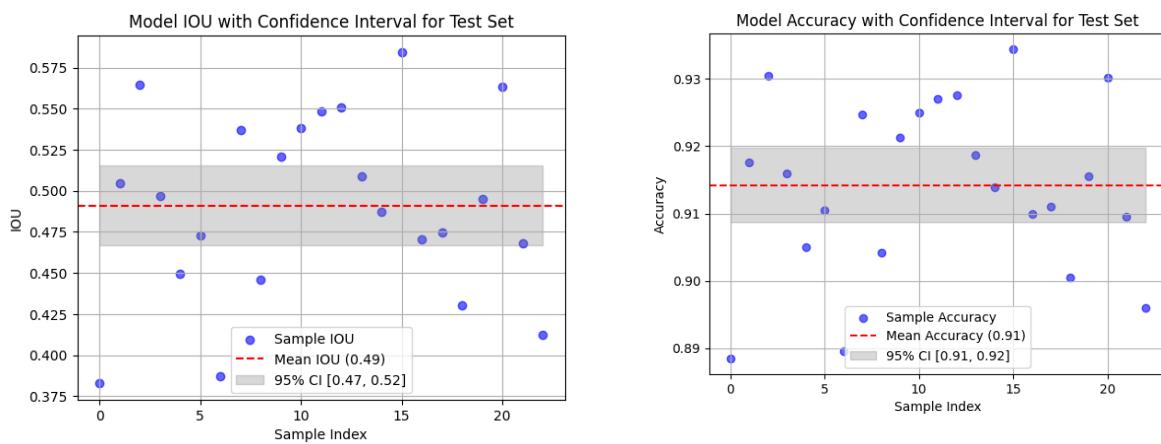


Figure 11. CI analysis using different metrics on the urban test set

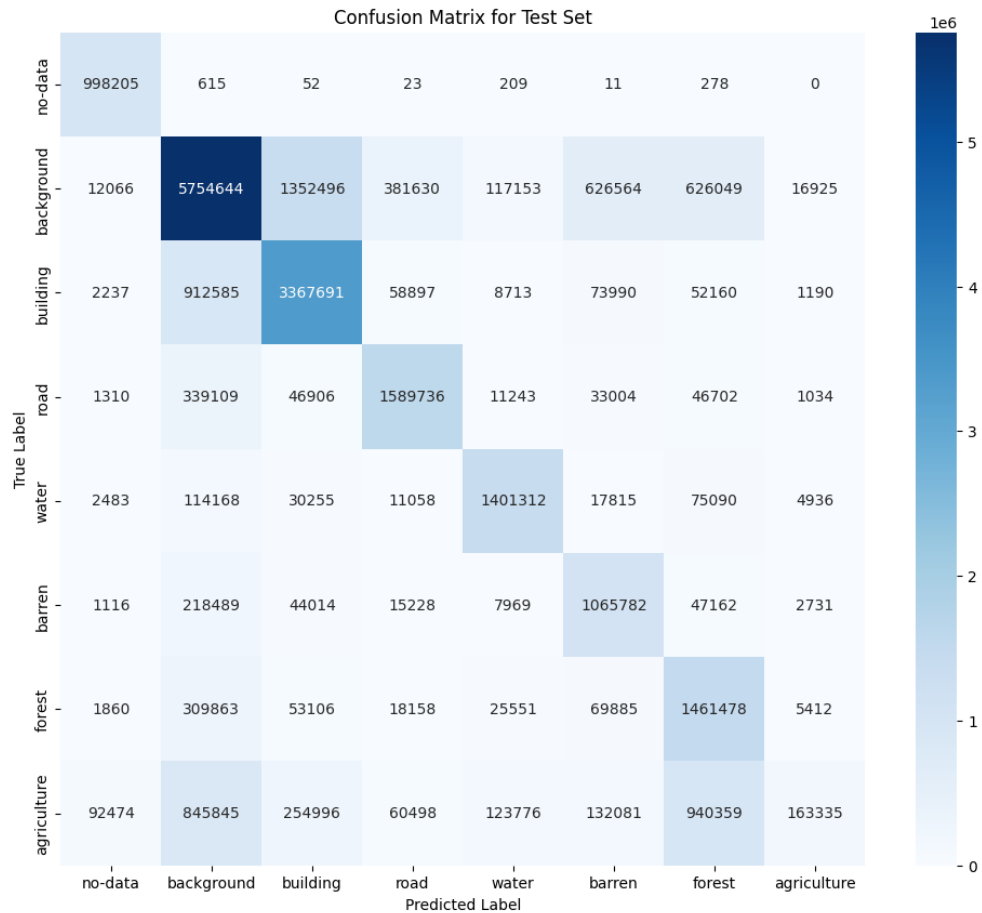


Figure 12. Confusion matrix of the urban test set

Overall, the model demonstrates strong classification performance for well-defined, structurally distinct classes, while accuracy declines sharply for classes with ambiguous boundaries or sparse representation. These findings highlight the importance of addressing data imbalance and intra-class variability when mapping high-resolution urban environments.

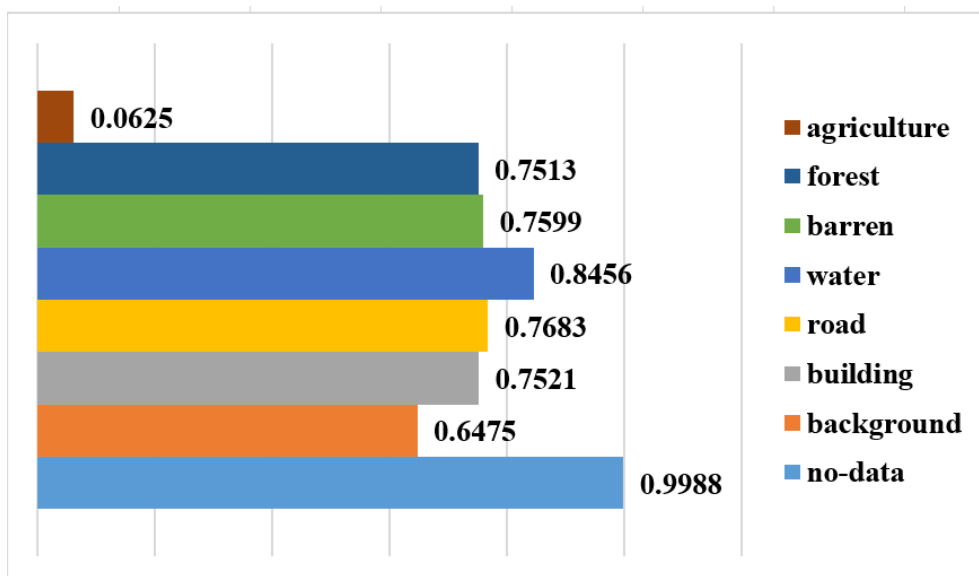


Figure 13. Class-wise Accuracy performance of the urban segmentation model

In urban areas, the model's attention was strongly focused on man-made structures such as roads and buildings, indicating that it successfully learns discriminative features relevant to urban classification. Generally, the Grad-CAM outputs for the urban test images show stronger, more concentrated activations for building, road, and water classes. Buildings and roads tend to have clear geometric edges and consistent textures in urban settings, which the model can more easily detect. This is visible in the Grad-CAM maps, where high-attention areas align closely with structured city blocks and well-defined roads.

Interestingly, despite the relatively high accuracy for barren and water classes, agriculture shows a drastically low performance. The corresponding Grad-CAM output reveals that the model does not effectively focus on agricultural zones, likely due to their visual similarity with barren or undeveloped land in urban fringe areas. This lack of focused attention is a key reason behind the low classification accuracy and supports the need for further domain-specific refinement. Figure 14 illustrates the activation map generated using Grad-CAM for a urban image.

4. DISCUSSION

Timely and accurate information on urban and rural land covers is critical for authorities. Urban and rural area analysis plays an important role in land cover change, population forecasting, environmental management, disaster management. (Guo & Du, 2017) HSR land cover data contribute to our interpretation of the geographical and ecological environment through remote sensing technology. These data often provide detailed information on the use of a particular area. Datasets in the literature are used to classify various types of land cover such as forests, agricultural areas, water bodies, settlements, etc.

In this study, we utilize the publicly available LoveDA HSR dataset, which presents challenges such as multi-scale objects, complex backgrounds, and inconsistent class distributions. The study focuses on comparing models and parameters across different geographical environments, specifically urban and rural areas. A total of 5,987 high-resolution (0.3 m) remote sensing images from Nanjing, Changzhou, and Wuhan cities are used for semantic segmentation tasks. Various models, loss functions, and optimization techniques are applied to perform land cover segmentation in both rural and urban settings. The study investigates the impact of different model configurations and parameter combinations on segmentation performance. Additionally, the influence of the geographic differences between rural and urban areas on segmentation success is explored, along with how the selected parameters reflect these differences.

Significant differences were observed between the segmentation results obtained from rural and urban areas using the models employed in this study. It was found that the type of area (rural or urban) influenced the performance of the DL models; however, the same model (UNet++) achieved higher performance in rural areas. This highlights the importance of optimizing the same model for different geographical contexts. The effect of optimization algorithms (Adam and SGD) on segmentation performance demonstrated notable differences in both rural and urban areas. The Adam optimizer exhibited faster and more stable convergence

in rural areas, which proved advantageous for optimizing the more complex and irregular land cover structures typically found in rural environments. The choice of loss function (Dice loss, CE loss, and WCE loss) also had a significant impact on segmentation success. Specifically, it was observed that the Dice loss function, when combined with the Adam optimizer, outperformed the SGD optimizer, particularly for smaller area segments. As a result, experiments in urban areas were conducted exclusively using the Adam optimizer. For both rural and urban areas, the combination of UNet++, the Adam optimizer, and Dice loss yielded the highest performance.

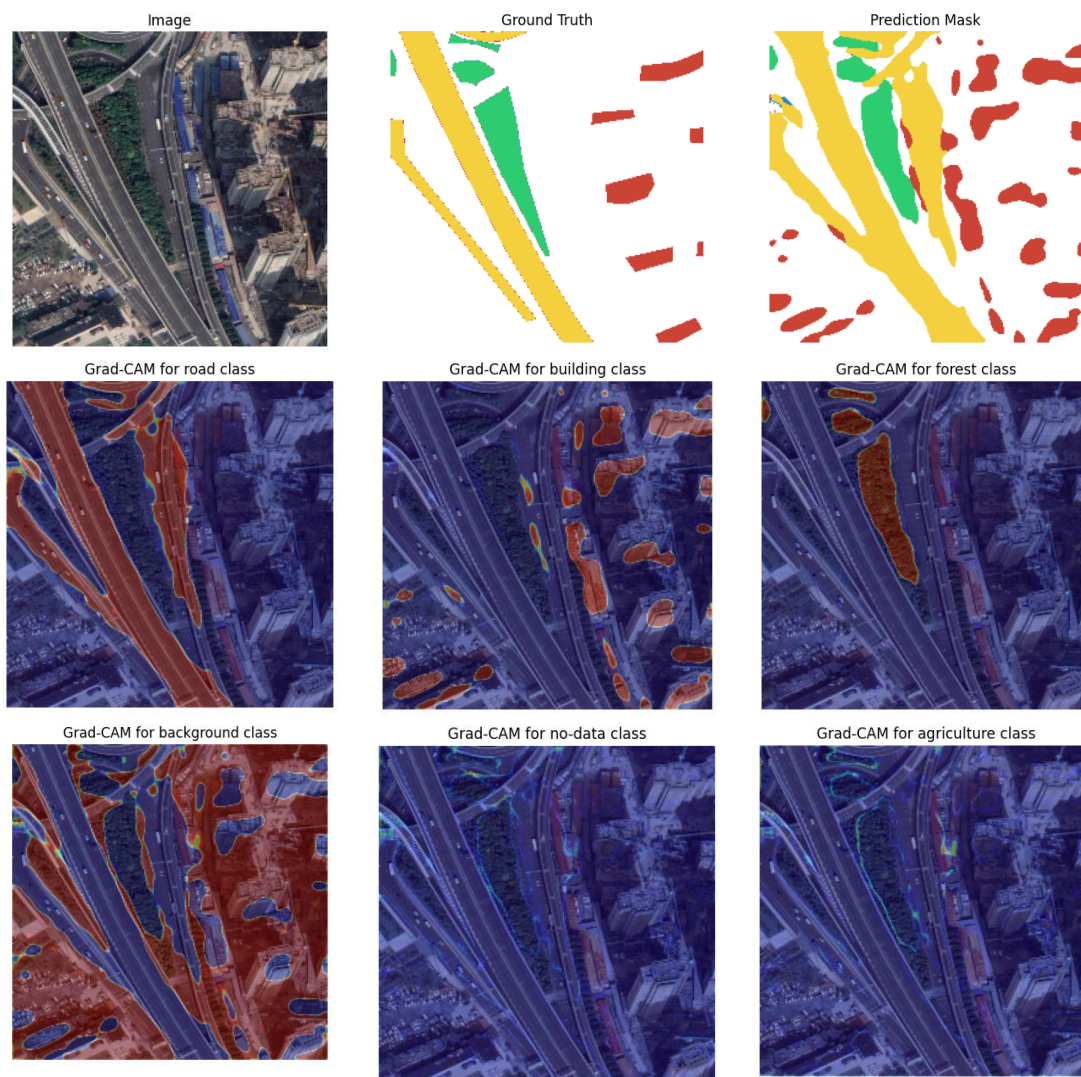


Figure 14. Grad-CAM visualization for an urban test image

In this study, ResNet50 was selected as the backbone architecture due to its strong balance between depth and computational efficiency, making it a widely adopted standard in various remote sensing and semantic segmentation tasks. Its residual connections enhance gradient flow and facilitate learning deeper representations, which are essential for capturing hierarchical features in complex scenes. Furthermore, all images were resized to 256×256 pixels prior to training. This resizing was a practical necessity to ensure feasible training under limited GPU resources and to maintain a consistent input size across batches. While

this transformation may introduce some degree of spatial information loss—particularly affecting small or elongated objects—visual inspection and empirical performance indicate that key semantic structures remain sufficiently preserved. Nevertheless, this trade-off is acknowledged as a limitation, and future work may explore adaptive tiling or patch-wise strategies to retain higher spatial fidelity without compromising computational feasibility.

To compare the performance of the LoveDA dataset with the existing literature, a weighted average of the combination (UNet++, Adam optimizer, and Dice loss) was calculated based on the number of test samples for both rural and urban groups. The final performance, in terms of mean Intersection over Union (IoU), was 62.14%. Table 2 presents a comparison of the proposed method with UNet-based techniques from the literature.

Table 2. Comparison of the proposed method with existing methods in the literature

Literature	Model	mean IoU (%)
(J. Wang et al., 2021)	UNet	47.84
(L. Wang et al., 2022)	UNetFormer	52.40
(Dimitrovski et al., 2024)	UNet Ensemble	57.36
The proposed method	UNet++	62.14

Wang et al. (J. Wang et al., 2021) obtained 47.84% mean IoU with UNet when introducing the LoveDA dataset. Wang et al. (L. Wang et al., 2022) achieved a mean IoU score of 52.40% with UNetFormer, a version of UNet combined with transformer attention mechanisms. Dimitri et al. (Dimitrovski et al., 2024) utilized a UNet ensemble model with three different backbones (Multi-Axis Vision Transformer, ConvFormer, and EfficientNet) to achieve a mean IoU of 57.36%. In this study, the combination of UNet++, which is a more advanced version of UNet with a hopping link structure, with Dice score and Adam optimizer proposed a mean IoU of 62.14%.

The comparative analysis between urban and rural regions highlights the sensitivity of classification performance to contextual landscape characteristics, particularly in the context of high spatial resolution (HSR) imagery. Rural areas exhibit notably higher accuracy in classes such as agriculture (82.25%), forest (83.33%), and background (72.20%), which can be attributed to the relative spectral homogeneity and spatial consistency of these land cover types in non-urban settings. In contrast, agriculture in urban regions is classified with extremely low accuracy (6.25%), underscoring one of the key challenges in HSR mapping the difficulty of detecting fragmented, spatially sparse, and spectrally ambiguous land cover patches. In urban environments, agricultural areas are often interspersed with other vegetated classes (e.g., parks, roadside greenery), and their small patch size further complicates the model's ability to learn distinctive representations. Likewise, building (80.20%) and water (83.08%) classes achieve better performance in rural settings compared to their urban

counterparts (75.21% and 84.56%, respectively). On the other hand, roads are more accurately classified in urban regions (76.83%) than in rural ones (69.35%), likely due to their more regular, linear structure and higher frequency of representation in urban high-resolution imagery. Interestingly, the barren class performs substantially better in urban areas (75.99%) than in rural ones (60.66%), which may reflect the more visually coherent and structured nature of urban barren areas such as construction sites or vacant lots, as opposed to the more heterogeneous barren landscapes in rural zones.

These findings emphasize the specific challenges of HSR mapping, especially in urban environments. Class confusion due to spectral overlap, high intra-class variability, fragmented spatial patterns, and limited class representation all contribute to reduced model performance for certain land cover types. Therefore, it is crucial to consider these HSR-specific limitations when interpreting classification results, and future work may benefit from incorporating class-aware sampling strategies, multi-modal data sources, or architecture-level enhancements to better address these issues.

The limitations of this study include the following: GT labels for the test data are not available in the dataset, necessitating the random generation of the test partition from the available data. Additionally, some of the images in the dataset exhibit spatial correlation, and the scenario was executed hierarchically due to the large number of experiments across all possible combinations.

5. CONCLUSION

DL approaches for land cover segmentation with HSR data yield more accurate, precise, and efficient results compared to manual methods such as visual interpretation or fieldwork. HSR data, with its high resolution and rich visual details, serves as a crucial data source for accurately distinguishing land cover classes. This study demonstrates that in land cover segmentation using HSR imagery, factors such as the differences between rural and urban patterns, as well as variations in model architectures, loss functions, and optimization techniques, significantly impact segmentation performance. It was observed that the geographic distinctions between rural and urban areas notably influence the optimization of model parameters and the overall success of the segmentation task. Future research should focus on further refining model configurations, incorporating larger datasets, and evaluating the developed models on diverse datasets to enhance generalizability.

AUTHOR CONTRIBUTIONS

Conceptualization: H.A., N.K.; Methodology: H.A., N.K.; Fieldwork and Data Collection: N.K.; Software and Implementation: H.A.; Title and Framing of the Study: N.K.; Validation and Experimentation: H.A.; Investigation and Literature Review: N.K.; Data Curation: H.A.; Writing – Original Draft Preparation: N.K.; Writing – Review and Editing: H.A.; Visualization and Figure Preparation: H.A., N.K.; Supervision and Project Administration: N.K. All authors have read and legally accepted the final version of the article published in the journal.

ACKNOWLEDGEMENTS

This study was conducted in the TÜBİTAK-BİLGEM. We would like to express our profound gratitude to TÜBİTAK-BİLGEM.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Abunnasr, Y., & Mhawej, M. (2023). Fully automated land surface temperature downscaling based on RGB very high spatial resolution images. *City and Environment Interactions*, 19, 100110.
- Aksoy, B., Çakmak, B., & Kumbasar, N. (2023). Muğla Wildfires Burn Severity and Vegetation Difference Analysis with Remote Sensing Techniques. 2023 31st Signal Processing and Communications Applications Conference (SIU), 1–4. <https://doi.org/10.1109/SIU59756.2023.10223948>
- Baek, W.-K., Lee, M.-J., & Jung, H.-S. (2024). Land Cover Classification From RGB and NIR Satellite Images Using Modified U-Net Model. *IEEE Access*, 12, 69445–69455. <https://doi.org/10.1109/ACCESS.2024.3401416>
- Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent. *Proceedings of COMPSTAT'2010: 19th International Conference on Computational Statistics Paris France*, August 22-27, 2010 Keynote, Invited and Contributed Papers, 177–186.
- Chan, L., Hosseini, M. S., & Plataniotis, K. N. (2021). A comprehensive analysis of weakly-supervised semantic segmentation in different image domains. *International Journal of Computer Vision*, 129(2), 361–384.
- Chen, L.-C., Papandreou, G., Schroff, F., & Adam, H. (2017). Rethinking Atrous Convolution for Semantic Image Segmentation. *ArXiv*, abs/1706.05587. <https://api.semanticscholar.org/CorpusID:22655199>
- Chen, L.-C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. *Proceedings of the European Conference on Computer Vision (ECCV)*, 801–818.
- Demir, I., Koperski, K., Lindenbaum, D., Pang, G., Huang, J., Basu, S., Hughes, F., Tuia, D., & Raskar, R. (2018). Deepglobe 2018: A challenge to parse the earth through satellite images. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 172–181.
- Dimitrovski, I., Spasev, V., Loshkovska, S., & Kitanovski, I. (2024). U-Net Ensemble for Enhanced Semantic Segmentation in Remote Sensing Imagery. *Remote Sensing*, 16(12), 2077.
- Fan, T., Wang, G., Li, Y., & Wang, H. (2020). MA-Net: A Multi-Scale Attention Network for Liver and Tumor Segmentation. *IEEE Access*, 8, 179656–179665. <https://doi.org/10.1109/ACCESS.2020.3025372>
- Goodfellow, I. (2016). *Deep learning*. MIT press.

- Guo, Z., & Du, S. (2017). Mining parameter information for building extraction and change detection with very high-resolution imagery and GIS data. *GIScience & Remote Sensing*, 54(1), 38–63.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- Hossain, M. D., & Chen, D. (2022). A hybrid image segmentation method for building extraction from high-resolution RGB images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 192, 299–314.
- Jadon, S. (2020). A survey of loss functions for semantic segmentation. *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, 1–7.
- Kampffmeyer, M., Salberg, A.-B., & Jenssen, R. (2016). Semantic segmentation of small objects and modeling of uncertainty in urban remote sensing images using deep convolutional neural networks. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 1–9.
- Kemker, R., Salvaggio, C., & Kanan, C. (2018). Algorithms for semantic segmentation of multispectral remote sensing imagery using deep learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 145, 60–77.
- Kingma, D. P., & Ba, J. (2014). Adam: A Method for Stochastic Optimization. *CoRR*, abs/1412.6980. <https://api.semanticscholar.org/CorpusID:6628106>
- Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3431–3440.
- Myint, S. W., Gober, P., Brazel, A., Grossman-Clarke, S., & Weng, Q. (2011). Per-pixel vs. object-based classification of urban land cover extraction using high spatial resolution imagery. *Remote Sensing of Environment*, 115(5), 1145–1161.
- Neupane, B., Horanont, T., & Aryal, J. (2021). Deep learning-based semantic segmentation of urban features in satellite images: A review and meta-analysis. *Remote Sensing*, 13(4), 808.
- Pan, Z., Xu, J., Guo, Y., Hu, Y., & Wang, G. (2020). Deep learning segmentation and classification for urban village using a worldview satellite image based on U-Net. *Remote Sensing*, 12(10), 1574.
- Peng, C., Li, Y., Jiao, L., Chen, Y., & Shang, R. (2019). Densely based multi-scale and multi-modal fully convolutional networks for high-resolution remote-sensing image semantic segmentation. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(8), 2612–2626.
- Pihur, V., Datta, S., & Datta, S. (2007). Weighted rank aggregation of cluster validation measures: a Monte Carlo cross-entropy approach. *Bioinformatics*, 23(13), 1607–1615.
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III* 18, 234–241.
- Sudre, C. H., Li, W., Vercauteren, T. K. M., Ourselin, S., & Cardoso, M. J. (2017). Generalised Dice overlap as a deep learning loss function for highly unbalanced segmentations. *Deep Learning in Medical*

- Image Analysis and Multimodal Learning for Clinical Decision Support : Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017 Quebec City, QC,..., 2017, 240–248.
<https://api.semanticscholar.org/CorpusID:21957663>
- Sun, W., & Wang, R. (2018). Fully convolutional networks for semantic segmentation of very high resolution remotely sensed images combined with DSM. *IEEE Geoscience and Remote Sensing Letters*, 15(3), 474–478.
- Tao, A., Sapra, K., & Catanzaro, B. (2020). Hierarchical multi-scale attention for semantic segmentation. *ArXiv Preprint ArXiv:2005.10821*.
- Wang, H., Dalton, L., Fan, M., Guo, R., McClure, J., Crandall, D., & Chen, C. (2022). Deep-learning-based workflow for boundary and small target segmentation in digital rock images using UNet++ and IK-EBM. *Journal of Petroleum Science and Engineering*, 215, 110596.
- Wang, J., Zheng, Z., Ma, A., Lu, X., & Zhong, Y. (2021). LoveDA: A Remote Sensing Land-Cover Dataset for Domain Adaptive Semantic Segmentation. In J. Vanschoren & S. Yeung (Eds.), *Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks (Vol. 1)*.
<https://datasets-benchmarks>
- Wang, L., Li, R., Zhang, C., Fang, S., Duan, C., Meng, X., & Atkinson, P. M. (2022). UNetFormer: A UNet-like transformer for efficient semantic segmentation of remote sensing urban scene imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 190, 196–214.
- Xia, H., Wu, J., Yao, J., Zhu, H., Gong, A., Yang, J., Hu, L., & Mo, F. (2023). A Deep Learning Application for Building Damage Assessment Using Ultra-High-Resolution Remote Sensing Imagery in Turkey Earthquake. *International Journal of Disaster Risk Science*, 14(6), 947–962.
- Yi-de, M., Qing, L., & Zhi-Bai, Q. (2004). Automated image segmentation using improved PCNN model based on cross-entropy. *Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, 2004.*, 743–746.
- Yuan, K., Zhuang, X., Schaefer, G., Feng, J., Guan, L., & Fang, H. (2021). Deep-learning-based multispectral satellite image segmentation for water body detection. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14, 7422–7434.
- Yu, B., Yang, L., & Chen, F. (2018). Semantic segmentation for high spatial resolution remote sensing images based on convolution neural network and pyramid pooling module. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(9), 3252–3261.
- Zhang, H., Liu, Y., Li, X., Feng, R., Gong, Y., Jiang, Y., Guan, X., & Li, S. (2023). Combining remote sensing information entropy and machine learning for ecological environment assessment of Hefei-Nanjing-Hangzhou region, China. *Journal of Environmental Management*, 325, 116533.
- Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N., & Liang, J. (2018). UNet++: A Nested U-Net Architecture for Medical Image Segmentation. In D. Stoyanov, Z. Taylor, G. Carneiro, T. Syeda-Mahmood, A. Martel, L. Maier-Hein, J. M. R. S. Tavares, A. Bradley, J. P. Papa, V. Belagiannis,

J. C. Nascimento, Z. Lu, S. Conjeti, M. Moradi, H. Greenspan, & A. Madabhushi (Eds.), *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* (pp. 3–11). Springer International Publishing.

Zou, S., Fan, X., Wang, L., & Cui, Y. (2024). High-speed rail new towns and their impacts on urban sustainable development: a spatial analysis based on satellite remote sensing data. *Humanities and Social Sciences Communications*, 11(1), 1–13.



Gazi University

Journal of Science

PART A: ENGINEERING AND INNOVATION

<http://dergipark.org.tr/guj.1660364>

Removal of Boron and Boron Compounds from Wastewaters

İpek ÇELEBİ GÜRSOY^{1*} İrfan AR² ¹ Eti Mine Works Ankara, Türkiye² Gazi University, Ankara, Türkiye

Keywords	Abstract
Boron Removal Wastewater Spring Water Adsorption Natural Adsorbents	Boron is an essential micronutrient for plants and animals. However, it is harmful and poisonous in high doses. This study aimed to conduct studies to remove boron from the wastewater of Eti Mine Works using natural and modified adsorbents such as waste clays, zeolites, walnut shells, waste corn cobs and waste sunflowers. Firstly, adsorption studies were carried out with the wastewater taken from Emet boron operations directorate with waste clays obtained from Emet, Bandırma, Kırka, Bigadiç operations directorates and zeolite obtained from site of the Bigadiç operations directorate in a batch system of shaking, thermostat water bath. Adsorption studies were carried out using different amount of adsorbent, pH and particle size to determine the effects of these parameters on adsorption. The adsorption processes were found to be insufficient to remove boron from wastewater with high boron content as 5883 ppm using natural adsorbents. Therefore, it was decided to apply first chemical precipitation and then adsorption by using natural adsorbents. Secondly, it was aimed to eliminate high boron content in Espey Spring Water with high boron content in the Espey open pit site to improve the quality of irrigation water within the framework of social responsibility. Espey Spring Water's initial boron value was 3810 ppm. After two-stage pre-treatment with lime, its boron value was 297 ppm (92% removal efficiency). Then, the adsorption studies were carried out at the same conditions as in the first part of this study using HCl modified almond shells as adsorbent. The optimum parameters reached were as 8 g of adsorbent, at pH 10, at 150 rpm mixing speed, at 25 °C and at 4 hours contact time. The adsorption process consisting all these optimum parameters; 66.7% removal efficiency was reached. Almond shells were found to be effective adsorbents for boron removal from springs waters, having high boron content in the scope of social responsibility considering environment and living health. These results were found to be important from an economic perspective because of evaluating natural wastes, as almond shells, in the adsorption process for boron removal.

Cite

Çelebi Gürsoy, İ., & Ar, İ. (2025). Removal of Boron and Boron Compounds from Wastewaters. *GU J Sci, Part A, 12(2)*, 503-540. doi:10.54287/guj.1660364

Author ID (ORCID Number)	Article Process
0000-0003-0830-1441	İpek ÇELEBİ GÜRSOY
0000-0002-6473-9205	İrfan AR
	Submission Date 18.03.2025
	Revision Date 07.04.2025
	Accepted Date 20.05.2025
	Published Date 30.06.2025

1. INTRODUCTION

1.1. General Information About Boron

Boron, located in III A group of periodic tables and shown as B, has semi-metallic and semi-conducting properties having atomic number as 5 and atomic weight as 10.81. It has two separate stable isotopes called B10 and B11 in nature. The most important feature that makes boron mineral important is that of being very hard and resistant to high temperatures (Baran, 2014).

In nature, boron is never found free (Baran, 2014). It tends to bond with oxygen. It is found as minerals bound to sodium, calcium and magnesium oxides and containing crystal water. These minerals are called boron salts

*Corresponding Author, e-mail: icelebi@etimaden.gov.tr

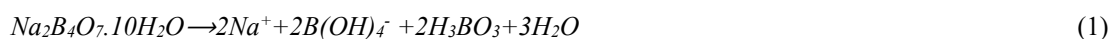
(Doğan et al., 2005). The general name of boron-oxygen compounds is borate. In this way, many boron compounds are used in different branches of industry.

Industrially important boron compounds are borax (tincal, sodium-based boron compounds), colemanite (calcium-based boron compounds) and ulexite (sodium-calcium-based boron compounds). In Turkey, sodium based tincal, calcium based colemanite and sodium+calcium based ulexite are commonly found boron minerals (Ateş, 2018).

Boron is used in many areas of life; it is widely used in glass, ceramics, agriculture, cleaning-detergent, wood protection, health and flame retardant. It is also included in armor materials through the use of boron carbide in the defence industry.

1.2. Boron Chemistry

When borates are dissolved in water, other polymeric species are present in the environment along with boric acid. For example, the following species are found in the solution of borax (sodium borate) in water (Uslu, 2010).



Boric acid and borates' solubility increases with temperature in water.

pH is an important parameter effecting the behaviour of water-soluble boron. Boron is present in water in the form of boric acid when the pH is less than 7 and in the form of borate at the pH values greater than 10.5. This situation can be seen in Figure 1.

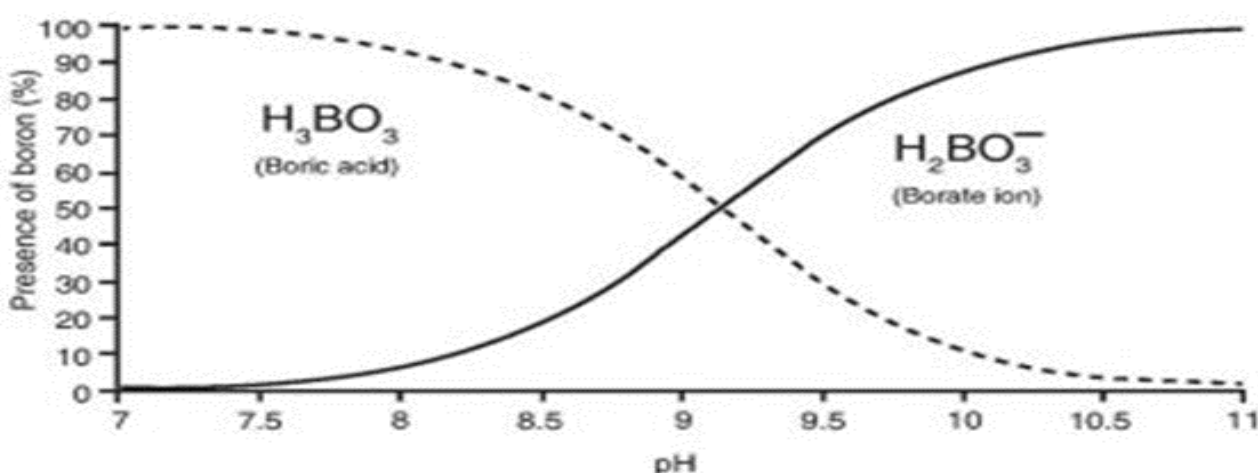


Figure 1. $H_3BO_3/B(OH)_4^-$ distribution as a function of pH (Rodarte & Smith, 2014)

That is, in any liquid system, pH value is the main parameter which determines the ratio of boric acid and borate. At high pH values, predominating ion is, the monovalent borate anion $B(OH)_4^-$, while at lower pHs, predominating one is unionized boric acid $B(OH)_3$ (Atalay Sönmez, 2014).

However, concentration is also an important parameter. The concentration constant (K) of boric acid increases with increasing concentration. If the total boron concentration in the aqueous solution is less than 0.025 M (0.1% B_2O_3 or 270 mg B/L), it is known that only H_3BO_3 and $B(OH)_4^-$ ions are present in the environment. Conversion to polyborates occurs between pH 6-11 and at higher concentrations (>0.025 M). Highly soluble polyborate ions such as $B_3O_3(OH)_4^-$, $B_4O_5(OH)_4^-$ and $B_5O_6(OH)_4$ are formed (Atalay Sönmez, 2014).

1.3. Boron in Turkey

73% of the world's boron reserves exist in Turkey. Eskişehir- Kırka, Kütahya-Emet, Balıkesir-Bigadiç and Bursa-Kestelek have known boron deposits (Yiğitbaşıoğlu, 2004; Yağmur, 2012). The most abundant boron minerals in terms of reserves in Turkey are tincal and colemanite.

With the 'Law on the Operation of Boron Salts No. 2840', the production, operation and marketing activities of boron mines in Turkey were given to the Eti Mine Works, which was established by Atatürk in 1935 under the name of Etibank. There are four operating directorates within Eti Mine Works: Kırka bor operation directorate-Eskişehir, Emet boron operation directorate-Kütahya, Bandırma boron and acid factories operation directorate-Balıkesir and Bigadiç boron operation directorate-Balıkesir (Tüzün et al., 2019). Figure 2 shows boron minerals extracted in Turkey (Ateş, 2018).



Figure 2. Bor minerals in Turkey **a)** tincal, **b)** colemanite, **c)** ulexite (Ateş, 2018)

1.4. The Impact of Boron on Environment

Air, soil and water are affected by boron therefore the effect of boron on water is evaluated as its effect on drinking water and its effect on agricultural water.

Boron occurs naturally in groundwater, also directly effects groundwater. Water filtered from the surface brings boron along with, and consequently, the composition of the groundwater is effected (Aslan, 2020).

Boron can be found in surface waters as an industrial pollutant or as a product of decayed plant materials and agricultural residues (Bilici Başkan & Atalay, 2014). Boron found in surface waters also originates from domestic wastewater (Bilici Başkan & Atalay, 2014; Aslan, 2020). Wastewaters generated during the operation of mines can cause soil and water pollution problems if precautions are not taken to ensure that. Boron concentration in these wastewaters exceeds the acceptable limit or values determined by the legislation, it can harm aquatic creatures or plants if used in agricultural irrigation and can be harmful to people who use them in their diet.

In Turkey, legislation has been created that includes various limits for the discharge of boron in order to protect the environment and also boron intake for human health. In this direction, limits for boron have been determined under three headings in the legislation. These are; “Water pollution control regulation (2004), Wastewater treatment plants technical procedures circular (2010) and Regulation on waters intended for human consumption (2005)”. The subject of classification of agricultural irrigation waters according to boron concentration has been determined by wastewater treatment plants technical procedures circular (2010). And according to the Regulation on water for human consumption (2005), the maximum boron concentration allowed in drinking water is 1 mg/L (Bilici Başkan & Atalay, 2014). Boron limit values vary depending on the receiving environment where the wastewater will be discharged. According to the ‘Water pollution control regulation (2004)’, 500 mg/L (ppm) is the discharge criterion to seawater. And the discharge criterion for wastewaters of factories producing chemical boron compounds is 50 ppm (Ünlü et al., 2013).

1.5. Boron Removal Methods

Different methods are used to remove boron from wastewater and water in order to prevent this pollution and possible damages, to use the water for agricultural purposes, and to eliminate the negative effects of the presence of boron in chemical processes. These are as reported in the study of Liu et al. (2022) as; “Chemical Precipitation and (Electric) Coagulation; Extraction Method; Membrane Processes which include forward osmosis (FO) process, Donnan dialysis (DD) process and membrane distillation (MD) process.; Adsorption Technologies which include adsorbents as activated carbon layered double hydroxides (LDHs), industrial waste materials, natural materials (such as eggshells), metal organic frameworks, and porous aromatic frameworks and other novel materials; capacitive deionization (CDI) process and electro deionization (EDI) process; Integrated Methods as AMF and the combination of the resin method and membrane processes etc.” Also the integration methods can be used as: RO separation combined with coagulation, MF/UF combined with adsorption, and complexing membrane filtration (CMF). These methods are considered to be promising techniques since being economically, ecologically profitable (Liu et al., 2022).

A study by Bara and Abba (2025) using nano-magnetite sorbent, at the adsorption section of the hybrid system, the wastewater was subjected to adsorption process for 250 minutes. The removal efficiencies of boron, turbidity, copper, and zinc were 74.39%, at 250 minutes contact time. However, this concentration is still above the standard limit (0.5-1.0 mg/L). Then, a nanocomposite membrane was employed as treatment process for

more efficiency. The boron concentration was significantly reduced to 0.43 mg/L, which is below the standard limit. Finally, it is found that the modified hybrid TiO₂ membrane is effective for boron removal from the wastewater.

Zeytuncu et al. (2023) conducted a study aimed at removing boron from industrial wastewaters and comparing the efficiency of chemical precipitation membrane filtration systems. Chemical precipitation, electrocoagulation and reverse osmosis processes were studied. Poly aluminium chloride (PACl), lime, calcium chloride (CaCl₂) and barium chloride (BaCl₂) and hydrogen peroxide as oxidant were used chemical precipitants. The study found that when solution pH was adjusted at 10, boron removal increased to 99.5%. With application of electrocoagulation, 45% of B were removed from spring water. Boron-removal increased 98% by increasing pH up to 10. DP-SWRO membranes system at pH 10 was found to be the best system. In summary, a hybrid membrane process which contains an electrocoagulation process as a pre-treatment method and a double pass RO membrane system is a very effective process for boron removal.

Bilen et al. (2018) reports that, in studies conducted with reverse osmosis, the removal efficiency was found to be 63-82% on average, and it was seen that the efficiency reached was 99% by increasing the pH to 10.5. In the reverse osmosis method, issues such as membrane stability, cost and coating of the membrane with CaCO₃ raise questions about the effectiveness of the method. In addition, this method produces a dirtier waste stream and its disposal poses a separate problem. Boron removal by electro-dialysis method has given partially successful results in synthetic solutions with low boron concentrations. However, due to the high boron concentration in real solutions, high conductivity and the effects of other ions, its application area is quite limited and it is generally used as the final treatment unit in the clean output streams of reverse osmosis treatment processes. In ion exchange studies, boron removal efficiencies in the range of 90-98% have been achieved, but there is an important problem in this method as the regeneration cost. Electrocoagulation has properties as low amount of sludge (compared to chemical coagulation), not giving extra anions to the solution (compared to chemical coagulation) and boron removal with high efficiency. This method is effective at high boron concentrations. Although these advantages are valid for synthetic solutions, this advantageous situation disappears in real solutions. Due to its high energy consumption, it is not appropriate for boron removal, especially in industrial applications. On the other hand, with its simple, industrially easy installation, wide application area and the usability of the boron-containing solid waste resulting from precipitation as raw material in the cement and ceramic industries, the chemical precipitation method is a useful process for boron removal (Bilen et al., 2018). Using hydroxyapatite, lime, and aluminium sulphate in the chemical precipitation process, 99.4% removal efficiency can be reached. Waste generation, high price, and problem with ultimate disposal are the disadvantages of these processes (Fadaei, 2022). Akkurt et al. (2022) stated in their study that as follows: “Eti Maden Kırka boron operations’ wastewater that contained 2752 ppm boron was investigated for boron removal using Ca(OH)₂ in the precipitation. The effect of pH and the amount of boron was studied.

It was found that the most important parameter was pH effect. And after adjustment of pH, addition of Ca(OH)_2 improved the precipitation. The best result was obtained in the two-stage addition of Ca(OH)_2 with the initial control of pH and in the second stage with additional use of aluminium sulphate. Under these conditions, boron content in the wastewater was lowered to 250 ppm". Also Bilen et al. (2018) stated in their study that as follows: "In the literature, general information can be found on the selection of the appropriate treatment process depending on the amount of boron contained in boron-containing water. In studies, in cases where the amount of boron oxide (B_2O_3) is more than 25 g/L ($\text{B} > 7850 \text{ mg/L}$), it is recommended to remove boron by forming low-solubility salts with substances such as Ca(OH)_2 , MgO , $(\text{NH}_4)_2\text{CO}_3$, Na_2CO_3 . It is mentioned that M(OH)_n , M: Al, Fe, Sn, Ti, Mn, Ni, Mg, Zn, Zr etc. sorbents or ion exchange resins can be used for boron removal in the situation when the amount of boron oxide (B_2O_3) is between these values as; $315 < \text{B} < 1000 \text{ mg/L}$. Finally, in cases where the amount of B_2O_3 is less than 1 g/L ($\text{B} < 315 \text{ mg/L}$), boron can be removed with various membrane or electrochemical methods according to the recent studies."

For removing boron from water and wastewater, there is no simple method (Zincir, 2013). Therefore, in ongoing projects around the world, an economical method is still sought for the removal of boron. One of the methods used for this purpose has been adsorption due to its ease of application (Zincir, 2013). Adsorption is one of the most versatile and widely used purification methods recently (Zincir, 2013). Adsorption is effective for many natural biological and chemical systems and is used in industrial applications to purify wastewater and drinking water.

However, attempts to carry out treatment with known adsorbents cause wastewater treatment costs to increase. For this reason, many researchers have directed their studies to cheap and abundant adsorbents (Zincir, 2013). Evaluating natural waste adsorbents is important from an environmental and economic perspective. In fact, some adsorbents used in boron removal can be regenerated and reused (Ateş, 2018). In addition, some adsorbents (meerscham and alumina) are added to the cement together with the boron they adsorb, thus preventing used adsorbents that are not regenerated from creating new boron pollution (Ateş, 2018).

Bhagyaraj et al. (2021) stated in their study that as follows: "Using waste-derived materials is highly significant for boron removal according to economic advantages. To determine boron adsorption performance; rice, walnut shells, wheat, various seeds, eggshells, and eggshell membrane, palm seed ash have been studied. At neutral pH, with a minor effect of reaction time and temperature on the adsorption process, the boron removal efficiency of the ash was found to be 46%. Maximum adsorption capacity of 4.23 mg/g B/L was obtained with the use of rice husks."

Liao et al. (2024) stated in their study that as follows: "Modified sugarcane bagasse was used for adsorption studies. The optimum conditions were with 5 ppm MgCl_2 addition at pH 9 according to the drinking water quality standards. The maximum boron adsorption capacity was 36 mg per 1 g char. Data fitted to Langmuir adsorption isotherm." That study was carried out with synthetically prepared boron solution.

Çelebi (2020) stated in his study that as follows: “In this study, the adsorption capacity of natural pumpkin seed shell was investigated. For the study, the effect of pH of the solution, pumpkin seed shell concentration and contact time on the adsorption process were evaluated under a constant mixing speed (150 rpm) and temperature (20 °C). Optimum conditions for adsorption were determined as adsorbent dose as 2 g/L, pH as 5.13 and contact time as 5 minutes. The maximum boron treatment efficiency of the pumpkin seed shell was obtained as approximately 88% under ideal conditions. The maximum adsorption capacity based on Langmuir isotherm is 12.61 mg/g and the adsorption of boron on pumpkin seed shell is defined by pseudo-second-order kinetic. As a result, it was noted that the boron element can be effectively removed from the aqueous solution using pumpkin seed shell.” That study was carried out with synthetically prepared boron solution having 1000 mg/L boron value.

Liu et al. (2022) stated in their study that as follows: “Natural materials have obvious advantages in relation to their sources and prices. Natural materials, such as bentonite, kaolinite, zeolite, waste calcite, wheat, rice, and walnut green shell were used in literature studies. The best performance was obtained with waste calcite and rice residue. Generally, organic adsorbents showed a higher boron adsorption capacity than mineral adsorbents. And also Liu et al. (2022) stated in their study that waste eggshells can also be used to remove boron. And also Liu et al. (2022) stated as follows: “Researchers have found that waste eggshells can also be used to remove boron successfully. This good boron adsorption capacity was attributed to the calcium oxide formed after the eggshell’s calcination, which not only adjusts the pH of water but also reacts with borate ions”.

Saleh Al-dhaw et al. (2023) stated in their study that as follows: “CRB05 was used as an adsorbent for boron removal from a synthesized solution. On boron removal, the effects of adsorbent dosage, contact time, boron concentration, and pH were investigated. It was found that pH was very important on boron removal. 98% removal was achieved with at pH 4.5, adsorbent dosage 1125 mg/L, time 255 minutes and concentration 1150 mg/L. And it was concluded that CRB05 is an effective adsorbent for boron removal from synthetic boron solutions”.

Öcal et al. (2024) stated in their study as follows: “By using different methods such as ion exchange resins, membrane processes and adsorption, boron removal has been studied from wastewater. With high operational costs, most of this methods produce excessive waste unfortunately. So, for purification of wastewater containing high concentrations of boron, more sustainable methods are required. In this study, a sustainable treatment process was developed for wastewater containing high concentrations of boron. This study investigated the removal of boron from wastewater by using $\text{Al}(\text{OH})_3$ as an adsorbent. The main goal of the study was the reuse of adsorbent ($\text{Al}(\text{OH})_3$) and the potential recovery of boric acid. It was seen that although lower concentrations of boron were obtained at pH 10.5 (980 mg/L and 635 mg/L at pH of 9.0 and 10.5, respectively), the amount of adsorbed boron at pH 9 was substantially higher (94.7 mg B/g $\text{Al}(\text{OH})_3$ and 27.8 mg B/g $\text{Al}(\text{OH})_3$ at pH of 9.0 and 10.5, respectively). This was attributed to the higher initial boron

concentrations and the formation of boric acid and polyborate complexes at pH 9.0. According to these results, polyborate species sorption was an outer sphere complex formation, which led to the desorption of boron as pH lowered. Adsorbed boron species to $\text{Al}(\text{OH})_3$ could effectively be desorbed at low pH values ($\text{pH} < 5.0$); which allows $\text{Al}(\text{OH})_3$ to be used in successive adsorption studies. Approximately 55% of boron recovery from pre-treated wastewater was possible with the effective reuse of adsorbent. Based on the amounts of chemical consumptions and boron recovery, a net profit of 2.85 \$/m³ could be obtained.”

Although boron has important micronutrient properties, it has harmful effects when taken in high doses. As a result, although boron is necessary, too much of it has a poisonous effect for living things. So, the importance of keeping the amount of boron in various water resources under control and removing boron from wastewater becomes evident.

Therefore, in this study, it is aimed to conduct studies on the removal of boron in wastewater taken from the Emet boron operations directorates affiliated to Eti Mine Works to determine the best removal method using natural and modified adsorbents and to reveal the relationship that represents the experimental data. In the studies in the literature, generally synthetically prepared boron solutions were used. Therefore, this study conducted with industrial wastewater makes it specific. Unlike the methods used in Eti Mine, the adsorption method was also tried to reduce the boron concentration in industrial boron waters below the regulation limit values of Turkey. Also carrying out adsorption with waste natural materials for purifying original boron water is also important in the sense of contributing these materials to the economy and for environmental protection. This also makes this study specific.

2. MATERIAL AND METHOD

2.1. Adsorption and Chemical Precipitation with Emet Boric Acid Factory Waste Dam

In the first part of study for adsorption studies, waste clays were obtained from Emet, Bandırma, Kırka, Bigadiç boron operations directorates. Clays were ground and 2 g of clay sample was used as adsorbent. And zeolites which exist at Bigadiç operations directorates site were obtained. The zeolites were crushed in a jaw crusher and sieved in a vibrating sieve and classified according to grain size as a result of sieve analysis. The fractionated zeolite samples were as; $D > 1\text{mm}$, $D < 0.125$, $D < 0.50$ and indexed as 1, 2, 3 respectively. In adsorption experiments 2 g of zeolite was used.

In addition, walnut shells, waste corn cobs and waste sunflower were also used in adsorption studies as natural waste adsorbents in the study.

First of all, walnut shells were broken, ground, sieved in a vibrating sieve and separated into fractions.

Two different acids, HCl and H_2SO_4 , were used to modify walnut shells. Ground walnut shells with +500 μ fraction were divided into two equal parts by weight and kept separately in 250 mL 0.1 M HCl and 250 mL 0.1 M H_2SO_4 solution for 24 hours. Then, they were filtered, washed with pure water, filtered again and then

heated at 110 °C. They were dried in the oven for 3.5 hours. The reason for choosing the +500 µ size was to repeat the adsorption process with other smaller sized walnut shells if successful removal was not achieved with this particle size. 2 g of ground walnut shells activated with HCl and H₂SO₄ were used. The aim was to understand whether the interaction with acid will had a positive contribution to adsorption. Also to understand which acid had positive effect. Studies continued with pH adjustment to see the effect of pH on adsorption. Adsorption studies were carried out in wastewater at different pH values (2,7,9,11). The pH of the initial boron wastewater was 4. 2 g of HCl-treated ground walnut shells (+500 µ) were used. Afterwards, in order to understand the effect of adsorbent amount on adsorption, adsorption studies were carried out in wastewater at pH 9 as a result of pH adjustment. 2 g and 4 g of HCl-activated ground walnut shells (+500 µ) were used.

Then waste corn cobs and waste sunflower were also used in adsorption studies. Waste corn cobs and waste sunflower stems were ground and divided into sections, treated with 0.1 M HCl and 0.05 M NaOH and kept for a 24 hours. Afterwards, they were filtered and washed with pure water, dried in the oven at 110 °C for 3.5 hours and made ready for adsorption.

2 g of waste corn cobs and waste sunflower stems were used.

2.1.1. Adsorption Studies with Emet Boric Acid Factory Waste Dam

Boron removal from the wastewater taken from the Emet boric acid factory waste dam was carried out. This wastewater had 5883 ppm boron value. The boron amount was determined according to the B₂O₃ analyse method used at Eti Mine Works according to the method in B₂O₃ Analyse Book (Eti Mine Works Quality Management System). pH was adjusted to 7.60 with 0.5 N HCl and 0.5 N NaOH. Afterwards, 5 g of mannitol added to the solution and the solution was titrated with 0.5 N NaOH until pH reaches 7.60. The amount of B₂O₃ was calculated from the consumption of NaOH. The amount of B₂O₃% (A) is calculated as a weight percentage with the formula below.

$$A = [V \times 0,017405 \times 100 \times f] \div m \quad (2)$$

Here; V: volume of sodium hydroxide solution spent in the titration after adding mannitol to the solution, mL; m: weight of the test sample, g; f: titration factor, 0.017405: Amount of B₂O₃ corresponding to 1 mL of 0.5 N NaOH. The elemental percentage [%B] of the test sample is calculated with the formula below.

$$\%B = \%B_2O_3 \div 3.22 \quad (3)$$

All adsorption studies were studied in batch system and carried out at 25 °C in a water bath for 2 hours with 170 rpm rapid mixing speed according to evaluation of the literature values (Öztürk & Kavak, 2004; Sütçü, 2005; Karahan et al., 2006; Liu et al., 2009; Bhagyaraj et al., 2021; Çelebi, 2020; Liao et al., 2024). 100 mL boron wastewater was used. Different amounts of adsorbents were used and different pH values were studied to see the effect of adsorbent amount and pH. Also effect of particle size was studied. Then, the samples were

filtered through blue striped filter paper and the filtrate was analysed by titration method, mentioned above, for boron analyses.

However, for zeolite adsorption, 1 hour; 1 hour and 45 minutes; 2 hours were also tried.

2.1.2. Chemical Precipitation with Emet Boric Acid Factory Waste Dam

In the second part of study, it was thought that it would be more appropriate to remove some of the boron in wastewater, having high boron content, by chemical precipitation with lime (Bilen et al., 2018) as a preliminary process in the removal of boron from wastewater and then to continue the removal by adsorption using environmentally friendly natural adsorbents. Boron removal from the wastewater taken from the Emet Boric Acid Factory Waste Dam was carried out, having 5883 ppm boron value, in a temperature-controlled magnetic stirrer heater a rotation speed of 760 rpm according to the literature values evaluation (Kökkılıç, 2003; Ateş, 2018; Bilen et al., 2018) at different temperatures as 25 °C, 55 °C and 88 °C. The wastewater sample had a value of 5883 ppm and the amount of boron in it was 5.83 g/L. For 200 mL water samples, studies were carried out with 1,5; 3, 6, 15, 30 g of lime. 1 hour working time, 1 hour waiting time were applied. Then, the samples were filtered through blue striped filter paper and the filtrate was analysed by titration method mentioned above to determine boron value left.

2.2. Chemical Precipitation and Adsorption with Espey Spring Water(a)

2.2.1. Chemical Precipitation with Espey Spring Water(a)

The spring waters with high boron content in the Espey open pit site belonging to the Emet boron operations directorate of Eti Mine Works, feed the Gelenbek stream in the region. This high boron content water is connected to the Susurluk basin in the following regions, negatively effecting the irrigation water quality of the basin. Since Espey Spring Water had high boron content (4590 ppm), removal of boron from this water was studied in the framework of social responsibility. Three different amounts of lime were added (5, 10, 20 g) to 200 mL Espey Spring Water samples. Studies were carried out in a temperature-controlled magnetic stirrer heater at a temperature of 25 °C and a rotation speed of 760 rpm according to boric acid factory wastewater studies mentioned above and the literature values evaluation (Kökkılıç, 2003; Ateş, 2018; Bilen et al., 2018). 1 hour working time, 1 hour waiting time were applied. Then, the samples were filtered through blue striped filter paper and the filtrate was analysed by titration method mentioned above to determine boron value left.

2.2.2 Adsorption with Espey Spring Water(a)

Afterwards, adsorption studies were carried out. 0.1 M HCl treated walnut shells were tried firstly. 2 g and 4 g of walnut shells were used in order to understand the effect of adsorbent amount on adsorption. Then it was decided to continue to the studies by adjusting pH. According to the former studies carried out with walnut shells in the scope of this study, pH was adjusted to 9. For adsorption, walnut shells in +500 μ fraction treated with 0.1 M HCl were used. 2 g and 4 g of walnut shells were used. In order to understand the effect of particle

size on adsorption, the remaining walnut shells were ground and sieved. The walnut shells in the $-63\ \mu$ fraction, which are called under-sieve, were treated with 0.1 M, 37.5% HCl for a period of 24 hours. During filtration with a vacuum filter, washing with pure water (was carried out to remove acid. Treated walnut shells were dried in an oven at $110\ ^\circ\text{C}$ for 3.5 hours. According to the results obtained with 2 and 4 g of walnut shells in the $+500\ \mu$ fraction activated with 0.1 M HCl and since the amount of substance was low, studies were carried out with 2 g of adsorbent in $-63\ \mu$ fraction. Afterwards, different pH values were studied to see the effect of pH change on adsorption in Espey Spring Water water to be treated. pH adjustment was performed with 0.1 M, 37.5% HCl and 0.5 M NaOH. The pH values studied were 11.30, which is the pH of the water after acid-treated walnut shell addition, and after pH adjustments as 12; 9; 7 and 4. For adsorption studies, 2 g of $-63\ \mu$ fraction walnut shells was used.

In order to use a different environmentally friendly adsorbent, the use of almond shells was considered. The almond shells were ground in a grinder, sieved in a vibrating sieve and classified according to particle size as a result of sieve analysis. In order to be similar to the walnut shell study, $+500\ \mu$ fraction and $-63\ \mu$ fractions were thought to be used in adsorption studies.

Almond shells in $+500\ \mu$ fraction were kept in 0.1 M HCl for 24 hours, washed with pure water to remove acid and dried. Different pH values were studied to see the effect of pH change on adsorption in wastewater. pH adjustment was made with 0.1 M, 37.5% HCl. The pH values studied were 12, 9,7 and 4. For adsorption studies, 2 g and 4 of $+500$ fraction walnut shells were used.

Then, the adsorption process was carried out with almond shells in the $-63\ \mu$ fraction, which were kept in 0.1 M HCl for 24 hours, washed with pure water to remove acid, and dried. Since there was not enough material for this fraction, the study was carried out with only 2 g. Different pH values were studied to see the effect of pH change on adsorption in wastewater. pH adjustment was made with 0.1 M, 37.5% HCl. The pH values studied were 12, 9,7 and 4.

Afterwards, it was decided to work with wood sawdust, which was a different environmentally friendly adsorbent. First of all, the boron amount of Espey Spring Water, which was previously subjected to chemical treatment and then filtered, was measured. Sawdust that was kept in 0.1 M NaOH for 24 hours and washed with pure water to remove the base then it was dried in the oven for 3.5 hours at $90\ ^\circ\text{C}$. 2 g and 4 g of sawdust were used in different adsorption studies to see the effect of amount of adsorbent on adsorption. Adsorption studies with sawdust were carried out by diluting the spring water with pure water at 1/2, 1/5 and 1/10 ratios, as well as with also not diluted pure Espey Spring Water.

All adsorption studies were carried out at $25\ ^\circ\text{C}$ in a water bath with 170 rpm rapid shaking for 2 hours. Then, the samples were filtered through blue striped filter paper and the filtrate was analysed by titration method mentioned above. 100 mL Espey Spring Water was used. Adsorption studies were carried out using different amount of adsorbent, pH and particle size to determine the effects of these parameters on adsorption.

Adsorption studies with sawdust were carried out by diluting the spring water with pure water at 1/2, 1/5 and 1/10 ratios with also not diluted pure Espey Spring Water.

Figure 3 shows adsorption studies for 0.1 M HCl treated walnut shells in shaking water bath.



Figure 3. Adsorption studies with almond shells in a shaking water bath

2.3. Chemical Precipitation and Adsorption with Espey Spring Water(b)

In the third part of this study, since there was no leftover water before adsorption, Espey Spring Water was brought from Emet again and the amount of boron was measured. Firstly, the stream water was filtered and the boron amount was measured. The boron amount of the water was 3810 ppm.

2.3.1. Chemical Precipitation with Espey Spring Water(b)

That time, studies were conducted with 10 g of lime. 2 hours of working time and 2 hours of waiting time were selected in order to increase boron removal. Studies were carried out in a temperature-controlled magnetic stirrer heater at a temperature of 25 °C and a rotation speed of 760 rpm according to part 2.2.1 *Chemical precipitation with Espey Spring Water(a)*. Then, the samples were filtered through blue striped filter paper and the filtrate was analysed by titration method mentioned above to determine boron value left.

2.3.2. Adsorption with Espey Spring Water(b)

Afterwards, waste activated carbon, used in water purification devices, was tried for adsorption studies, which was a different environmentally friendly adsorbent. Waste activated carbon was ground and treated by immersing and maintained it in 0.1 M NaOH for 24 hours. Afterwards, it was filtered and washed with distilled water until drained water pH became 7 and then dried in the oven at 110 °C for 3.5 hours and made ready for adsorption. In addition, another environmentally friendly adsorbent pine cane was tried for boron removal. Pine cane was ground and treated with 0.1 M NaOH and kept for a 24 hours. Afterwards, it was filtered and washed with distilled water, dried in the oven at 90 °C, since it being more heat sensitive, for 3.5 hours and made ready for adsorption. All adsorption studies were carried out at 25 °C in a water bath with 170 rpm rapid shaking for 2 hours. Then, the samples were filtered through blue striped filter paper and the filtrate was

analysed by titration method mentioned above. 100 mL Espey Spring Water was used. In that time in order to increase effectiveness, 10 g of adsorbent was used for both adsorbents and adsorption was carried out in the same time for both adsorbents to acquire better results.

Adsorption studies were continued with almond shells used in part 2.2.2 *Adsorption with Espey Spring Water (a)* in order to be compatible with the studies of this part. For almond shells; the effect of the amount of adsorbent as 2, 4, 8, 10, 20 g; the effect of pH as 4, 7, 10, 11, 12; the effect of mixing speed as 120, 150, 170, 180, 200 rpm, the effect of temperature as 15 °C, 25 °C, 40 °C, 55 °C, 70 °C and effect of time as 1, 2, 3, 4, 5 hours were studied to see the effect of these parameters on adsorption efficiency. All adsorption studies were carried out at in the shaking water bath. 100 mL Espey Spring Water was used.

3. RESULTS AND DISCUSSION

3.1. Adsorption and Chemical Precipitation with Emet Boric Acid Factory Waste Dam

3.1.1. Adsorption Studies with Emet Boric Acid Factory Waste Dam

The results of adsorption studies which were carried out with waste clays (obtained from Emet, Bandırma, Kırka, Bigadiç operations directorates and indexed as 1,2,3,4,5 respectively) are shown in Table 1.

Table 1. Adsorption results using clay samples

Clay Samples	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
1	19.307	19,307	5.994	5994
2	35.736	35,736	11.098	11,098
3	56.802	56,802	17.642	17,642
4	40.976	40,976	12.725	12,725
5	278.682	278,682	86.547	86,547

The results in Table 1 indicates that, the desired boron removal efficiency could not be achieved in studies conducted with clays. Comparison of initial (5883 ppm) and final boron contents of wastewater indicates an increase in the boron content of the wastewater. This phenomenon can be explained as follows.

According to these results, the desired boron removal efficiency could not be achieved in studies conducted with clays considering the initial boron value of wastewater was 5883 ppm. There was an increase in the amount of boron in the water compared to the initial value.

Baran (2014) have mentioned the following studies as follows: “Keren et al. (1981) and Karahan et al. (2006), montmorillonite, kaolinite, illite, sepiolite, bentonite and their modified forms with nonylammonium chloride were used as adsorbents for boron removal. In these studies, experiments were carried out using initial boron concentrations of 1-50 mg/L. As a result of batch system studies, it was observed that the boron retention capacity of modified clay surfaces was higher than that of unmodified clay mineral”.

As reported by Seyhan et al. (2007), iron-rich Çamlıca bentonites CB1, CB2 were used for boron removal from water. Boron was determined by the fluorimetric-azomethine-H method, which is a new method. Using the factorial design method, the optimum conditions were found to be 0.250 g clay and 20 mL sample at pH 10, 45 °C. Under these conditions, boron removal is 80% for CB1 and 30% for CB 2. This high removal value can be due to the lower boron values of boron water and clay used was not saturated with boron.

As mentioned above, in these previous studies boron values in water are in low amounts. However, in this study boron in water was very high (5883 ppm). Also the increase in boron concentration after adsorption according to the Table 1, can be explained by the release of boron during adsorption from clay minerals that extracted from the same mine and therefore saturated with boron. The results of adsorption studies which were carried out by using natural zeolite as adsorbent are shown in Table 2-3-4.

Table 2. Adsorption results using natural zeolite for 1 hour

Zeolite Samples	B₂O₃ [g/L]	B₂O₃ [ppm]	B [g/L]	B [ppm]
1	18.677	18,677	5.800	5800
2	18.588	18,588	5.773	5773
3	18.588	18,588	5.773	5773

Table 3. Adsorption results using natural zeolite for 1 hour and 45 minutes

Zeolite Samples	B₂O₃ [g/L]	B₂O₃ [ppm]	B [g/L]	B [ppm]
1	18.659	18,659	5.795	5795
2	18.552	18,552	5.761	5761
3	18.570	18,570	5.767	5767

Table 4. Adsorption results using natural with zeolite for 2 hours

Zeolite Samples	B₂O₃ [g/L]	B₂O₃ [ppm]	B [g/L]	B [ppm]
1	18.659	18,659	5.795	5795
2	18.552	18,552	5.761	5761
3	18.570	18,570	5.767	5767

In studies conducted with 3 different fractions of zeolite over different time periods, it was seen that time and particle size changes did not have a significant effect on the amount of adsorption and most importantly, the amount of boron in the wastewater after adsorption was still at very high levels. As a result of the boron-rich zeolite found in the same area with boron in the boron local soils, it was thought that it did not have sufficient capacity in terms of boron adsorption, and to use a new adsorbent, walnut shell was decided.

Also, Yüksel and Yürüm (2009) also stated that for boron removal, while natural zeolite (clinoptilolite) and demineralized lignite were not successful, 90% boron was removed by using fly ash.

Adsorption results which were carried out with Walnut Shell Particles Treated with HCl and H₂SO₄, (+500 µ) are shown in Table 5.

According to these results, the amount of boron remaining in the water showed that the effect of both acid-treated walnut shells on the adsorption efficiency was almost the same. Considering that the initial boron concentration in water was 5883 ppm, as can be seen from Table 5, an efficient adsorption result was not achieved unfortunately.

Adsorption Results which were carried out with HCl activated walnut shell Particles at different pH's, (+500 µ) are shown in Table 6.

Table 5. Adsorption results using walnut shell particles treated with HCl and H₂SO₄, (+500 µ)

pH	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
2	18.340	18,34	5.696	5696
7	17.774	17,774	5.520	5520
9	16.731	16,731	5.196	5196
11	15.406	15,406	4.784	4784

Table 6. Adsorption results using HCl activated walnut shell particles at different pH's, (+500 µ)

	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
HCl treated walnut shells	18.818	18,818	5.844	5844
H ₂ SO ₄ treated walnut shells	18.800	18,8	5.839	5839

According to these results, as the pH value increased, the boron removal efficiency from boron-containing wastewater had increased, albeit slightly. However, it should be taken into consideration that as a result of pH adjustment, there was turbidity and precipitation after adsorption in the samples at pH 9 and 11, and that the samples were diluted as the amount added as a result of pH 9 and 11 adjustment was slightly more than the other samples and the total volume exceeded 100 mL. Because of the fact that, the unit of measurement was mg/L, ppm.

Adsorption Results with Different Amounts of Walnut Shell Particles (Activated with HCl) at pH 9, (+500 µ) are shown in Table 7.

Table 7. Adsorption results using different amounts of walnut shell particles (activated with hcl) at ph 9, (+500 μ)

Amount of Adsorbent	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
2	16.225	16,225	5.039	5039
4	16.102	16,102	5.000	5000

According to these results, considering that the initial boron concentration was 5883 ppm, a small amount of removal was achieved. And as the amount of adsorbent increased at pH 9, the amount of adsorption increased, as seen in Table.7. This was because as adsorbent amount increased, the number of adsorption sites on the adsorbent surface increased.

Due to the inadequacy of walnut shells for boron removal, waste corn cobs and waste sunflower stems were used.

Adsorption Results with (acid/base) treated and untreated waste sunflower stems and waste corn cobs are shown in Table 8.

It can be seen from these results that, output ppm values obtained as a result of the studies were still around 5800. This showed that ground waste corn cobs and ground waste sunflower stems, as well as their acid and base modified forms, were unfortunately insufficient to remove boron from wastewater by adsorption, considering the initial value of the water was 5883 ppm which was a very high boron value.

For this reason, it was thought that it would be more appropriate to remove some of the boron in the wastewater by chemical precipitation with lime as a preliminary process, then to conduct adsorption studies with environmentally friendly natural adsorbents.

Table 8. Adsorption results using (acid/base) treated and untreated waste sunflower stems and waste corn cobs

Adsorbent Type	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
Waste Sunflower Stems	18.739	18,739	5.819	5819
Waste Corn Cobs	18.896	18,896	5.868	5868
0.1 M HCl treated Waste Sunflower Stems	18.931	18,931	5.879	5879
0.1 M HCl treated Waste Corn Cobs	18.896	18,896	5.868	5868
0.05 M NaOH treated Waste Sunflower Stems	18.896	18,896	5.868	5868
0.05 M NaOH treated Waste Corn Cobs	18.913	18,913	5.864	5864

3.1.2. Chemical Precipitation with Emet Boric Acid Factory Waste Dam

Results of chemical precipitation with different amounts of lime at 55 °C temperature are shown in Table 9.

It can be seen from these results that the boron removal increased as the amount of lime increased. The reason for this is as explained below.

With the addition of lime to boron wastewater, the pH of the water increased to value of around 11-12. And in this pH range, as explained in the *I. Introduction* part, at high pH's, boron is present in the form of $B(OH)_4^-$ in water. Therefore, the reaction between lime and boron is as follows:

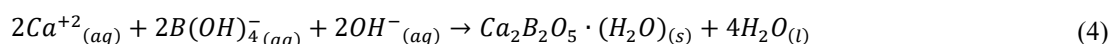


Table 9. Results of Chemical Precipitation Study Performed using Different Amounts of Lime at 55 °C Temperature

Lime Amount [g]	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
1.5	18.861	18,861	5.858	5858
3	16.346	16,346	5.076	5076
6	10.199	10,199	3.167	3167
15	2.480	2480	0.771	771
30	2.235	2235	0.694	694

According to the above reaction, as the amount of lime (Ca^{+2}) increases, the amount of precipitate ($Ca_2B_2O_5 \cdot (H_2O)_s$) and the lime surface area that can interact with boron will increase. As a result, the boron removal efficiency from water will increase with the increasing amount of lime.

However, as the amount of lime increases, solution mixing becomes difficult and as the amount of filtrate decreases, sludge formation increases. Therefore, studies were limited to 30 g of lime. According to these results, while boron removal was not sufficient for 1.5 and 3 g lime (0.4% and 13% respectively), it was moderate for 6 g lime (46%); and for 15 and 30 g lime, it is promising and productive with 771 (87%) and 694 (88%) ppm, respectively. Filtration after chemical precipitation with different amounts of lime ($Ca(OH)_2$) at 55 °C can be seen from Figure 4.

Studies were continued at 25 °C to see the effect of temperature on precipitation with lime. Since the study at 1.5 and 3 g of lime by mass was insufficient at 55 °C, studies at 25 °C were conducted by using higher amount of lime.

The results of chemical precipitation with different amounts of lime at 25 °C temperature are shown in Table 10.

It can be seen from Table 10 that, as the amount of lime increased, the boron removal increased (about 49 ppm, although there is a slight difference between 6 and 15 g). While the boron removal in the work done with 6 g to 15 g of lime was very low (almost no removal), the work done with 30 g lime could be considered as moderately efficient with 55% removal.



Figure 4. Filtration After Chemical Precipitation with Different Amounts of Lime at 55 °C

Table 10. Results of chemical precipitation study performed using different amounts of lime at 25 °C temperature

Lime Amount [g]	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
15	0.593	593	0.184	184

Studies were continued at 88 °C to see the effect of temperature on precipitation with lime. The amount of lime to be worked at 88 °C was based on the amount of lime worked at 55 °C, as the results of the study at 55 °C were more efficient than 25 °C. And since 1.5 and 3 g were insufficient in the studies carried out at 55 °C, and 30 g could cause more sludge formation, the amount of lime to be worked at 88 °C was chosen as 15 g. The results obtained are presented in Table 11.

Table 11. Results of chemical precipitation study performed using lime at 88 °C temperature

Lime Amount [g]	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
6	18.966	18,966	5.890	5890
15	18.809	18,809	5.841	5841
30	8.453	8453	2.625	2625

According to this result; 184 ppm was a very efficient result with 97% removal.

In order to better understand the effect of reaction temperature on boron removal, it is first necessary to determine whether the reaction between lime and boron (Equation 4), given above, is exothermic or endothermic. The ΔH calculation for this reaction at 298 K is made below according to the values shown in Table 12 using Equation 5 and Equation 6 (Bilen et al., 2018).

Table 12. Enthalpies of reactants and products at 298 k (bilen et al., 2018)

Ion	ΔH_{298} (kJ/mole)
Ca^{+2}	-542.83
$\text{B(OH)}_{4\text{(aq)}}$	-1344.03
$\text{OH}^{-}\text{(aq)}$	-229.99
$\text{Ca}_2\text{B}_2\text{O}_5 \cdot (\text{H}_2\text{O})_s$	-3041.80
$\text{H}_2\text{O}_{(l)}$	-242.00

$$(\Delta H)_{298} = [1 \times \Delta H_{\text{Ca}_2\text{B}_2\text{O}_5 \cdot (\text{H}_2\text{O})_{(s)}} + 4 \times \Delta H_{\text{H}_2\text{O}_{(l)}}] - [2 \times \Delta H_{\text{Ca}^{+2}\text{(aq)}} + 2 \times \Delta H_{\text{B(OH)}_4^{-}\text{(aq)}} + 2 \times \Delta H_{\text{OH}^{-}\text{(aq)}}] \quad (5)$$

$$(\Delta H)_{298} = [1 \times (-3041.80) + 4 \times (-242.00)] - [2 \times (-542.83) + 2 \times (-1344.03) + 2 \times (-229.99)] = 223.90 \text{ kJ/mole} \quad (6)$$

The (ΔH) value of 223.90 kJ/mole shows that the reaction between Ca(OH)_2 and boron is endothermic and the reaction will shift towards the products as the temperature increases. According to this, in this study, it is an expected result that the amount of precipitated matter increases with the increase in temperature.

Considering the literature, Selimoglu and Boncukoğlu (2019) studied boron removal from synthetically prepared boron solutions by chemical precipitation with lime. It was concluded that by increasing the temperature and the amount of Ca(OH)_2 , boron removal increased. And the removal achieved was 98%.

In summary, regarding the studies carried out, in order to remove boron from the wastewater of Emet boron operations directorate waste dam with high boron content, since the adsorption process carried out with environmentally friendly adsorbents were unsuccessful, chemical precipitation with lime (Ca(OH)_2) was used as a pre-treatment, and efficient results were obtained with chemical precipitation process as 97% of removal.

3.2. Chemical Precipitation and Adsorption with Espey Spring Water(a)

3.2.1. Chemical Precipitation with Espey Spring Water(a)

Since efficient results were obtained with the chemical precipitation method using lime (Ca(OH)_2) for the removal of boron from the wastewater of the Emet boron operation directorate waste dam, the studies for the purification of Emet Espey Spring Water continued with chemical precipitation as a pre-treatment in this study period as well as purification of Emet boron operation directorate waste dam wastewater. It was aimed to continue the work with the treatment of spring water by adsorption, which would be obtained with optimum values to be reached.

The initial boron value of Emet Espey Spring Water, which feeds Gelenbek stream, was 4590 ppm and the pH was 6.86.

To remind that, efficient results were obtained with chemical precipitation for the removal of boron from the wastewater of the Emet boron operation directorate waste dam at 88 °C and 55 °C. Chemical treatment was continued at 25 °C because of the fact that practically working at high temperatures creates energy costs in applications temperatures and will indirectly harm the environment (fossil fuel use, etc.).

Results of chemical precipitation study of Espey Spring Water, conducted with different amounts of lime at 25 °C temperature are shown in Table 13.

Table 13. Results of chemical precipitation study of Espey Spring Water, conducted with different amounts of lime at 25 °C

Lime Amount [g]	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
5	1.366	1366	0.424	424
10	1.297	1297	0.403	403
20	1.210	1210	0.376	376

It can be seen from Table 13 that, changing the amount of lime did not have much effect on the removal, and 5 g would be sufficient for treatment. Because as the amount of lime increased, the amount of sludge, which was the waste produced as a result of chemical precipitation, would also increase. This was a negative situation for the environment in water treatment. The decrease from 4590 ppm to 424 ppm with 91% percentage removal could be considered as a very good result in terms of chemical treatment with lime. The boron value of this output water was 424 ppm and its pH was 11.67.

3.2.2. Adsorption with Espey Spring Water(a)

Afterwards adsorption studies were carried out. The adsorption results of Espey Spring Water with different amounts of walnut shell particles (activated with HCl) for +500 µ particle size at 25 °C are shown in Table 14.

Table 14. Adsorption results of Espey Spring Water using different amounts of walnut shell particles (activated with hcl), (+500 µ) at 25 °C

Amount of Adsorbent [g]	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
2	1.320	1320	0.410	410
4	1.293	1293	0.401	401

Considering that the initial boron concentration was 424 ppm, a slight removal was achieved with output value of 410 ppm for 2 g of adsorbent (%3.3 removal efficiency) and output value of 401 ppm for 4 g of adsorbent (%5.4 removal efficiency). And as the amount of adsorbent increased, the amount of adsorption increased, albeit slightly. This was because as adsorbent amount increased, the number of adsorption sites on the

adsorbent surface increased. However, with these results, it could be considered that boron removal was not sufficient. Because of this, it was decided to continue to the studies by adjusting pH.

According to the former studies carried out with walnut shells in the scope of this study, pH was adjusted to 9. The adsorption results of Espey Spring Water with different amounts of walnut shell particles (activated with HCl) for +500 μ particle size at 25 °C, at pH 9 are shown in Table 15.

Table 15. Adsorption results of Espey Spring Water using different amounts of walnut shell particles (activated with hcl), (+500 μ) at 25 °c, at pH 9

Amount of Adsorbent [g]	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
2	1.246	1246	0.387	387
4	1.219	1219	0.379	379

It can be seen from these results that, considering that the initial boron concentration was 424 ppm, a slight removal was achieved with output value of 387 ppm (9% removal efficiency) for 2 g of adsorbent and an output value of 379 ppm (11% removal efficiency) for 4 g of adsorbent. And as the amount of adsorbent increased, the amount of adsorption increased, albeit slightly. Compared to the previous study without pH adjustment (pH: 11.67), as seen in Table 14, the amount of boron removal from wastewater increased by adjusting the pH to 9.

In summary, for the adsorption studies carried out with HCl activated walnut shells of +500 μ particle size, as the amount of adsorbent increased, a very small decrease was observed. Again, in the study conducted by adjusting the pH of Espey Spring Water to 9, more efficient results were obtained compared to the study with pH 11.67, original pH. This was because, as explained in the above sections, pH is an important parameter effecting the behaviour of water-soluble boron. Boron is present in the form of non-ionized boric acid, H₃BO₃, in water when the pH is less than 7, and behaves as (-) charged borate, B(OH)₄⁻, at pH values greater than 10.5. At higher pH values, hydroxyl ions compete with (-) ions on the adsorbent surface and (+) charged surfaces turn into (-) charges. This can be considered as the reason for the lower boron removal efficiency in the previous study (pH: 11.67) without pH adjustment.

For +500 μ particle size walnut shells, the highest yield was obtained with 4 g adsorbent at pH 9 with 379 ppm (11% removal efficiency).

In order to understand the effect of particle size on adsorption, studies continued with the walnut shells in the -63 μ fraction. Also different pH values were tried. The adsorption results of Espey Spring Water with 2 g of walnut shell particles (activated with HCl,) for -63 μ particle size at 25 °C, at different pHs are shown in Table 16.

According to these results, with an output value of 346 ppm (18% removal efficiency), the optimum removal pH could be considered as 9 for -63 μ fraction. It could be thought that as the degree of acidity or alkalinity increased, adsorption was negatively affected. Because of that walnut shell particles were modified with HCl, they had a surface with (+) charged. So in this situation, for effective adsorption, the molecule to be adsorbed should have (-) charge. And pH is an important parameter affecting the behaviour of water-soluble boron. When the pH is less than 7, boron is present in water in the form of non-ionized boric acid, H_3BO_3 and at pH values greater than 10.5 behave as, (-) charged borate, $B(OH)_4^-$ as seen in Figure 1. So, in this conditions borate was more favourable for (+) charged surfaces. At higher pH's as 11.30 and 12 as seen in Table 16, the reason for the decrease in boron adsorption was that the hydroxyl ion competed with the (-) ions on the surface of adsorbent. Because at higher pH's, (+) charged surfaces turned to (-) charge.

Table 16. Adsorption results of Espey Spring Water using 2 g of walnut shell particles (activated with hcl), (-63 μ) at 25 °C, at different pHs

pH	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
12	1.307	1307	0.406	406
11.30	1.289	1289	0.400	400
9	1.115	1115	0.346	346
7	1.150	1150	0.358	358
4	1.185	1185	0.368	368

In summary, the results of the two studies on HCL treated walnut shells at the +500 μ fraction (2 g of adsorbent amount) and HCL treated walnut shells at the -63 μ fraction (2 g adsorbent amount) at 25 °C were compared at close pH's as 11.67 and 11.30. It was observed that when the particle size decreased, the adsorption efficiency increased slightly, as seen from Table 14 and Table 16, from the 410 ppm (+500 μ fraction) (3.3% removal efficiency) and 400 ppm (-63 μ fraction) (5.6% removal efficiency) values. Because of as particle size decreased, the surface area of adsorbent increased. And because of the fact that resistance to reach pores would be shortening and easier pore diffusion would occur. Decrease in diffusional resistance became.

Then, since adsorption process carried out with walnut shells were unsuccessful, another environmentally friendly adsorbent, almond shells were tried for adsorption studies.

The adsorption results of Espey Spring Water with 2 g of almond shell particles (Activated with HCl), for +500 μ particle size at 25 °C, at different pHs are shown in Table 17.

The adsorption results of Espey Spring Water with 4 g of almond shell particles (Activated with HCl), for +500 μ particle size at 25 °C, at different pHs are shown in Table 18.

Table 17. Adsorption results of Espey Spring Water using 2 g of almond shell particles (activated with hcl), (+500 μ) at 25 °C, at different pHs

pH	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
12	1.117	1117	0.347	347
9	1.064	1064	0.330	330
7	1.134	1134	0.353	353
4	1.134	1134	0.353	353

Table 18. Adsorption results of Espey Spring Water using 4 g of almond shell particles (activated with hcl), (+500 μ), at 25 °C, at different pHs

pH	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
12	1.099	1099	0.341	341
9	1.046	1046	0.325	325
7	1.099	1099	0.341	341
4	1.117	1117	0.347	347

When the remaining boron values after adsorption in wastewater at different pHs were evaluated, considering that the initial boron concentration was 424 ppm, it was seen that a small amount of boron removal was achieved.

Comparison of results given in Table 17 and 18 showed that the boron removal increases with the increasing amount of adsorbent. This may be explained by the increasing surface area and consequent increase of active sites available for adsorption with the increasing amount of adsorbent.

The optimum removal pH was determined as 9 for almond shells for +500 μ fraction activated with 0.1 M HCl as output value 330 ppm (for 2 g adsorbent) (22 % removal efficiency) and as output value 325 ppm (for 4 g adsorbent) (23 % removal efficiency). That was because of the reason explained for results discussion of Table 16.

Then, the adsorption process was carried out with almond shells in the -63 μ fraction to see the effect of particle size on adsorption. Adsorption results of Espey Spring Water with 2 g of almond shell particles (activated with HCl), for -63 μ size at 25 °C, at different pHs are shown in Table 19.

It can be seen from these results that, when the remaining boron values after adsorption in Espey Spring Water at different pHs were evaluated, considering that the initial boron concentration was 424 ppm, a small amount of boron removal was achieved. The optimum removal pH was 9 with output value as 324 ppm (23.5% removal efficiency). That was because of the reason explained for the discussion of results given in Table 16.

Table 19. Adsorption results of Espey Spring Water using 2 g of almond shell particles (activated with hcl), (-63 μ) at 25 °C, at different pHs

pH	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
12	1.077	1077	0.334	334
9	1.042	1042	0.324	324
7	1.095	1095	0.340	340
4	1.112	1112	0.345	345

Comparison of the results given in Table 17 (+ 500 μ) and Table 19 (- 63 μ), showed that, when the particle size decreased, the adsorption efficiency increased. Because of as particle size decreased, the surface area of adsorbent increased. Also decrease in particle size causes decrease in diffusional resistance due to the shortened path to reach the active sites interior parts of the adsorbents through the pores. This implies that the adsorption of boron having a diffusion resistance controlling mechanism.

Then, another environmentally friendly adsorbent, saw dust was tried for adsorption studies. First of all, the concentration of boron in Espey Spring Water, which was previously subjected to chemical treatment and then filtered, was measured. The boron value of water was 404 ppm. The reason for the decrease from 424 ppm to 404 ppm in spring water was due to calcium borate precipitation formed at the bottom of Espey Spring Water over time. The adsorption results of Espey Spring Water by using 2 g of saw dust at 25 °C, at different dilution rates are shown in Table 20.

Table 20. Adsorption results of Espey Spring Water using 2 g of saw dust at 25 °C, at different dilution Rates

Dilution Rate	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
0	1.233	1233	0.383	383
1.Feb.	0.616	616	0.191	191
1.May.	0.205	205	0.064	64
1.Oct.	0.120	120	0.037	37

It can be seen from these results that, the amount of boron remaining in the wastewater after the adsorption experiment carried out under the same conditions but without dilution, was 383 ppm. The boron removal achieved was 21 ppm (5.2% removal efficiency), which is much higher than the discharge upper limit value specified in the wastewater regulation of Turkey. For this reason, it was decided to carry out adsorption experiments by diluting the Espey Spring Water.

After adsorption using 1/2 diluted Espey Spring Water, the boron concentration remaining in the wastewater was found to be 191 ppm. With dilution, the initial boron concentration in wastewater, which was 202 ppm before adsorption, decreased to 191 ppm. In other words, 11 ppm (5.4% removal efficiency) removal was

achieved. Since it was observed that boron removal increased with dilution, it was decided to conduct adsorption studies by increasing the dilution rates.

After adsorption using 1/5 diluted Espey Spring Water, the boron concentration remaining in the wastewater was found as 64 ppm. Considering that the initial boron concentration with dilution was 81 ppm, it could be seen that 17 ppm (26.6% removal efficiency) boron removal was achieved by adsorption.

After adsorption using 1/10 diluted Espey Spring Water, the remaining boron concentration in the wastewater was found as 37 ppm. Considering that the concentration of wastewater before adsorption as a result of dilution was 40 ppm, 3 ppm (7.5% removal efficiency) boron removal was achieved via adsorption.

When all results were evaluated for 2 g adsorbent, boron removal was achieved, albeit slightly, as a result of adsorption. The reason for this could be explained by the fact that sawdust is a material with a porous cellulosic structure and complexation with the hydroxylic, carboxylic and phenolic groups in its structure. In addition, it could be mentioned that treatment with NaOH had a positive contribution to adsorption in terms of removing impurities in the adsorbent and allowing easier access of boron to the pores.

As explained in the literature, concentration is an important parameter effecting the behaviour of boron in water. The concentration constant (K) of boric acid increases with increasing concentration. It is known that if the total boron concentration in aqueous solution is less than 0.025 M (0.1% B₂O₃ or 270 mg B/L), only H₃BO₃ and B(OH)₄⁻ ions are present in the environment. At higher concentrations (>0.025 M), polyborate ions such as B₃O₃(OH)₄⁻, B₄O₅(OH)₄⁻ and B₅O₆(OH)₄⁻ are formed.

The reason for the positive effect of dilution on adsorption is that, only H₃BO₃ and B(OH)₄⁻ ions present in water at low concentrations; whose adsorption is easier than large molecules such as polyborates,

These results showed that boron removal increases with dilution up to a certain dilution, but beyond this dilution ratio, boron removal started to decrease again. In short, it was observed that there was an optimum dilution ratio for boron removal. The reason for this is related to the theorem that as the boron concentration decreases, the driving force for the adsorption decreases.

Based on these results, it could be concluded that the active area of the adsorbent might be insufficient at high boron concentrations, in other words, since all active areas were occupied by boron, the adsorption phenomenon reached a dynamic equilibrium and the amount of adsorbent was increased and the adsorption studies were carried out with different dilution ratios. It was decided to repeat adsorption studies for different dilution ratios by increasing the amount of adsorbent in the same conditions.

The studies were repeated for 4 g of adsorbent and the results were given in Table 21.

Table 21. Adsorption results of Espey Spring Water using 4 g of saw dust at 25 °C, at different dilution rates

Dilution Rate	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
0	1.182	1182	0.367	367
1/2	0.599	599	0.186	186
1/5	0.171	171	0.053	53
1/10	0.103	103	0.032	32

As a result of the adsorption experiment performed without dilution, the amount of boron remaining in the wastewater is 367 ppm, i.e. 37 ppm (9.2%) boron removal was achieved. This concentration is much higher than the discharge upper limit value specified in the regulations. Therefore, it was decided to carry out the adsorption experiments by using diluted wastewater.

The results of adsorption experiments by using 1/2, 1/5 and 1/10 diluted wastewater showed that, the boron concentration remaining in wastewater are 186 ppm, 53 ppm and 32 ppm respectively. In another words, at these dilution ratios amount of boron removed 16 ppm, 81 ppm, 8 ppm and removal efficiencies are 7.9%, 34.6% and 20% respectively. These results very well confirmed with the adsorption experiments carried out by using 2 g of adsorbent by indicating there is an optimum dilution ratio for adsorption.

Also comparison of results given in Table 20 and Table 21 showed that, as the amount of adsorbent increased, the amount of adsorption increased due to the increase in surface area and hence increase of active sites on it.

When the results of the studies conducted, 0, 1/2, 1/5 and 1/10 dilution ratios by using 2 g and 4 g adsorbent, were compared, 4.0%, 2.5%, 8% and 12.5% excess removal were obtained for the case of 4g adsorbent than the 2 g adsorbent cases. This may be explained by the increase in area and consequently in active sites available for adsorption with the increasing amount of the adsorbent.

Since the positive effect of increasing the amount of adsorbent was seen, it was considered to continue the studies with the increase in the amount of adsorbent.

When the studies about saw dust in literature was considered; Jaouadi (2021) explains the study as follow, “Pinewood sawdust is a good precursor for activated carbon production. Activated carbon was prepared from sawdust by chemical activation using H₃PO₄ at 300°C. Sawdust and activated carbon were investigated as bio-adsorbents for boron removal from freshwater. Sawdust showed the highest boron adsorption uptake (1.58 mg g⁻¹) compared to activated carbon (0.97 mg g⁻¹). This higher capacity was caused by the presence of hydroxylic, carboxylic, and polyphenolic groups. FTIR analysis before and after adsorption of boron suggested that the mechanism involved in the boron removal might be complexation.” Sasaki et al. (2011) explains the study as follow: “Effective immobilization of boron in groundwater is a major challenge. Permeable reactive barrier (PRB) column tests for removal of borate have been investigated using MgO agglomerates as the primary

reactive material over 40 weeks. Additionally, saw dust was also blended with MgO agglomerates to facilitate for borate removal in this system. Boron accumulation was more than 1.6 times greater in the presence of saw dust, although MgO alone performed well. Increased boron accumulation in the presence of saw dust was primarily due to higher porosity of the PRB column, decreasing the impact of secondary $\text{Mg}(\text{OH})_2$ passivating layers and leaving more reactive sites on MgO agglomerates. In addition, Mg^{2+} ions released from MgO agglomerates are complexed with carboxylic acids leached from saw dusts. This sequestration prevents the formation of bulky $\text{Mg}(\text{OH})_2$ which is an ineffective sorbent for borate and covers the surfaces and passivating reactive sites on the MgO agglomerates. The morphologies of $\text{Mg}(\text{OH})_2$ precipitated in the PRB column were also significantly affected by the presence of saw dust, with crystallization of needle-like particles of $\text{Mg}(\text{OH})_2$ was prevented by Mg^{2+} ions–organic ligand complexation.”

As mentioned in literature studies explained above and according to the results in this study, saw dust can be good candidate for boron removal from diluted boron waters according to presence of hydroxylic, carboxylic, and polyphenolic groups.

3.3. Chemical Precipitation and Adsorption with Espey Spring Water(b)

3.3.1. Chemical Precipitation with Espey Spring Water(b)

The results of chemical precipitation with 10 g of lime at 25 °C temperature are shown in Table 22.

Table 22. Results of chemical precipitation study performed with 10 g of lime at 25 °C temperature

Lime Amount [g]	B_2O_3 [g/L]	B_2O_3 [ppm]	B [g/L]	B [ppm]
10	0.957	957	0.297	297

The decrease from 3810 ppm to 297 ppm (92% removal efficiency) can be considered as a very good result in terms of chemical purification.

3.3.2. Adsorption with Espey Spring Water(b)

Adsorption results of Espey Spring Water(b) with 10 g of waste activated carbon (activated with 0.1 M NaOH) and with 10 g of pine cane (activated with 0.1 M NaOH), at 25 °C are shown in Table 23.

Table 23. Adsorption results of Espey Spring Water(b) using 10 g of Waste Activated Carbon (activated with 0.1 M NaOH) and using 10 g of Pine Cane (activated with 0.1 M NaOH)

Adsorbent Type	B_2O_3 [g/L]	B_2O_3 [ppm]	B [g/L]	B [ppm]
Waste Activated Carbon	0.905	905	0.281	281
Pine Cane	0.853	853	0.265	265

It can be seen from these results that for waste activated carbon, the amount of boron remaining in the wastewater was 281 ppm. Considering that the initial boron concentration in water was 297 ppm, the boron

removal achieved was 16 ppm (5.4% removal efficiency) which was not sufficient. And also, for pane cane the amount of boron remaining in the wastewater was 265 ppm. Considering that the initial boron concentration in water was 297 ppm, the boron removal achieved was 32 ppm (10.8% removal efficiency) which was not sufficient removal result also.

However, when the percentage removal efficiencies were considered, it was seen that waste activated carbon and pine cane were not appropriate adsorbents for boron removal. This can be explained by the fact that the pore structure of the activated carbon was not suitable for boron adsorption or the lack of interaction between boron and the functional groups that might be present in the adsorbent.

Because of these inefficient results, it was decided to continue the study with the adsorbent almond shells since efficient and could be improved results were obtained in the second part of this study.

Adsorption Results of Espey Spring Water with different amounts of almond shell particles (Activated with HCl), at 25 °C are shown in Table 24.

The results of adsorption experiments carried out using different amounts of adsorbent with an initial boron value of 297 ppm were as follows: for 2 g of adsorbent as a result of the experiment, boron removal was 130 ppm (43.8% removal efficiency); for 4 g of adsorbent, boron removal was 157 ppm (52.9% removal efficiency); for 8 g of adsorbent, boron removal was 163 ppm (54.9% removal efficiency); for 10 g of adsorbent boron removal was 163 ppm (54.9% removal efficiency); for 20 g of adsorbent, boron removal was 163 ppm (54.9% removal efficiency).

Table 24. Adsorption results of Espey Spring Water using different amounts of almond shell particles (activated with hcl), at 25 °C

Amount of Adsorbent [g]	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
2	0.540	540	0.167	167
4	0.450	450	0.140	140
8	0.433	433	0.134	134
10	0.433	433	0.134	134
20	0.433	433	0.134	134

According to these results, as the amount of adsorbent increased, the removal efficiency increased up to a certain amount of adsorbent till 8 g of adsorbent. The reason for this could be explained by the increase in the available surface area and the increase in active sites in the increased surface area as the amount of adsorbent increased. At 8 g of adsorbent and later, the adsorption removal efficiency remained constant. The reason for this was thought to be that the adsorption event reached a dynamic equilibrium.

When literature data was examined about the effect of effect of adsorbent on adsorption, according to the data exists in the study of Giray (2010), Öztürk and Köse (2008), Çengeloğlu et al. (2007), Uslu (2010), Arıkan (2019), Öztürk and Kazak (2004) and Atalay Sönmez (2014), as the adsorbent amount increased, boron removal increased up to a level, then remained constant due to the dynamic equilibrium reached.

The studies were continued with 8 g, which was the amount of adsorbent that yielded the best result in the previous study.

The adsorption results of Espey Spring Water with 8 g of almond shell particles (Activated with HCl), at 25 °C, at different pHs are shown in Table 25.

Table 25. Adsorption results of Espey Spring Water using 8 g of almond shell particles (activated with HCl), at 25 °C, at different pHs

pH	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
4	0.554	554	0.170	170
7	0.450	450	0.140	140
10	0.346	346	0.107	107
11	0.398	398	0.124	124
12	0.467	467	0.145	145

The results of adsorption experiments carried out at different pHs with an initial boron value of 297 ppm were as follows: The boron removal at pH 4 was 127 ppm (42.8% removal efficiency), at pH 7 was 157 ppm (52.9% removal efficiency), at pH 10 was 190 ppm (64% removal efficiency), at pH 11 was 173 ppm (58.3% removal efficiency) and at pH 12 was 152 ppm (51.2% removal efficiency).

According these results, with an output value of 107 ppm (190 ppm removal, 64% boron removal efficiency), the optimum removal pH could be considered as pH 10. That was because of the reason explained in the discussion of the results given in Table 16.

When literature data was examined about the effect of pH on adsorption, optimum pH values for boron removal using different adsorbents are pH 7-9 with using modified perlite (Zincir, 2013); pH 7 with using cotton cellulose (Liu et al., 2007); pH 10 with using non-activated waste sepiolite (NAWS) and HCl-activated waste sepiolite (AWS) (Öztürk & Kavak, 2004); pH 7.5-8.5 with using vermiculite clay, pH 7.5-9 with using perlite (Baran, 2014); pH 10 with using granular iron hydroxide (GDH) (Atalay Sönmez, 2014); pH 8.2-8.5 with using aluminum based adsorbent (Irawan et al., 2011); pH 8-10 with using clay and modified clay (Karahana et al., 2006); pH 8-10 for all adsorbents as fly ash, zeolite, and demineralized lignite (Yüksel & Yürüm, 2009); pH 10 with using sepiolite (Öztürk & Kavak, 2004); pH 5.13 with using natural pumpkin seed shell (Çelebi, 2020); pH 9 with using mineral adsorbents while pH 7 with using walnut shells and rice residues and pH 8 with using wheat residues (Jalali et al., 2016).

The studies were continued with 8 g of adsorbent and pH 10 which were the optimum values obtained.

The adsorption results of Espey Spring Water using 8 g of almond shell particles (activated with HCl), at 25 °C, at pH 10, at different mixing speeds are shown in Table 26.

Table 26. Adsorption results of Espey Spring Water with 8 g of almond shell particles (Activated with HCl), at 25 °C, at pH 10, at different mixing speeds

Mixing Speed [rpm]	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
120	0.363	363	0.113	113
150	0.346	346	0.107	107
170	0.346	346	0.107	107
180	0.346	346	0.107	107
200	0.346	346	0.107	107

The results of adsorption experiments carried out at different mixing speeds with an initial boron value of 297 ppm were as follows: The boron removal at 120 rpm was 184 ppm (62% removal efficiency), at 150 rpm was 190 ppm (64% removal efficiency), at 170 rpm was 190 ppm (64% removal efficiency), at 180 rpm was 190 ppm (64% removal efficiency), and at 200 rpm was 190 ppm (64% removal efficiency).

According to these results, it was seen that the mixing speed had almost no significant effect on the adsorption efficiency. Only at 120 rpm, less efficiency was obtained compared to other and stable results. In summary, it could be said that the adsorption efficiency was independent of the mixing speed after 150 rpm.

These results very well confirm with the results reported in the literature for effect of mixing speed on adsorption for boron removal. For example, Korkmaz (2011) stated that using boron removal efficiency did not change with mixing speed in the study of boron removal using Purolite S 108 resin.

The studies were continued by using 4 g of adsorbent, at pH 10 and 150 rpm mixing speed at which the optimum values were obtained.

The adsorption results of Espey Spring Water with 8 g of almond shell particles (activated with HCl), at pH 10, at 150 rpm, at different temperatures are shown in Table 27.

The results of adsorption experiments carried out at different temperatures with an initial boron value of 297 ppm were as follows: The boron removal at 15 °C was 144 ppm (48.5%), at 25 °C was 190 ppm (64%), at 40 °C 122 ppm (41.1%), at 55 °C was 100 ppm (33.7%) and at 70 °C was 79 ppm (26.6%).

According to these results, the effect of temperature on adsorption efficiency was significant. As the temperature increased, the adsorption efficiency decreased. From these results, it was concluded that the

adsorption was exothermic which is a characteristics of physical adsorption. The temperature at which adsorption was most efficient was 25 °C.

Table 27. Adsorption Results of Espey Spring Water using 8 g of almond shell particles (activated with HCl), at pH 10, at 150 rpm, at different temperatures

Temperature [°C]	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
15	0.492	492	0.153	153
25	0.346	346	0.107	107
40	0.563	563	0.175	175
55	0.633	633	0.197	197
70	0.703	703	0.218	218

When literature data was examined about the effect of temperature on adsorption, Atalay Sönmez (2014) stated in her study as follows: “In the study conducted by Öztürk et al. (2005) on boron adsorption from water using batch and column experiments and fly ash, maximum boron removal occurred at temperature 25 °C. Öztürk et al. (2008) investigated the effect of cerium oxide as an adsorbent material on boron removal from water using factorial experimental design method. Maximum boron removal occurred at temperature of 40 °C. In the study conducted by Kıpçak et al. (2012) on boron adsorption from aqueous solutions with waste calcined magnesite, the optimum temperature was 45 °C. In the study conducted by Kavak (2009), boron adsorption on calcined alunite was studied using a 23 factorial design. Boron removal decreased as temperature increased. The optimum optimum temperature was found to be 25 °C. In the study conducted by Polowczyk et al. (2013) using fly ash, it was found that boron adsorption increased with increasing temperature.” Che Ku Mohd and İsmail (2020) stated that, maximum boron removal was obtained at 27°C with using the (MAZC) composite beads of mangrove bark, alginate and zeolite (MAZC) as an adsorbent. Thermodynamic studies have shown that this adsorption process was a spontaneous exothermic process.

Uslu (2010) stated as follows: “From the data obtained as a result of the experiments, it was observed that the adsorption behavior did not cause a significant change in the adsorption of boron by PVC-NMG in the temperature range examined (20-40 °C).” Kluczka et al. (2013) stated that using a new adsorbent natural clinoptilolite and amorphous zirconium dioxide (ZrO₂), the adsorption was exothermic. Bhagyaraj et al. (2021) stated as follows: “Haddabi et al. explored that for boron removal efficiency of date palm seed ash temperature on the adsorption process had a minor effect. Al-Ghouti et al. investigated the use of eggshell membrane (ESM) and modified eggshell membrane (MESM) as biosorbents for boron. The results indicated that the adsorption of boron preferentially occurred at 35 °C, which favor a spontaneous endothermic reaction.”

The studies were continued with 8 g, pH 10, 150 rpm shaking speed and 25 °C which were the optimum values obtained.

The adsorption results of Espey Spring Water with 8 g of almond shell particles (Activated with HCl), at pH 10, at 150 rpm, at 25 °C, at different contact times are shown in Table 28.

Table 28. Adsorption results of espey spring water using 8 g of almond shell particles (activated with HCl), at pH 10, at 25 °C, 150 rpm, at different contact times

Contact time [h]	B ₂ O ₃ [g/L]	B ₂ O ₃ [ppm]	B [g/L]	B [ppm]
1	0.370	370	0.115	115
2	0.346	346	0.107	107
3	0.352	352	0.109	109
4	0.317	317	0.099	99
5	0.317	317	0.099	99

The results of adsorption experiments carried out as a function of time with an initial boron value of 297 ppm were as follows: The boron removal at 1 hour was 182 ppm (61.3%); at 2 hours was 190 ppm (64%); at 3 hours was 188 ppm (63%); at 4 hours was 198 ppm (66.7%); and at 5 hours was 198 ppm (66.7%).

When the adsorption results were evaluated, as the contact time increased, the adsorption efficiency increased from 1 hour to 4 hours. At the end of the 4 hours contact time, the system reached dynamic equilibrium.

When literature data was examined about the effect of contact time on adsorption;

Jalali et al. (2016) stated as follows: “B removal from aqueous solutions was investigated using natural easily available materials including bentonite, kaolinite, zeolite, waste calcite, and residues of wheat, rice and green shell of walnut as native and chemically modified with FeCl₃. The optimum time obtained for mineral sorbents and organic sorbents was 24 and 48 h, respectively.” Baran (2014) stated that optimum time values were 15 hours using vermiculite clay and 4 hours using perlite. In other studies; optimum contact times were found as 30 minutes using new adsorbent natural clinoptilolite and amorphous zirconium dioxide (ZrO₂) (Kluczka et al. 2013) and 24 hours with using the composite (MAZC) consisting of mangrove bark, alginate and zeolite as an adsorbent (Che Ku Mohd & İsmail, 2020). Atalay Sönmez (2014) stated in her study as follows: “According to the study conducted by Ferreira et al. (2006); boron removal from water was studied using Mg-Al and Mg-Fe hydrotalcite as adsorbent material. In the study, the time to reach equilibrium was found to be 120 minutes. In the study conducted by Wei et al. (2011); the design and production of environmentally friendly adsorbent materials for boron removal were studied. In the study, N-methyl glucamine, a natural biopolymer, and chitosan were functionalized by atom transfer radical polymerization and used as environmentally friendly adsorbent materials. Boron adsorption was 90% at the end of 8 hours of contact time. According to the study conducted by Ferreira et al. (2006); boron removal from water was studied using Mg-Al and Mg-Fe hydrotalcite as adsorbent material. In the study, the time to reach equilibrium was found to be 120 minutes. In the study conducted by Kıpçak et al. (2012) on boron adsorption from aqueous solutions with waste calcined

magnesite, the optimum contact time was determined as 420 minutes.” Bhagyaraj et al. (2021) stated as follows: “Haddabi et al. explored that for boron removal efficiency of date palm seed ash contact time on the adsorption process had a minor effect.” Saleh Al-dhawi et al. (2023) stated that with using CRB05 as the adsorbent optimum contact time reached was 255 minutes. Çelebi (2020) stated that with using pumpkin seed shell as adsorbent, optimum contact time reached was 5 minutes.

4. CONCLUSION

In the first part of this study, boron removal from wastewater taken from the Emet boric acid factory waste dam of Eti Mine Works was studied. Different adsorption studies were carried out in shaking water bath. Natural clays also zeolite which exit in the sites of Eti Mine Works and natural adsorbents as modified walnut shells, modified waste sunflower stems and modified waste corn cobs were used as adsorbents. Unfortunately, these adsorbents were found to be insufficient for the removal of boron from wastewater with high boron content as 5883 ppm.

An additional method was needed before adsorption. Since, the adsorption process with environmentally friendly adsorbents were insufficient for removing boron from industrial wastewater. It was thought that it would be more suitable to remove some boron from wastewater by chemical precipitation with lime then to continue to removal by adsorption using environmentally friendly adsorbents.

In the second part of this study, it was decided to study removal of boron from Espey Spring Water in the scope of social responsibility. The initial boron value of this water was 4590 ppm. After chemical precipitation with lime, the boron decrease from 4590 ppm to 424 ppm (as output value), 91% of removal, could be considered as a very good result in terms of purification. Afterwards, the adsorption process was started. 0.1 M HCl treated walnut shells, 0.1 M HCl treated almond shells and also 0.1 M NaOH treated sawdust were used separately as adsorbents. For walnut shells and almond shells, the effect of adsorbent amount, pH and particle size on adsorption were tried in the studies. And for saw dust, different dilution ratios were tried. Among the natural waste adsorbents used for boron removal of Espey Spring Water, the best results were obtained with 4 g of -63 μ almond shells at pH 9 with pure Emet Spring water with 23.5% removal efficiency and with 4 g of saw dust using 1/5 ratio diluted Espey Spring Water with 34.6% removal efficiency.

In the third part of this study, because of the fact that there wasn't not enough Espey Spring Water for further studies, spring water was provided from Emet, this time boron removal from 3810 ppm to 297 ppm (92% removal) was reached with two-stage chemical precipitation process. So, that was a good result again.

For adsorption studies, firstly modified waste activated carbon and modified pine cane were used as adsorbents. For waste activated carbon, boron removal achieved was 5.4% removal efficiency and for pine cane boron removal achieved was 10.8% removal efficiency. These were unfortunately inefficient results.

Since, improvable results were obtained with almond shells using not diluted but pure spring water in the second part of this study, studies were continued with almond shells used in part 2.2.2 *Adsorption with Espey Spring Water(a)*. To see the effect of these parameters on adsorption efficiency; the effect of the amount of adsorbent as 2, 4, 8, 10, 20 g; the effect of pH as 4, 7, 10, 11, 12; the effect of mixing speed as 120, 150, 170, 180, 200 rpm, the effect of temperature as 15 °C, 25 °C, 40 °C, 55 °C, 70 °C and effect of time as 1, 2, 3, 4, 5 hours were studied.

The best results were reached with using 8 g of adsorbent, at pH 10, at 150 rpm mixing speed, at 25 °C and at 4 hours contact time. *For adsorbent amount effect*, best results were obtained with 8 g of adsorbent (54.9% removal efficiency); *for pH effect* (with 8 g of adsorbent, at 25 °C), best results were obtained with pH 10 (64% removal efficiency); *for the effect of mixing speed effect* (with 8 g of adsorbent, at 25 °C, at pH 10), best results were obtained with pH 150 rpm (64% removal efficiency); *for the effect of temperature* (with 8 g of adsorbent, at pH 10, 150 rpm), best results were obtained with 25 °C (64% removal efficiency); *for the effect of contact time* (with 8 g of adsorbent, at pH 10, at 150 rpm, at 25 °C), best results were obtained with 4 hours (66.7% removal efficiency). According to these results; as amount of adsorbent increased, boron removal increased. Since the surface area and active sites available for adsorption increased. Also, experimental results showed that the boron removal increased with the decreasing particle size. This implies that the adsorption may have the diffusion controlling mechanism. pH was found to be very important because of the fact that behaviour of boron in a solution highly depends on pH. Also, adsorption was found to be exothermic. And this is physical adsorption's characteristic property. As the shaking time increased, the interaction time of the adsorbent and adsorbed metals increased. So, the adsorption increased and reached equilibrium after a while.

With 66.7% removal efficiency, it was concluded that modified waste almond shell particles could be good candidate for boron removal from boron wasters having high boron content.

Compared to the results in the second part of this study (23.5% removal efficiency) more efficient results were obtained with 66,7% removal efficiency with almond shells for boron removal. The reason for this was in the second part of this study, initial boron after single chemical precipitation boron value of Espey Spring Water was 424 ppm. However, in the third part of this study, initial boron after two-stage chemical precipitation boron value of Espey Spring Water was 297 ppm. As explained in literature part, 1.2 Boron Chemistry, "if the total boron concentration in the aqueous solution is less than 270 mgB/L, it is known that only H_3BO_3 and $B(OH)_4^-$ ions are present in the environment. Conversion to polyborates occurs between pH 6-11 and at higher concentrations ($>0.025 M$).". So, in the second part of this study boron value of spring water was 424 ppm before adsorption and at this boron value, polyborates might occurred since boron concentration was higher than 270 ppm and initial pH was 11.67 (after chemical precipitation). However, in the third part of this study, boron concentration of spring water was 297 ppm and initial pH was 11.34 (after chemical precipitation). And at this 297 ppm, it was thought that mainly H_3BO_3 and $B(OH)_4^-$ ions existed in the water compared to polyborates which was thought to be as newly started to occur. Because of this reason and also with the effect

of the optimum parameters applied such as adsorbent amount, pH, mixing speed, temperature and contact time, more efficient results were obtained compared to the second part of this study consisting adsorption with almond shells.

In further studies, it is recommended to determine which adsorption model the results fit into and also to determine thermodynamic parameters and the kinetic model which the adsorption fits into.

In summary, in the studies in the literature, in general, synthetically prepared boron solutions were studied for boron removal. From environmental and economic perspective and in the scope of social responsibility; evaluating natural waste adsorbents, as almond shells for boron removal from spring waters which have high boron content in the adsorption process, makes this study specific.

For further studies, it is hoped that unlike the methods used in Eti Mine Works, adsorption which is a useful, easy and economic method, will also ensure that the boron concentration in boron waters will be lowered below the regulation limit values in Turkey.

AUTHOR CONTRIBUTIONS

Methodology, İ.Ç.G and İ.A.; laboratory work, İ.Ç.G.; manuscript-original draft, İ.Ç.G.; manuscript-review and editing, İ.Ç.G and İ.A. All authors have read and legally accepted the final version of the article published in the journal.

ACKNOWLEDGEMENT

This study is a part of PhD thesis of İ.Ç.G. under the supervision of İ.A and also some parts of this study were presented in International Conference on Advanced Materials Science & Engineering and High-Tech Devices Applications, ICMATSE. (Çelebi Gürsoy & Ar, 2024). All experiments were carried out in laboratory belongs to Technology Development Department of Eti Mine Works. The authors thank very much to Tacettin Akbunar for all endless support and data share. And also the authors thank for Eti Mine Works, administrators as General Manager Deputy Derya Maraşlıoğlu, Head of Technology Development Department Emin Türk and Manager Emine Özlem Yavuklu Cindaş, Manager Çiğdem Arat Yılmaz and Manager Harun Avbaz for laboratory study permission. Also authors thank for Ayşegül Al, Dürdane Tosun and Ayşe Bagana for support and also thank for all laboratory team.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

Anonim (2010). TC Çevre ve Şehircilik Bakanlığı, Su Kirliliği Kontrolü Yönetmeliği, Dec.31, Resmi Gazete, 25687, 2010. <https://webdosya.csb.gov.tr/db/cygm/icerikler/skky-2018-20180226120233.docx>

- Arıkan, T. (2019). Sulu Çözeltilerdeki Bor İyonlarının Sepiyolit İle Giderimi. *KSÜ Tarım ve Doğa Dergisi*, 22(2), 333-338. <https://doi.org/10.18016/ksutarimdoga.v22i42694.498284>
- Akkurt, F., Yörükoğlu, A., Aktosun, Z., & Yargıçoğlu, P. (2022). Reducing the Amount of Boron in the Wastewater of a Boron Processing Plant by Chemical Precipitation. *Iranian Journal of Chemistry and Chemical Engineering*, 41(10), 3400-3406. <https://doi.org/10.30492/ijcce.2022.529108.4698>
- Aslan, K. (2020) *Borik Asitin Derişik Sulu Çözeltilerinden Modifiye Aktif Karbon ve Aljinat Bazlı Adsorbanlar ile Giderilmesi*, MSc Thesis, Süleyman Demirel University.
- Atalay Sönmez, N. (2014) *Sulardan Adsorpsiyon Yöntemi İle Bor Giderimi*, MSc Thesis, Pamukkale University.
- Ateş, Ç. (2018) *Ardışık Kimyasal ve Elektrokimyasal Arıtma ile Bor İçerikli Atık Sudan Bor Giderimi*. MSc Thesis, Ankara University.
- Bara, M. A., & Abba, M. U. (2025). Boron removal from industrial wastewater using hybrid adsorption-membrane system. *International Journal of Agricultural and Veterinary Science*, 7(1). <https://doi.org/10.70382/mejav.v7i1.030>
- Baran, P. (2014) *Atık Sulardan Bor Gideriminde Modifiye Doğal Kil Minerallerinin Kullanılabilirliğinin Araştırılması*. PhD Thesis, İTÜ.
- Bhagyaraj, S., Al-Ghouti, M. A., Kasak, P., & Krupa, I. (2021). An updated review on boron removal from water through adsorption processes. *emergent mater.* (4), 1167–1186. <https://doi.org/10.1007/s42247-021-00197-3>
- Bilen, M., Ateş, Ç., & Bayraktar, B. (2018). Yanıt yüzey yöntemi ile bir bor fabrikası atık suyu kimyasal arıtma sürecinde optimum koşulların belirlenmesi. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 33(1). <https://doi.org/10.17341/gazimmfd.406798>
- Bilici Başkan, M., & Atalay, N. (2014). Boron contamination in drinking - irrigation water and boron removal methods. *Pamukkale Univ Muh Bilim Derg*, 20(3), 78-84. <https://doi.org/10.5505/pajes.2014.47955>
- Çelebi, H. (2020). Doğal Kabak Çekirdeği Kabuğunun Bor Adsorpsiyon Kapasitesi, *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 9(4), 1698–1710. <https://doi.org/10.17798/bitlisfen.667331>
- Çelebi Gürsoy, İ., & AR, İ. (2024). *Removal of Boron and Boron Compounds From Waste Waters*. In: International Conference on Advanced Materials Science & Engineering and High Tech Devices Applications Oct. 24-26 2024, Ankara, Türkiye.
- Che Ku Mohd, C. K. N. A., & Ismail, M. H. S. (2020). Optimization of boron removal from aqueous solution via adsorption using composite beads of mangrove bark, alginate and zeolite. *Journal of Chemical Engineering and Industrial Biotechnology*, 3(1), 1-12. <https://doi.org/10.15282/jceib.v3i1.3739>
- Çengelloglu, Y., Tor, A., Arslan, G., Ersoz, M., Gezgin, S. (2007). Removal of boron from aqueous solution by using neutralized red mud, *Journal of Hazardous Materials*, 142(1-2), 412–417. <https://doi.org/10.1016/j.jhazmat.2006.08.037>
- Doğan, G., Sabah, E., & Erkal, T. (2005, 9-12 June). *Borun Çevresel Etkileri Üzerine Türkiye'de Yapılan Bilimsel Araştırmalar*. In: Türkiye International Mining Congress, İzmir

Eti Mine Works Quality Management System. *B₂O₃ Analyse Book*.

Fadaei, Abdolmajid (2022). A Comparison of Techniques of Boron Removal from Water and Wastewater. *Journal of Environmental Science and Sustainable Development*, 5(2), 404-435. <https://doi.org/10.7454/jessd.v5i2.1151>

Giray, G. (2010). *Sülfonamid Esaslı N-Metil-D-Glukamin Fonksiyonlu Polimerik Sorbentle Çözeltilerden Bor Giderimi*, MSc Thesis, İTÜ, İstanbul.

Jaouadi, M. (2021). Characterization of activated carbon, wood sawdust and their application for boron adsorption from water. *International Wood Products Journal*, 12(1), 22–33. <https://doi.org/10.1080/20426445.2020.1785605>

Jalali, M., Rajabi, F.H., & Ranjbar, F. (2016). The removal of boron from aqueous solutions using natural and chemically modified sorbents. *Desalination and Water Treatment*, 57, 8278-8288. <https://doi.org/10.1080/19443994.2015.1020509>

Karahan, S., Yurdakoç, M., Seki, Y., & Yurdakoç, K. (2006). Removal of boron from aqueous solution by clays and modified clays. *Journal of Colloid and Interface Science*, 293(1), 36-42. <https://doi.org/10.1016/j.jcis.2005.06.048>

Keren, R., & Mezuman, U. (1981). Boron Adsorption by Clay Minerals Using a Phenomenological Equation. *Clays and Clay Minerals*, 29(3), 198-204. <https://doi.org/10.1346/ccmn.1981.0290305>

Kluczka, J., Korolewicz, T., Zołotajkin, M., Simka, W., & Raczek, M. (2013). A new adsorbent for boron removal from aqueous solutions. *Environmental technology*, 34(9-12), 1369–1376. <https://doi.org/10.1080/09593330.2012.750380>

Korkmaz, M. (2011) *Bor İçeren Sulardan Purolite S 108 Reçinesi Kullanarak Bor Giderimi*. MSc Thesis, Balıkesir Üniversitesi.

Kökkılıç, O. (2003) *Kırka Bor İşletmesi Atık Göletlerindeki Sulardan Borun Kazanılması*. MSc Thesis, İTÜ.

Liao, L., Chen, H., He, C., Dodbiba, G., & Fujita, T. (2024). Boron Removal in Aqueous Solutions Using Adsorption with Sugarcane Bagasse Biochar and Ammonia Nanobubbles. *Materials*, 17(19), 4895. <https://doi.org/10.3390/ma17194895>

Liu, R., Ma, W., Wang, L. (2007). Effect of pH on biosorption of boron onto cotton cellulose. *Desalination*, 207, 257–267. <https://doi.org/10.1016/j.desal.2006.07.012>

Liu, H., Qing, B., Ye, X., Li, Q., Lee, K., & Wu, Z. (2009). Boron adsorption by composite magnetic particles. *Chemical Engineering Journal*, 151(1-3), 235-240. <https://doi.org/10.1016/j.cej.2009.03.001>

Liu, X., Xu, C., Chen, P., Li, K., Zhou, Q., Ye, M., Zhang, L., & Lu, Y. (2022). Advances in Technologies for Boron Removal from Water: A Comprehensive Review. *International Journal of Environmental Research and Public Health*, 19(17), 10671. <https://doi.org/10.3390/ijerph191710671>

Öcal, Z. B., Öncel, M. S., Keskinler, B., Khataee, A., & Karagündüz, A. (2024). Sustainable treatment of boron industry wastewater with precipitation-adsorption hybrid process and recovery of boron species. *Process Safety and Environmental Protection*, 182, 719-726. <https://doi.org/10.1016/j.psep.2023.12.006>

- Öztürk, N., & Kavak, D. (2004). *Adsorption of Boron from Aqueous Solutions by Sepiolite using Full Factorial Design: I. Batch Studies*. In: 2nd International Boron Symposium, Eskişehir, Türkiye.
- Öztürk, N., Köse T. E. (2008). Boron removal from aqueous solutions by ion exchange resin: Batch studies, *Desalination*, 227(1-3), 233–240. <https://doi.org/10.1016/j.desal.2007.06.028>
- Rodarte, D., Smith, R.S. (2014). A Comparison of Methods for Boron Removal From Flowback and Produced Waters. <https://jpt.spe.org/comparison-methods-boron-removal-flowback-and-produced-waters>
- Saleh Al-dhawi, B. N., Mohammad Kutty, S. R., Baloo, L., Mohammed Yahya Almahbashi, N., Mohsen Alawag, A., Saeed Ghaleb, A.A., Hussaini Jagaba, A., Abdulrahman Al-Mekhlafi, A. (2023). Boron removal from produce water through adsorption. *BIO Web of Conferences*, 73, 02002. <https://doi.org/10.1051/bioconf/20237302002>
- Selimoğlu, V., & Boncukcuoğlu, R. (2019). Bor Çözeltilerinden Kimyasal Çöktürmeyle Bor Giderimi. *Journal of the Institute of Science and Technology*, 9(3), 1343-1350. <https://doi.org/10.21597/jist.515526>
- Sütçü, L. (2005) *Removal of Boron From Waters Using Fly Ash*. MSc Thesis, Graduate School of Engineering and Science of İzmir Institute of Technology.
- Seyhan, S., Seki, Y., Yurdakoç, M., & Merdivan, M. (2007). Application of iron-rich natural clays in Çamlıca, Turkey for boron sorption from water and its determination by fluorimetric-azomethine-H method. *Journal of Hazardous Materials*, 146(1-2), 180–185. <https://doi.org/10.1016/j.jhazmat.2006.12.008>
- Uslu, A. S. (2010) *Fonksiyonel Hale Getirilmiş Pvc Esaslı Adsorban İle Sulu Çözeltilerden Bor Giderimi*, MSc Thesis, İTÜ.
- Ünlü, M. İ., Bilen, M., & Gürü, M. (2013). Kütahya-Emet Bölgesi Yeraltı Sularında Bor Ve Arsenik Kirliliğinin Araştırılması. *Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 26(4).
- Yağmur, Y. (2012) *Birlikte Çöktürme-Mineralizasyon Yöntemi ile Sulu Çözeltilerden Bor Kazanımı*, MSc Thesis, İTÜ.
- Yiğitbaşıoğlu, H. (2004). Türkiye İçin Önemli Bir Maden: Bor. *Coğrafi Bilimler Dergisi*, 2(2), 13-25. https://doi.org/10.1501/Cogbil_00000000046
- Yüksel, S., & Yürüm, Y. (2009). Removal of Boron from Aqueous Solutions by Adsorption Using Fly Ash, Zeolite, and Demineralized Lignite. *Separation Science and Technology*, 45(1):105-115. <https://doi.org/10.1080/01496390903256042>
- Zeytuncu, B., Pasaoglu, M. E., Eryildiz, B., Kazak, A., Yuksekdağ, A., Korkut, S., Kaya, R., Turken, T., Ceylan, M., & Koyuncu, I. (2023). Application of different treatment systems for boron removal from industrial wastewater with extremely high boron content. *Journal of Water Process Engineering*, 55(104083). <https://doi.org/10.1016/j.jwpe.2023.104083>
- Zincir, E. (2013) *Modifiye Edilmiş Perlit ile Atık Sulardan Bor Giderimi*, MSc Thesis, İTÜ.



Optimizing Choke Size to Minimize Sand Production in Oil Wells: A Machine Learning Approach

Sunday AGBONS IGBINERE^{1*} Ikponmwosa OHENHEN¹ Edobor Frankie CHRISTOPHER¹

¹ University of Benin, Benin City, Edo State, Nigeria

Keywords	Abstract
Sand Production	This study investigates the relationship between operational parameters like choke size etc. and sand production in some oilfields, aiming to optimize efficiency while minimizing sand-related challenges. Data visualization revealed trends in sand cut behaviour under varying gross rate, net rate, and BS&W conditions. Three machine learning models artificial neural network (ANN), Random Forest (RF) and extreme gradient boosting (XGBoost) were developed, with XGBoost achieving the highest accuracy. Extreme gradient boosting (XGBoost) outperformed the others by achieving the highest R-squared value of 0.952 and the lowest mean absolute error (MAE) and mean squared error (MSE), demonstrating its superior accuracy in predicting sand cut values. Shapley additive exPlanations (SHAP) analysis highlighted key parameters like manifold pressure, gas rate, and BS&W in predicting sand cut. Optimization using the Mealpy Genetic Algorithm yielded an optimal configuration gross rate of 750blpd, sand cut of 0.25pptb, and BS&W of 5.5%. Sensitivity analysis emphasized monitoring separator pressure and gas rate. The findings demonstrate the potential of integrating machine learning and optimization to enhance decision-making, reduce risks, and improve production efficiency. Recommendations for implementation and future research are provided to ensure sustainable operations.
Machine Learning	
SHAP	
ANN	
XGBoost	
Random Forest	
Optimization	
Choke	
Genetic Algorithm	
Sand Cut	

Cite

Igbinere, S.A., Ohenhen, I., & Christopher, E.F. (2025). Optimizing Choke Size to Minimize Sand Production in Oil Wells: A Machine Learning Approach. *GU J Sci, Part A, 12(2)*, 541-561. doi:10.54287/guj.1669814

Author ID (ORCID Number)	Article Process
0000-0003-1908-389X	Sunday AGBONS IGBINERE
0000-0002-5938-8973	Ikponmwosa OHENHEN
0009-0000-5319-891X	Edobor Frankie CHRISTOPHER
	Submission Date 03.04.2025
	Revision Date 08.04.2025
	Accepted Date 02.05.2025
	Published Date 30.06.2025

1. INTRODUCTION

Sand production is a pervasive challenge in oil and gas wells, particularly in poorly consolidated sandstone formations. It occurs when the mechanical strength of the reservoir rock is insufficient to withstand the stresses induced by fluid flow during production. Sand production can lead to severe operational and economic consequences, including erosion of surface and subsurface equipment, formation subsidence, sand accumulation in surface facilities, and increased maintenance costs due to frequent well interventions (Dickson et al., 2003; Completion Tech., 1995). These issues are exacerbated in high-rate production wells, where the fluid flow velocity exceeds the critical threshold required to mobilize sand particles. The causes of sand production are multifaceted, influenced by factors such as the degree of rock consolidation, pore pressure reduction, production rate, reservoir fluid viscosity, and water cut. For instance, younger, poorly consolidated formations, such as those found in the Gulf of Mexico, Nigeria, and the North Sea, are

*Corresponding Author, e-mail: agbons.igbinere@uniben.edu

particularly prone to sanding due to their low compressive strength (Completion Tech., 1997). Additionally, increased water production can destabilize the sand arch around perforations, further exacerbating sand production (Carlson et al., 1992). During production, reduction in pore pressure due to fluid withdrawal leads to increase of grain to grain effective stress and is directly proportional to the decrease in pore pressure, shear failure of the rock occurs when effective stress reaches the threshold value (Tiab & Donaldson, 2004).

To mitigate these challenges, various sand control methods have been developed, ranging from mechanical exclusion techniques (e.g., screens and gravel packs) to chemical consolidation and operational strategies like rate control and selective perforation. Among these methods is rate control, and it involves regulating the flow rate through choke size variations in order to help minimize sand production and has gained attention due to its simplicity and cost-effectiveness. By optimizing the choke size, operators can control the pressure drawdown and fluid velocity, thereby reducing the likelihood of sand creation and mobilization. However, traditional rate control methods often rely on empirical guidelines and trial-and-error approaches, which may not account for the complex interplay of reservoir properties, fluid dynamics, and operational constraints. This limitation underscores the need for a more systematic and data-driven approach to optimize choke size and minimize sand production.

When a well is first put into production, there is typically an initial surge in sand production due to the mobilization of perforation debris alongside the reservoir fluid. After reaching certain flow rates, a steady state is achieved, leading to transient sand production (Shabdirova et al., 2024). Extensive research has been conducted to model the onset of sand production, resulting in various analytical, numerical, and empirical models that predict the critical flow conditions that trigger sand production to varying extents (Morita et al., 1989; Kessler et al., 1993; Weingarten & Perkins, 2007; Wang & Dusseault, 2010; Han et al., 2011; Al-Shaabi et al., 2013; Fuh & Morita, 2013; Wu et al., 2016; Papamichos & Furui, 2019). Laboratory experiments are instrumental in studying the mechanisms behind sand production, though their outcomes are highly influenced by boundary conditions, and may not always be directly applicable to field-scale scenarios (Skjaerstein et al., 1997; Nouri et al., 2006a; van den Hoek et al., 2007; Fattahpour et al., 2012; Wu et al., 2016; Kozhagulova et al., 2021; Shabdirova et al., 2022). These experimental results are primarily used to calibrate and validate analytical and numerical sand production models.

Despite significant progress in analytical, numerical, and experimental methods for predicting sand production, these approaches are frequently constrained by the need for simplifying assumptions, high computational demands, and difficulties in translating laboratory-scale results to real-world field conditions. In recent years, machine learning (ML) has emerged as a powerful tool for addressing complex problems in the oil and gas industry (Anirbib et al., 2021). ML algorithms, such as artificial neural networks (ANNs), support vector machines (SVMs) as well as fuzzy logic (FL) as classification and regression problem tools (Ani et al., 2016), and particle swarm optimization (PSO), have been successfully applied to predict sand production and optimize well performance. The prediction of sanding onset can be framed as a classification

problem, with two possible outcomes: whether sanding will occur or not under specific conditions. Multiple studies have effectively utilized machine learning algorithms to predict sanding onset and evaluate the importance of various input variables. For example, Kanj and Abousleiman, (1999) pioneered the use of ANNs to predict sand onset, while subsequent studies have demonstrated the effectiveness of ML techniques in estimating critical drawdown pressure (CDDP) and identifying sand-prone zones (Azad et al., 2011; Khamsehchi et al., 2014). Khamsehchi et al. (2014) analyzed sand production onset in 23 field datasets from the North Adriatic Sea, using simple regression and artificial neural network (ANN) algorithms to predict critical total drawdown (CTD). Gharagheizi et al. (2017) applied support vector machine (SVM) to predict sanding onset using data from 31 wells in the Northern Adriatic Basin, with the algorithm demonstrating high predictive performance as assessed by the receiver operating characteristics (ROC) curve. Ketmalee and Bandyopadhyay (2018) utilized ANN to generate synthetic logs for sand prediction tools in the Bongkot field, Thailand, successfully validating the model with real well data. Ngwashi et al. (2021) compared the performance of ANN with backpropagation and SVM in predicting sanding onset in the Niger Delta region, achieving accuracy rates of 80% and 100%, respectively. Abdelghany et al. (2022) employed a probabilistic neural network to predict sand production onset by inferring reservoir properties, achieving higher accuracy compared to conventional models. Song et al. (2022) applied four machine learning algorithms to predict sand production in gas-hydrate sediments, recommending XGBoost for early-stage predictions. These advancements highlight the potential of ML to enhance sand management strategies by providing accurate, real-time predictions and optimizing operational parameters. This study proposes a machine learning approach to optimize choke size for minimizing sand production in oil wells. By leveraging historical production data, reservoir properties, and operational parameters, we aim to develop a predictive model that can identify the optimal choke size for a given well configuration. The proposed approach not only addresses the limitations of traditional rate control methods but also provides a scalable solution for improving well productivity and reducing sand-related costs. The integration of ML into sand management represents a significant step forward in the quest for sustainable and efficient oil and gas production as well as assisting in solving other reservoir engineering issues (Anifowose et al., 2017a, b).

2. MATERIAL AND METHOD

2.1. Data collection

The dataset used in this study was obtained from a producing asset located in the Niger Delta region. It comprises production data from eight wells spanning six different oil fields, each with varying geological and operational characteristics. The data was sourced from field operational logs and production monitoring systems, capturing a wide range of field conditions that reflect both stable and transient production phases. Key parameters recorded include: Gross Rate in barrel of liquid per day (blpd), Net Rate in barrel of oil per day (bopd), Sand Cut in parts per thousand barrels (pptb), BS&W in percentage (%), Gas Rate in million standard cubic feet (mmscf), manifold flowing pressure (MFP) in pounds per square inch (psi). The summary of the utilised dataset is shown in Table 1.

Table 1. Summary Statistic of Data

	count	mean	std	min	25%	50%	75%	max
CHOKE SIZE(/64")	330.0	20.436364	2.344990	16.0000	18.000000	22.000000	22.000000	26.00000
MFP (PSI)	330.0	253.589744	111.937716	60.0000	182.500000	253.589744	380.000000	440.00000
SEP P(PSI)	330.0	188.973915	86.819321	60.0000	150.000000	188.973915	190.000000	372.51000
GROSS RATE(blpd)	330.0	213.116532	229.767879	0.0000	0.000000	213.116532	289.200000	1026.00000
NET RATE (bopd)	330.0	131.272319	168.329870	0.0000	0.000000	131.272319	154.364940	918.18888
GAS RATE Inst (mmscf)	330.0	0.700557	0.905126	0.0000	0.000000	0.502500	0.700557	4.14590
WATER RATE (bwpd)	330.0	86.648385	150.197165	0.0000	0.000000	18.457600	86.648385	933.66000
BS&W(%) (FIELD LAB)	330.0	22.467722	25.667581	0.0500	1.350833	22.467722	22.467722	99.40000
DP Across Orifice(in.H2O)	330.0	284.750025	338.998713	14.6700	55.591667	284.750025	284.750025	1669.71000
SAND CUT (pptb)	330.0	0.302100	0.110532	0.1077	0.302100	0.302100	0.302100	0.87500
TOTAL_PRODUCTION	330.0	431.037236	457.616062	0.0000	0.000000	431.037236	578.400000	2052.00000
GROSS RATE(blpd)_PERCENTAGE	216.0	47.272800	11.011984	0.0000	49.442720	50.000000	50.000000	61.88253
NET RATE (bopd)_PERCENTAGE	216.0	31.514746	16.613604	0.0000	30.454983	30.454983	44.417500	100.00000
WATER RATE (bwpd)_PERCENTAGE	216.0	21.212454	22.027163	0.0000	5.582500	20.102297	20.102297	100.00000

2.1.1. Data quality and pre-processing

Before analysis, the dataset underwent rigorous quality checks to address missing values, outliers, and inconsistencies in measurement units. These pre-processing steps were essential to ensure data reliability for machine learning model training and optimization. To further understand the underlying data behaviour, the distribution of key variables—Sand Cut, Gross Liquid Rate, Water Rate, and BS&W—were visualized using histograms and kernel density estimation (KDE) plots (Figure 1-4). These visualizations revealed key distribution patterns and potential modelling implications. For instance, Gross Rate exhibited a right-skewed (log-normal) distribution, justifying the use of tree-based models like Random Forest or XGBoost, which are robust to non-Gaussian input features.

In particular, the distribution of Sand Cut (Figure 2) shows a sharp peak around 0.3 pptb, with a long tail extending toward higher values. This indicates that most wells operate within a narrow, low-sand range, while a few instances of higher sand production may reflect critical operational anomalies or edge cases. Marking a critical threshold of 0.5 pptb helps identify data points that exceed acceptable sand production limits and might require intervention or optimization of flow parameters. Upon data pre-processing, a Pearson correlation matrix was generated to examine the linear relationships among all numerical features in the dataset, as shown in Figure 5. This analysis aids in identifying potential multicollinearity and understanding how input variables interact. Notably, choke size (in 64ths of an inch) shows a negative correlation with sand cut (−0.23), suggesting that increasing the choke size may reduce sand production. This is consistent with fluid dynamics principles—larger choke openings reduce flow velocity, which can lower the mobilization of sand particles, aligning with Darcy’s law. Additionally, net oil rate exhibits strong positive correlations with gross liquid rate (0.75) and gas rate (0.67), which reinforces its dependency on both fluid phases. Conversely, water rate and BS&W (Basic Sediment and Water) show a significant inverse relationship with oil production, which is critical for sand-prone reservoirs where higher water production can increase sand-related issues. Figure 2 showed a right-skewed pattern with most values clustered around 0.3 pptb. A critical threshold of 0.5 pptb is highlighted, beyond which sand production may pose operational risks.

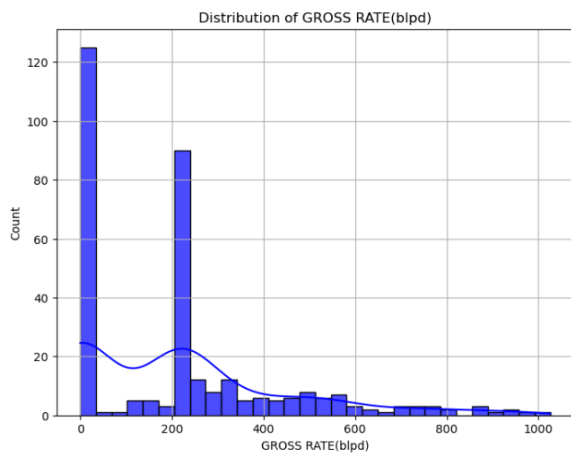


Figure 1. Distribution of Gross Rate

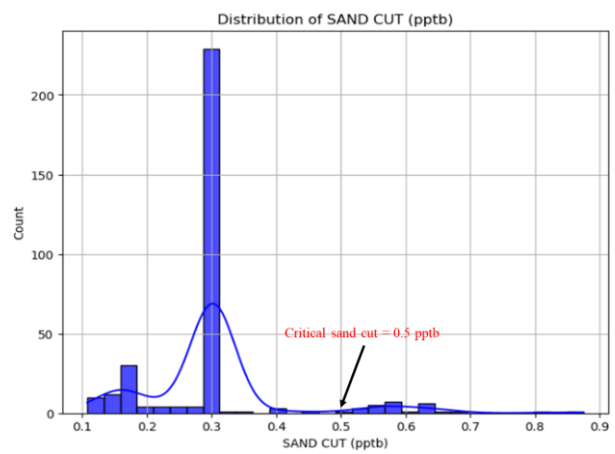


Figure 2. Distribution of Sand Cut

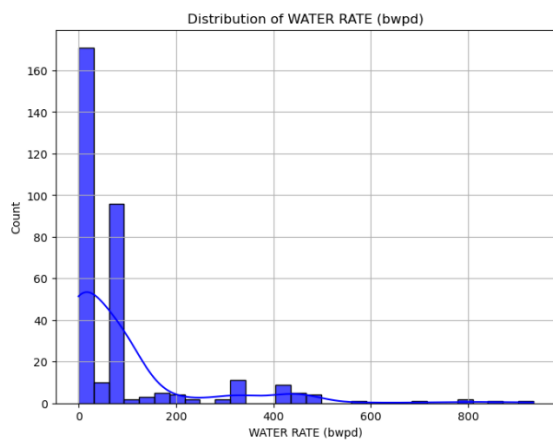


Figure 3. Distribution of Water Rate

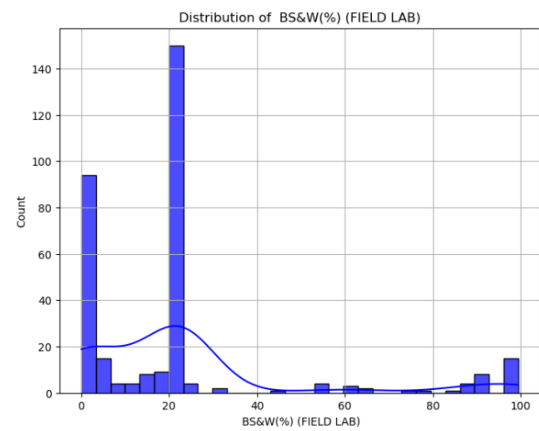


Figure 4. Distribution of BS&W

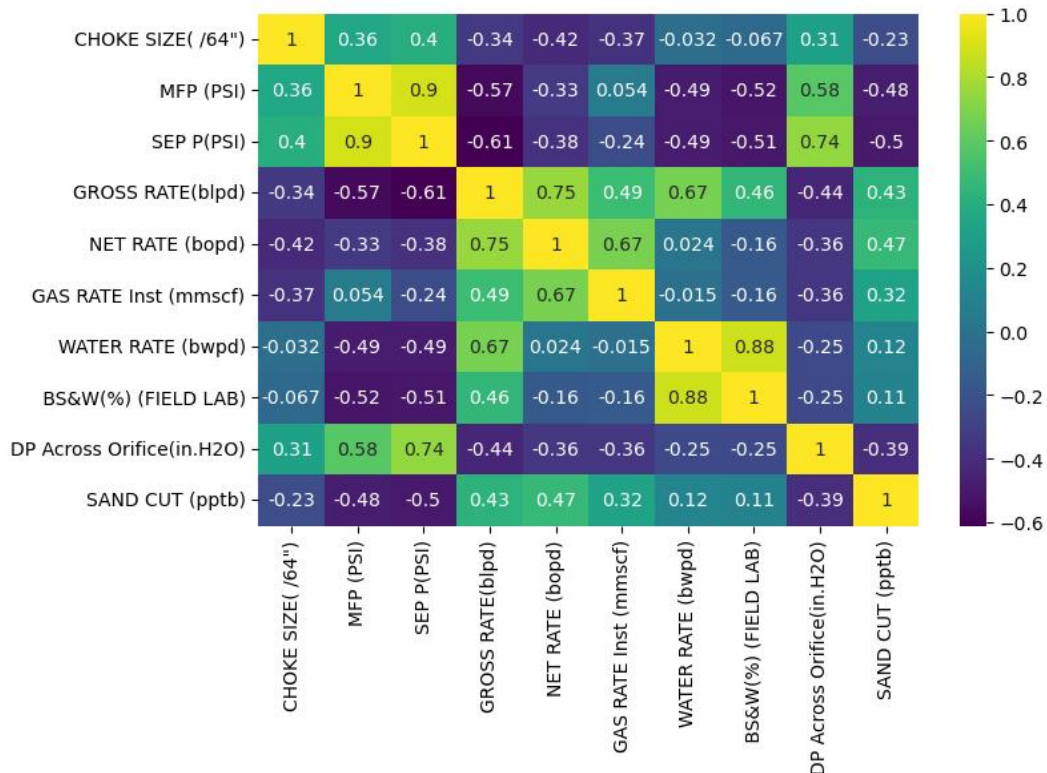


Figure 5. Pearson Correlation Chart

2.2. Machine learning model development

Figure 6 presents the workflow adopted for machine learning model development in this study. The dataset was first imported into a Python environment (Jupyter Notebook), where it underwent several pre-processing steps. Data cleaning was performed to address missing values and eliminate inconsistencies, ensuring the quality of inputs fed into the models. The cleaned dataset was then partitioned into training and testing sets using a 70:30 split, allowing the models to learn patterns from the majority of the data while preserving a portion for unbiased evaluation. Feature scaling was applied using the Standard Scaler, particularly for models sensitive to the magnitude of input features such as Support Vector Machines (SVM) and Artificial Neural Networks (ANN). This step was omitted for tree-based models like Random Forest (RF) and Extreme Gradient Boosting (XGB), which are inherently robust to feature scaling.

Model development was carried out using SVM, ANN, RF, and XGB algorithms. To optimize each model's performance, hyperparameter tuning was conducted using GridSearchCV with five-fold cross-validation. This approach systematically searched through specified parameter combinations to identify configurations that yielded the best learning outcomes. After training, the models were integrated into a framework that enabled comparison and further interpretability through feature importance analysis using SHAP values, enhancing transparency in understanding which input variables most influenced the predicted sand cut.

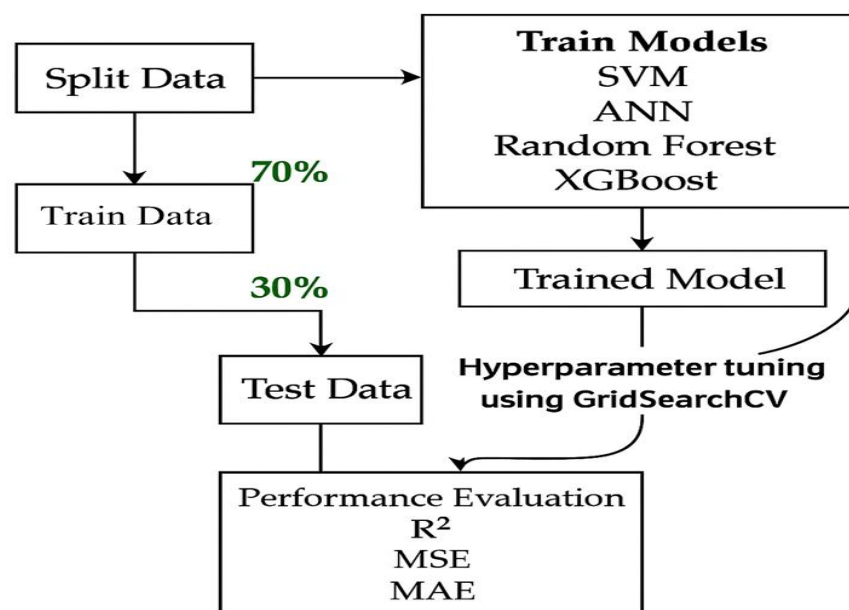


Figure 6. Machine Learning Workflow

2.2.1. Random forest (RF)

Random Forest (RF) is a widely used machine learning technique for building predictive models across various research domains. It operates on the principle of bagging, which combines bootstrapping and aggregation to enhance model performance (Ali et al., 2020). A key objective in predictive modelling is often to minimize the number of variables required for accurate predictions, thereby reducing data collection efforts and improving efficiency (Speiser et al., 2019). In RF, each decision tree is trained on a randomly selected subset of input data (a bootstrap sample) and develops independently. The model's accuracy stems

from the collective predictions of these individual trees, leveraging their combined strength. The RF process involves three main steps: (1) generating n bootstrap samples from the input variables, (2) constructing unpruned regression trees using the optimal predictor split, and (3) aggregating predictions from the n trees (Nosakhare et al., 2024).

For a random data subset $d(x, y)$ the decision tree iteratively partitions the variable space x as samples with close targets are aggregated. The data at node m is represented by d_m with N_m samples, and d_m may be partitioned into subsets denoted as d_m^{right} and d_m^{left} using each candidate split $\theta(p, t_m)$ where p and t_m refers to a feature and threshold, respectively as shown in equation 1.

$$\begin{cases} d_m^{left} = \{(x, y) | x_p \leq t_m\} \\ d_m^{right} = d_m / d_m^{left} \end{cases} \quad (1)$$

The candidate split is defined as a loss function $H(\bullet)$, as shown in equation 2.

$$H(d_m) = \frac{1}{N_m} \sum_{y \in d_m} (y - \bar{y}_m)^2 \quad (2)$$

$$\text{Where } \bar{y}_m = \frac{1}{N_m} \sum_{y \in d_m} y$$

The minimization of Equation (2) yields the parameters $\theta(p, t_m)$.

$$G(d_m, \theta) = \frac{N_m^{left}}{N_m} H(d_m^{left}(\theta)) + \frac{N_m^{right}}{N_m} H(d_m^{right}(\theta)) \quad (3)$$

The recursion of equation 3 continues for d_m^{right} and d_m^{left} until the maximum depth is achieved. The prediction of the RF model is thus obtained from equation 4.

$$f(x) = \frac{1}{K} \sum_{k=1}^K DT_k(x) \quad (4)$$

where K is the number of decision trees (DT) in the random forest.

2.2.2 Extreme gradient boosting (XGB)

Extreme Gradient Boosting (XGBoost) is a highly scalable and efficient tree-based machine learning algorithm widely adopted across diverse data analysis fields. As an advanced implementation of gradient boosting, XGBoost excels in both regression and classification tasks. Extreme gradient boosting is a type of boosting tree that has been used to determine nonlinear and local relationship in sand production prediction (Song et al., 2022). Its foundation lies in the boosting principle, which integrates the predictions of multiple weak learners through an additive training process to create a strong, accurate model. This approach not only boosts predictive performance but also reduces overfitting caused by insufficient data, small learning rate and large tree depth that captured noise rather than generalizable trends. XGB also enhances computational efficiency (Alshboul et al., 2022; Sheridan et al., 2016). The general function of the forecasting is set up at step p , as presented in equation 5

$$f_i^{(p)} = \sum_{k=1}^p f_k(x_i) = f_i^{(p-1)} + f_p(x_i) \quad (5)$$

where $f_p(x_i)$ denotes the learner at step p , $f_i^{(p)}$ denotes the prediction at p , $f_i^{(p-1)}$ denotes the prediction at $p-1$, and x_i denotes the input features.

To balance overfitting while maintaining computational efficiency, XGBoost employs a refined analytical formula, as shown in equation 6, to assess the model's "goodness of fit" to the original function. This method optimizes performance by controlling model complexity and improving predictive accuracy.

$$Objective^{(p)} = \sum_{k=1}^n l(\tilde{y}_i, y_i) + \sum_{k=1}^p \sigma(f_i) \quad (6)$$

where l presents the loss function, n presents the number of observations utilized, and σ presents the regularization term as represented in equation 7.

$$(f) = \theta T + 0.5\lambda\omega^2 \quad (7)$$

where ω expresses vector scores in leaves, T expresses the minimal loss necessary to divide the leaf node further, and λ expresses the regularization parameters.

2.2.3. Artificial neural network (ANN)

ANN is a computational model inspired by the structure and function of the biological neural networks that make up the brain. The ANN model used for this study was a feed-forward neural network consisting of n hidden layers each with a set of neurons that can estimate a given output response (y) from provided input data (x). Figure 7 illustrates the architecture of the Artificial Neural Network (ANN) model used in this study. The network consists of an input layer, four hidden layers, and a single output layer. Each hidden layer utilizes the ReLU (Rectified Linear Unit) activation function to introduce non-linearity and enhance model learning. The number of hidden layers and their respective neuron configurations were selected through hyperparameter optimization using GridSearchCV, aiming to achieve an optimal balance between model complexity and generalization performance. Dropout layers were not incorporated, as the dataset was sufficiently regularized through tuning and early stopping.

The mathematical concept of ANN involves a set of interconnected nodes or neurons that process and transmit information. Each neuron calculates a linear combination (z) of n input features (x_i) with corresponding weight (w_i) and bias (b) as shown in equation 8. The weights of the connections between the neurons are adjusted during training to minimize the error between the predicted output and the true output. The selection of the right activation function is very essential in the ANN modelling process as it is chiefly responsible for the mapping of the inputs to the outputs.

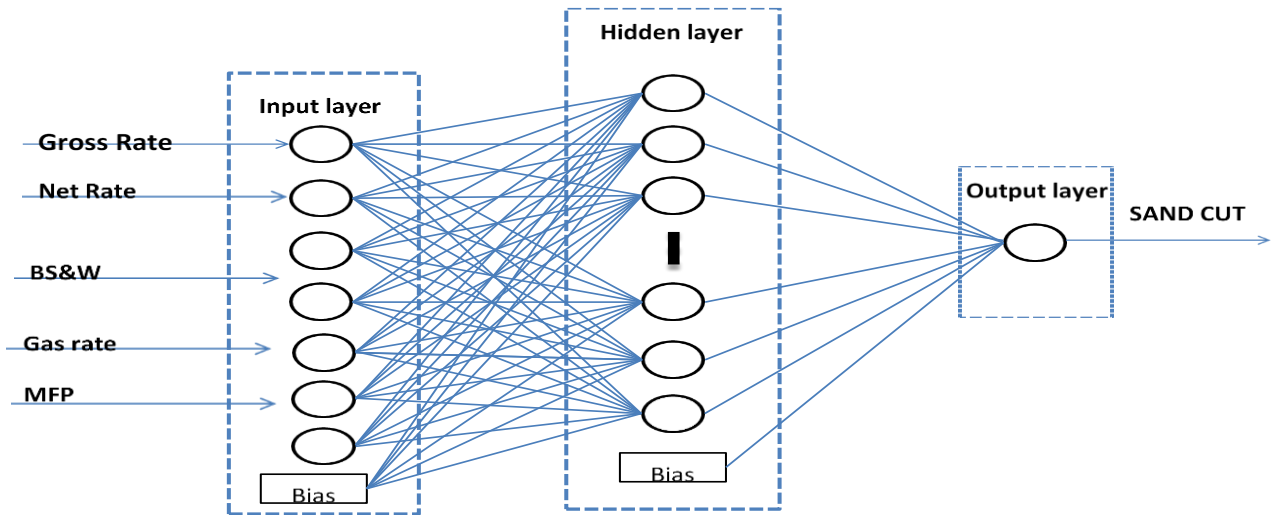


Figure 7. ANN Architecture for the Implemented Dataset

$$z = \sum_{i=1}^n x_i w_i + b \quad (8)$$

During the ANN process, the transformed version of the output data is forwarded from one hidden layer to another. The network is trained through an updating process that tunes the model weights to minimize model loss (L) which is typically taken as the root mean square error (RMSE) as shown in equation 9, where $\hat{y}_i$ is the model prediction. The minimization is done through a gradient descent method where the weights are manipulated as a function of L and scaled using the learning rate (λ) as shown in equation 10.

$$L = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (9)$$

$$w_{t+1} = w_t - \lambda \frac{\partial L}{\partial w_t} \quad (10)$$

To avoid overfitting, a k-fold cross-validation was carried out using the output dataset. The ANN model was implemented in Tensor Flow version 2.11.0 library. The ANN hyper parameters were optimized using a specially created Python script that compared the prediction accuracy of different networks trained with different combinations of hyper parameters.

2.2.4. Genetic algorithms

Adaptive heuristic search algorithms, or genetic algorithms (GAs) (Figure 8), make up the majority of evolutionary algorithms. Natural selection and genetics serve as the foundation for genetic algorithms. These are clever uses of sporadic searches that are aided by past data to focus the search on areas of the solution space where performance is higher. They are frequently employed to produce excellent solutions for search and optimization issues.

Because genetic algorithms mimic the process of natural selection, species that are able to adapt to changes in their surroundings have a better chance of surviving, procreating, and evolving into new generations. To

put it another way, they mimic the "survival of the fittest" among people from successive generations in order to address an issue. Every generation is made up of a population of individuals, and each individual stands for a potential solution or point in the search space. Every person is represented as a string made up of bits, characters, integers, and floats. The chromosome is comparable to this string.

2.3. SHapley additive explanations (SHAP)

SHAP (SHapley Additive exPlanations) is a unified framework for interpreting machine learning model outputs. It is based on game theory and assigns each feature an importance value for a particular prediction, offering insights into the contribution of individual input features to the model's predictions. This study employs SHAP to enhance the explainability of the machine learning models used for predicting sand cut.

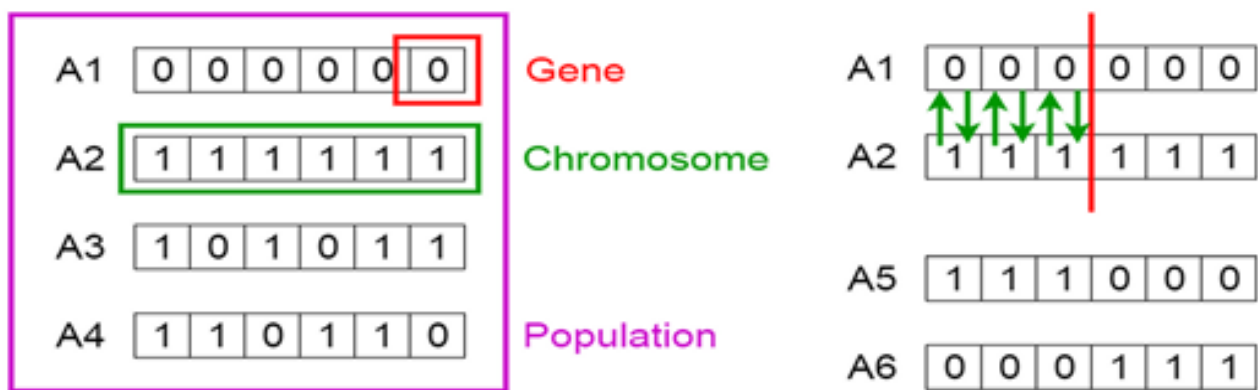


Figure 8. Simple Implementation of Genetic Algorithms

2.3.1. Theoretical framework

SHAP leverages concepts from cooperative game theory, particularly Shapley values, to fairly distribute the "payout" (i.e., prediction) among the "players" (i.e., input features). Each feature's contribution is calculated by considering all possible subsets of features, ensuring a comprehensive assessment of its impact on the prediction. The main properties of SHAP values include:

1. **Local Accuracy:** The sum of SHAP values for all features equals the model's prediction.
2. **Consistency:** If a model changes such that a feature contributes more to predictions, its SHAP value will not decrease.
3. **Missingness:** Features that do not influence the prediction will have a SHAP value of zero.

2.3.2. Implementation in this study

In this research, SHAP was applied to interpret the outputs of the three machine learning models namely; artificial neural Network (ANN), random forest (RF), and extreme gradient boosting (XGBoost). The SHAP framework was particularly valuable for:

1. **Identifying Key Features:** Determining the most influential parameters affecting sand production.
2. **Feature Interaction:** Exploring how different features interact and contribute to predictions.

3. **Global and Local Interpretations:** Understanding the overall behaviour of the model (global) and the specific reasoning behind individual predictions (local).

3. RESULTS AND DISCUSSION

The field production data was used to train multiple machine learning (ML) models, optimizing their respective hyperparameters to enhance predictive performance. Table 1 presents the optimal hyperparameters obtained for each model through a combination of 5-fold cross-validation and a full-factor experimental grid search. The performance of these models was evaluated using key statistical metrics, which are discussed in section 3.2.

3.1. Hyperparameter tuning

To ensure optimal model performance, Grid Search CV with 5-fold cross-validation was employed for hyperparameter tuning. The best-performing hyperparameters for each model are summarized in Table 2.

3.2. Model performance evaluation

Presented in Table 3 are the mean absolute error (MAE), mean squared error (MSE) and the R-squared or coefficient of determination (R^2), and these parameters helps to determine the differences between actual and predicted values as well as how well regression model explains the variance in the dependent variables. Among the models shown in Table 3, extreme gradient boosting (XGBoost) demonstrated the best performance, with the lowest values for mean absolute error and mean squared error as well as the highest R^2 , indicating its superior predictive capability for the dataset. Figure 9 and 10 showed the actual and predicted sand cut values using the extreme gradient boosting model and the artificial neural network model.

3.3. Sand production optimization

The relationship between net rate (bopd) and sand cut (pptb), as shown in Figure 11, reveals two distinct regimes. At lower net rates (<500 bopd), sand cut exhibits significant variability, likely due to unstable sand arches in the near-wellbore region that collapse intermittently under subcritical flow velocities (Tixier, 1949). However, beyond 500 bopd, sand cut stabilizes (Figure 11), suggesting that exceeding this critical flow velocity minimizes sand arch collapse and promotes consistent sand transport. This stabilization implies that maintaining net rates above 500 bopd (shaded as the 'Safe Zone' in Figure 11) could reduce unpredictability in sand production, offering a practical threshold for operational optimization. Larger choke openings increase production rates (and vice versa), enabling operators to target the stabilized 'Safe Zone' for sand management. While higher net rates improve predictability, further analysis is required to balance sand mitigation with reservoir pressure sustainability. These findings align with field observations where sustained high-rate production often correlates with reduced sand-related downtime.

3.4. SHAP sensitivity analysis

The SHAP analysis (Figure 12–13) quantifies the relative importance and directional effects of operational parameters on sand mitigation. Key findings include:

1. Prioritization of Critical Parameters:

- MFP (psi) emerged as the most influential parameter (SHAP value: +0.04), exerting 4× greater impact than Choke Size (/64") (SHAP value: +0.01) (Figure 12). This aligns with field observations where pressure fluctuations correlate strongly with sand ingress.
- Practical Insight: To mitigate sand production, operators should prioritize real-time MFP monitoring and stabilization protocols over secondary adjustments like choke size optimization.

2. Directionality of Feature Impact:

- The beeswarm plot (Figure 13) reveals that higher MFP values (associated with stable reservoir conditions) consistently reduce sand cut, whereas elevated gas rates correlate with increased sand risk.
- Field Action: Maintaining optimal MFP thresholds (e.g., within $\pm 5\%$ of baseline) is recommended to suppress sand activity, while gas rate surges should trigger mitigation workflows (e.g., choke throttling, screen inspections).

3. Interaction Effects (Future Work):

While the current SHAP analysis focuses on individual parameter impacts, the reviewer rightly highlights the need to explore **interaction effects** (e.g., how MFP and gas rate jointly influence sand dynamics). Although partial dependence plots (PDPs) are beyond the scope of this study, future work will integrate PDPs to quantify these synergies and refine multi-variable control strategies.

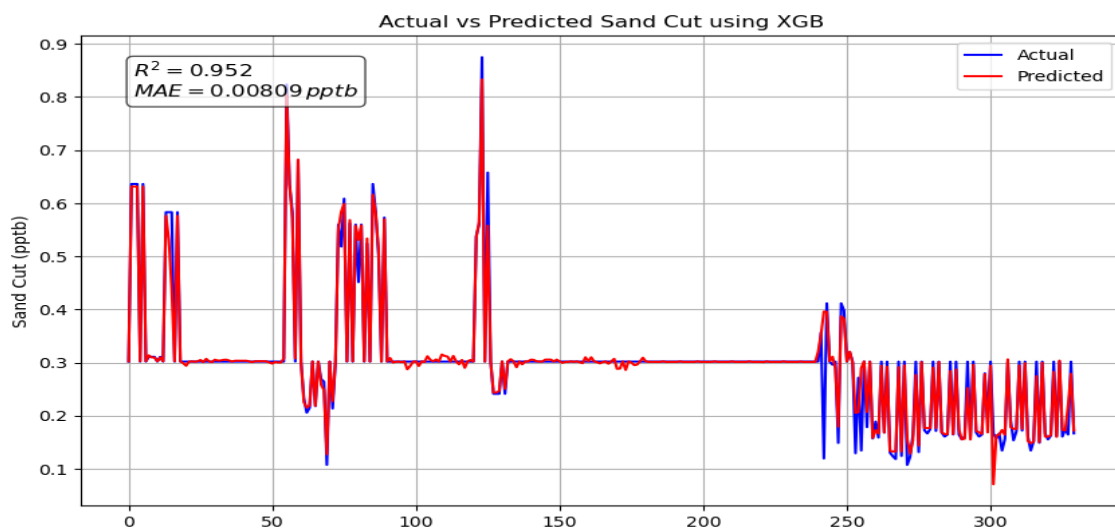


Figure 9. Actual vs Predicted Sand Cut Values using XGBoost

3.5. Optimization using genetic algorithm (GA)

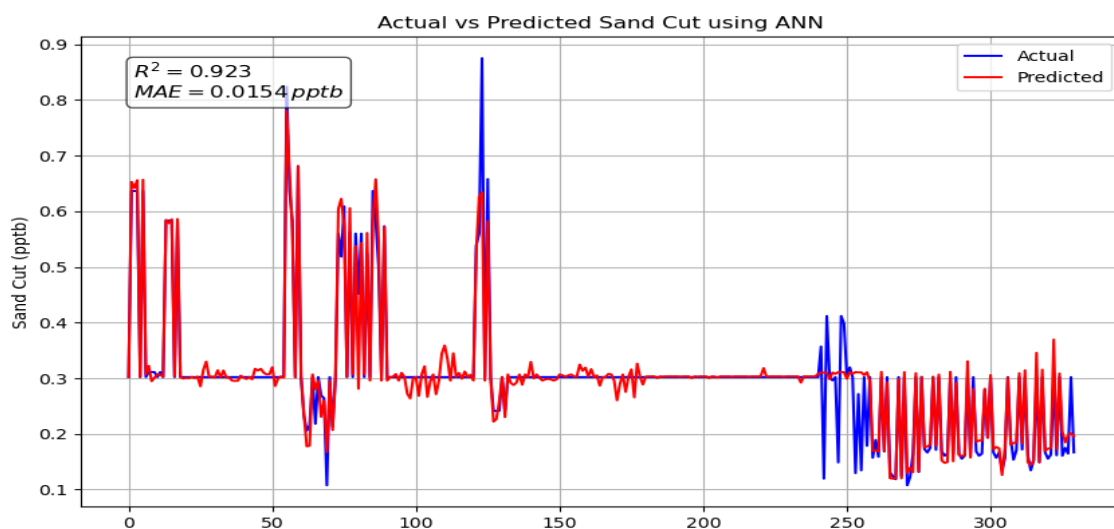
The Mealpy Genetic Algorithm (GA) was applied to optimize operational parameters to minimize sand production while maintaining production targets. The optimization objectives included maximizing gross rate and minimizing sand cut and BS&W.

Table 2. Optimal Hyperparameters for the Utilised Machine Learning Models

Model	Hyperparameters	Final Optimized Value
ANN	Kernel_initializer	Normal
	model optimizer	Lbfgs
	Epochs	100
	layer 1 number of neurons	32
	layer 1 activation function	Relu
	layer 2 number of neurons	64
	layer 2 activation function	Relu
	layer 3 number of neurons	16
	layer 3 activation function	Relu
	layer 4 number of neurons	8
	layer 4 activation function	Relu
RF	output layer number of neurons	1
	output activation function	Relu
XGB	n estimators	500
	max depth	5
	max features	3
	max depth	5
	γ	0.01
	n estimators	500
	Subsample	1
	random state	9

Table 3. Model Performance Metrics

Model	MAE	MSE	R ²
ANN	0.0154	0.0011	0.923
RF	0.0114	0.00086	0.942
XGBoost	0.0081	0.00058	0.952

**Figure 10.** Actual vs Predicted Sand Cut Values using ANN

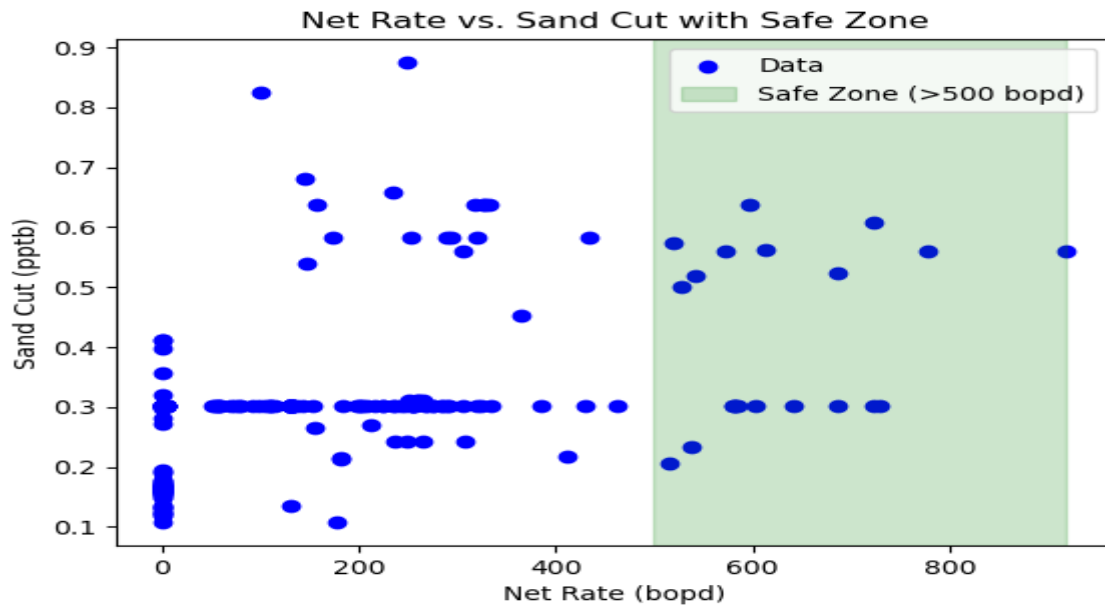


Figure 11. Sand Production Prediction

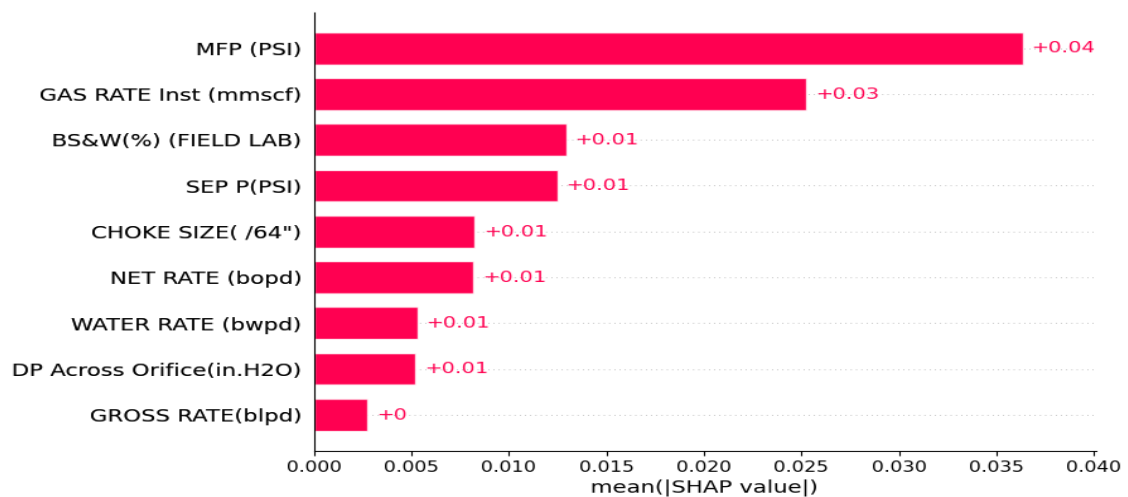


Figure 12. Feature importance plot using SHAP

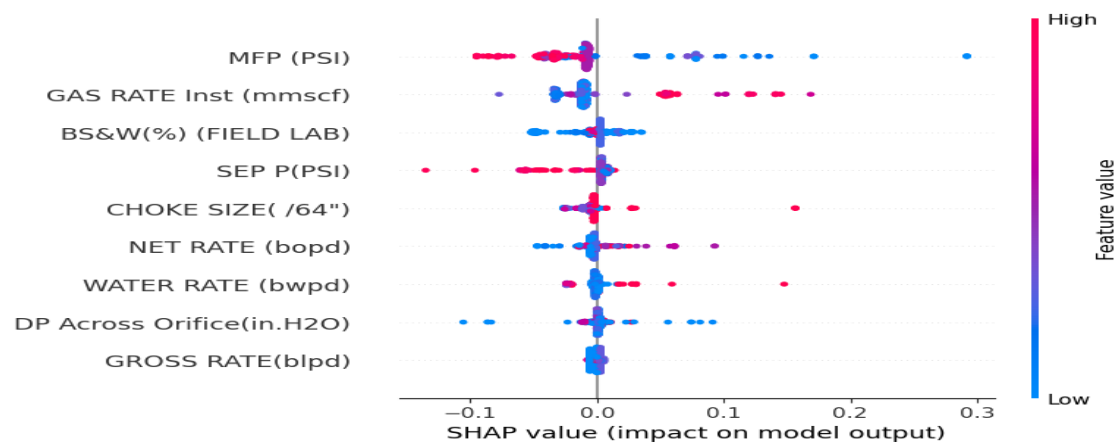


Figure 13. Summary plot of model impact on sand cut using SHAP

3.5.1. Optimization results

Table 4 shows the optimal parameters of the optimization algorithm used in minimising sand cut produced.

Table 4. *Optimal Parameters of the Genetic Algorithm*

Optimization Algorithm	Hyperparameters	Final Optimized Value
Genetic Algorithm	Epoch	100
	population size	100
	Pc	0.75
	Pm	0.25

The GA achieved an optimal configuration of parameters, yielding: optimal net rate of 750blpd minimized sand cut of 0.197pptb and minimized BS&W of 5.5%. The results indicate that GA effectively balances production efficiency with operational safety, reducing risks associated with excessive sand production. The genetic algorithm's convergence is shown in Figure 14, where the fitness value (representing sand mitigation efficacy) improves steadily over generations. This indicates robust optimization toward operational parameters that minimize sand production. The stabilization of fitness after ~40 generations suggests diminishing returns beyond this point, aligning with field observations that sand management gains plateau once critical thresholds (e.g., net rate >500 bopd) are sustained.

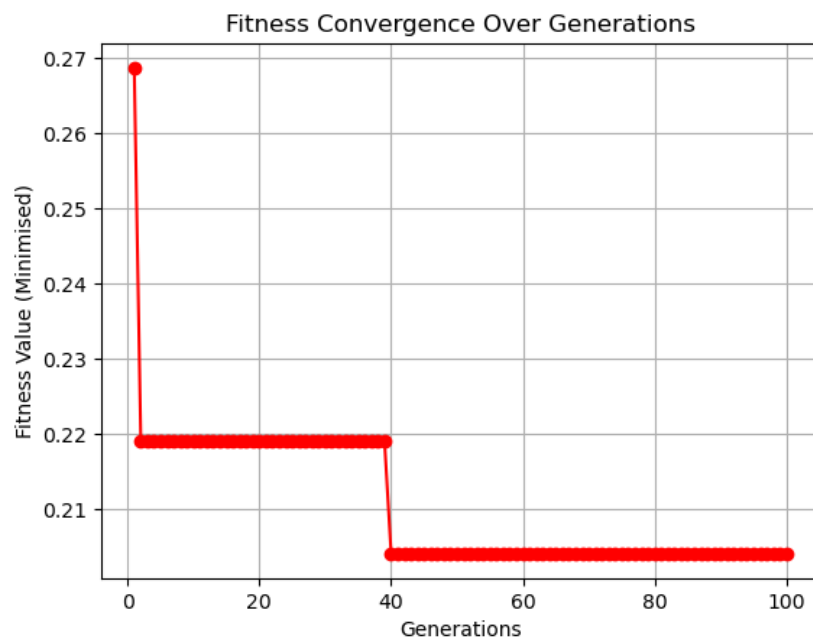


Figure 14. *Fitness convergence over 100 generations, demonstrating optimization progress toward maximizing sand mitigation*

3.6. Comparative analysis with previous studies

To further validate the performance of the developed XGBoost model, its predictive accuracy was compared with those reported in existing literature as shown in Figure 15. The XGBoost model in this study achieved an R^2 of 0.952, which is notably higher than those obtained in similar works. For example, Ketmalee and

Bandyopadhyay (2018) reported an R^2 of 0.908 using an Artificial Neural Network (ANN), while Ngwashi et al. (2021) also employed ANN but recorded a lower R^2 of 0.800. Song et al. (2022), who used a boosting tree algorithm similar to XGBoost, reported an R^2 of 0.940, slightly below the performance achieved in this study. Furthermore, Shabdirova et al. (2022) applied a Support Vector Machine (SVM) model and obtained an R^2 of 0.870. The superior performance of the present model can be attributed to the combination of advanced feature selection through SHAP analysis and hyperparameter tuning via GridSearchCV, as well as the integration of a Genetic Algorithm for optimization. These methodological enhancements have resulted in a more robust and reliable model capable of capturing the complex nonlinearities associated with sand production in oilfield operations.

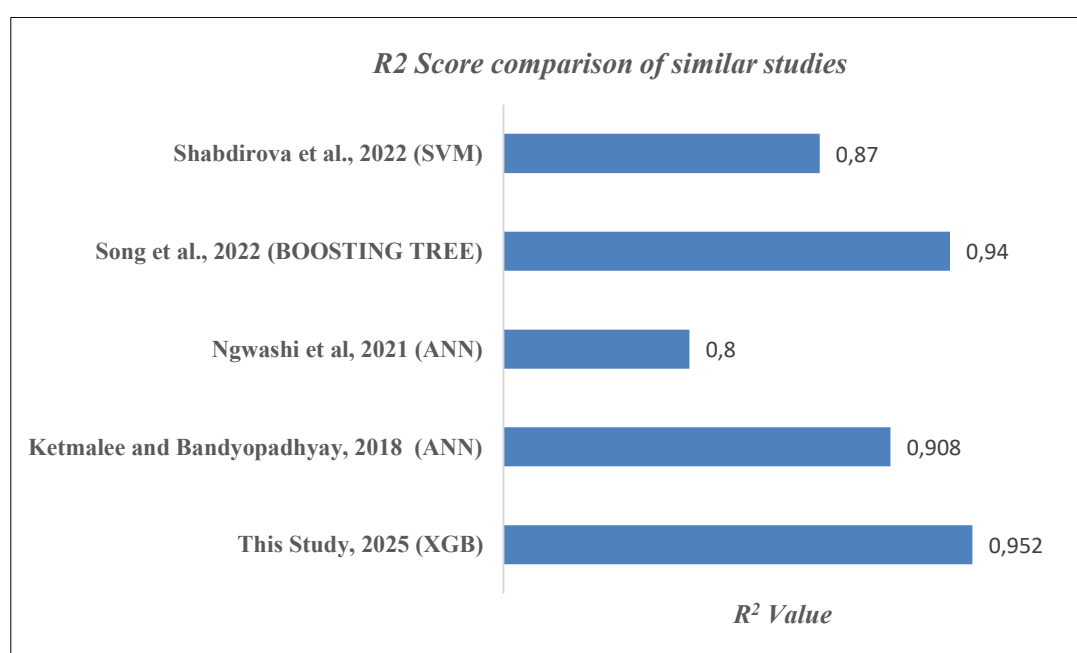


Figure 15. Comparison of R^2 values obtained and selected related works in sand production prediction

4. CONCLUSION

This study developed and validated an optimized machine learning framework to predict and minimize sand production in oilfield operations, with a focus on improving operational decision-making. Utilizing a multi-feature dataset obtained from eight wells across six different fields in the Niger Delta, the study implemented multiple supervised learning algorithms and employed advanced optimization techniques.

Among the models tested, Extreme Gradient Boosting (XGBoost) delivered the best performance, achieving an R^2 of 0.952, a mean absolute error (MAE) of 0.00809 pptb, and a mean squared error (MSE) of 0.00058 pptb². These results indicate its strong ability to capture complex nonlinear relationships inherent in sand production dynamics. Feature importance analysis conducted using SHAP values identified manifold flowing pressure (MFP), gas rate, and basic sediment and water (BS&W) as the most critical features influencing sand cut. These insights offer actionable guidance for production optimization and operational planning.

Furthermore, optimization using a Genetic Algorithm (GA) yielded a gross production rate of 750 blpd, BS&W of 5.5%, and minimized sand cut of 0.197 pptb as the ideal operational conditions. Sensitivity analysis further reinforced the critical role of MFP and gas rate, emphasizing the necessity of real-time monitoring to maintain effective sand control strategies. Consequently, the following inferences were arrived at from this study;

1. Three machine learning models, artificial neural network (ANN), random forest (RF) and extreme gradient boosting (XGBoost) were developed and evaluated.
2. Extreme gradient boosting (XGBoost) outperformed the others by achieving the highest R-squared value of 0.952 and the lowest mean absolute error (MAE) and mean squared error (MSE), demonstrating its superior accuracy in predicting sand cut values.
3. Feature importance analysis using SHAP values identified manifold flowing pressure (MFP), gas rate and BS&W are the most influential factors affecting sand production.
4. Optimization using a Genetic Algorithm (GA) successfully balanced production efficiency and sand control, yielding an optimal gross rate of 750blpd, minimized sand cut of 0.197pptb, and a reduced BS&W of 5.5%.
5. Sensitivity analysis further highlighted the critical role of manifold flowing pressure and gas rate, emphasizing the need for real-time monitoring to sustain effective sand control strategies optimal operational conditions.

Limitations and Future Work

- The dataset is limited to a specific set of wells within the Niger Delta region; expanding the model to include data from other regions and reservoirs would enhance generalizability.
- The integration of real-time sensor data is recommended to improve prediction accuracy and support dynamic decision-making.
- Future research should explore the use of hybrid and ensemble models, as well as incorporating domain-specific constraints and physics-based knowledge into machine learning architectures for more robust and interpretable predictions.

AUTHOR CONTRIBUTIONS

The three authors of this research study were involved in all the different stages of the work, and also contributed immensely in the reporting of this article for publication.

ACKNOWLEDGEMENT

The authors of this work acknowledge the immense contributions of Mr. Ifeanyi Cyriacus Azubuike and Mr. Shedrach Ikpeiminoghena Igemhokhai of the Department of Petroleum Engineering, University of Benin, Benin City during the research stages of this work.

CONFLICT OF INTEREST

The authors of this research article copiously and clearly state that there is no conflict of interest before the conceptualization of this work, during the different stages involved in the study and after the completion of the work.

REFERENCES

- Abdelghany, W. K., Hammed, M. S., Radwan, A. E., & Nassar, T. (2022). Implications of machine learning on geomechanical characterization and sand management: A case study from Hilal Field, Gulf of Suez, Egypt. *Journal of Petroleum Exploration and Production Technology*. <https://doi.org/10.1007/s13202-022-01551-9>
- Ali, M., Prasad, R., Xiang, Y. & Yaseen, Z. M. (2020). Complete Ensemble Empirical Mode Decomposition Hybridized with Random Forest and Kernel Ridge Regression Model for Monthly Rainfall Forecasts, *J. Hydrol.* 584(2020) 124647, <https://doi.org/10.1016/J.JHYDROL.2020.124647>
- Al-Shaabi, S. K., Al-Ajmi, A. M., & Al-Wahaibi, Y. (2013). Three dimensional modeling for predicting sand production. *Journal of Petroleum Science and Engineering*, 109, 348–363.
- Alshboul, O., Shehadeh, A., Almasabha, G., & Almuflih, A. S. (2022). Extreme Gradient Boosting-Based Machine Learning Approach for Green Building Cost Prediction. *Sustainability*, 14(11), 6651. <https://doi.org/10.3390/su14116651>
- Ani, M., Oluyemi, G., Petrovski, A., & Rezaei-Gomari, S. (2016). Reservoir Uncertainty Analysis: the Trends from Probability to Algorithms and Machine Learning. *SPE Intelligent Energy International Conference and Exhibition*, 6-8 September 2016, Aberdeen, UK.
- Anifowose, F. A., Labadin, J. & Abdulraheem, A. (2017a). Hybrid Intelligent Systems in Petroleum Reservoir Characterization and Modeling: The Journey so Far and the Challenges Ahead. *Journal of Petroleum Exploration and production Technology*, 7(1) (2017), 251-263.
- Anifowose, F. A., Labadin, J. & Abdulraheem, A. (2017b). Ensemble Machine Learning: An Untapped Modeling Paradigm for Petroleum Reservoir Characterization. *J. Petrol. Sci. Eng.*, 151(2017), 480-487.
- Anirbid, S., Kriti, Y., Kamakshi, R., Namrata, B. & Hemangi, O. (2021). Application of Machine Learning and Artificial Intelligence in Oil and Gas Industry. *Petroleum Research*, 6(4), 379-391. <https://doi.org/10.1016/j.ptlrs.2021.05.009>
- Azad M., Zargar G. & Arabjamaloei, R. (2011). A New Approach to Sand Production Onset Prediction Using Artificial Neural Networks. *Petroleum Science and Technology*, 29:19, 1975-1983. <https://doi.org/10.1080/10916460903551081>
- Carlson, J., Gurley, D., King, G., Price-Smith, C., & Waters, F. (1992). Sand Control: Why and How. *Oilfield Review*; Netherlands, 4(4).
- Completion Technology for Unconsolidated Formations, Revision 2, (1995)

Completion Technology for Unconsolidated Formations, Revision 3, (1997). <https://fr.scribd.com>

Dickson, R. L., Raymond, C. A., Joe, W., & Wilkinson, C. A. (2003). Segregation of Eucalyptus Dunnii Logs using Acoustics. *Forest Ecology and Management*, 179(1-3), 243-251.

Fattahpour, V., Moosavi, M., & Mehranpour, M. (2012). An experimental investigation on the effect of rock strength and perforation size on sand production. *Journal of Petroleum Science and Engineering*, 86–87, 172–189.

Fuh, G., & Morita, N. (2013). Sand production prediction analysis of heterogeneous reservoirs for sand control and optimal well completion design. *International Petroleum Technology Conference*. <https://doi.org/10.2523/16940-MS>

Gharagheizi, F., Mohammadi, A., Arabloo, M., & Shokrollahi, A. (2017). Prediction of sand production onset in petroleum reservoirs using a reliable classification approach. *Petroleum*, 3(2), 280–285. <https://doi.org/10.1016/j.petlm.2016.02.001>

Han, G., Shepstone, K., Harmawan, I., Er, U., Jusoh, H., Lin, L. S., Pringle, D., et al. (2011). A comprehensive study of sanding rate from a gas field, from reservoir to completion, production, and surface facilities. *SPE Journal*, 16(2), 463–481.

van den Hoek, P. J., Hertogh, G. M. M., Kooijman, A. P., de Bree, Ph., Kenter, C. J., & Papamichos, E. (2007). A new concept of sand production prediction, theory and laboratory experiments. *SPE Drilling & Completion*, 15(4), 261–273.

Kanj, M. Y. & Abousleiman, Y. (1999). Realistic Sanding Predictions: A Neural Approach. Paper Presented at the *SPE Annual Technical Conference and Exhibition*, October 3–6, 1999. SPE-56631-MS. <https://doi.org/10.2118/56631-MS>

Kessler, N., Wang, Y., & Santarelli, F. J. (1993). A simplified pseudo 3D model to evaluate sand production risk in deviated cased holes. *SPE Annual Technical Conference and Exhibition*.

Ketmalee, T., & Bandyopadhyay, P. (2018). Application of neural network in formation failure model to predict sand production. In *Offshore Technology Conference Asia 2018* (pp. 1–10). <https://doi.org/10.4043/28506-ms>

Khamehchi, E. Kivi, I. R. & Akbari, M. (2014). A Novel Approach to Sand Production Prediction using Artificial Intelligence. *Journal of Petroleum Science and Engineering*. 123(2014), 147-154. <https://doi.org/10.1016/j.petrol.2014.07.033>

Kozhagulova, A., Shabdirova, A., Minh, N. H., & Zhao, Y. (2021). An integrated laboratory experiment of realistic diagenesis, perforation and sand production using a large artificial sandstone specimen. *Journal of Rock Mechanics and Geotechnical Engineering*.

- Morita, N., Whitfill, D. L., Fedde, O. P., & Lovik, T. H. (1989). Parametric study of sand-production prediction: Analytical approach. *SPE Production Engineering*, 4(1), 25–33. <https://doi.org/10.2118/16990-PA>
- Ngwashi, A. R., Ogbe, D. O., & Udebhulu, D. O. (2021). Evaluation of machine-learning tools for predicting sand production. In *SPE Nigeria Annual International Conference and Exhibition 2021, NAIC 2021*, 1–16. <https://doi.org/10.2118/207193-MS>
- Nosakhare, A., Igemhokhai, S., Aimhanesi, S., Ugboodu, F. & Iyore, N. (2024). Heliyon Data-Driven Intelligent Modeling, Optimization, and global sensitivity analysis of a xanthan gum biosynthesis process. *Heliyon*, 10(3), e25432. <https://doi.org/10.1016/j.heliyon.2024.e25432>
- Nouri, A., Vaziri, H., Belhaj, H., & Islam, R. (2006). Sand-production prediction, a new set of criteria for modeling based on large-scale transient experiments and numerical investigation. *SPE Journal*, 11(2), 26–29.
- Papamichos, E., & Furui, K. (2019). Analytical models for sand onset under field conditions. *Journal of Petroleum Science and Engineering*, 172, 171–189.
- Shabdirova, A., Minh, N. H., & Zhao, Y. (2022). Role of plastic zone porosity and permeability in sand production in weak sandstone reservoirs. *Underground Space (China)*. <https://doi.org/10.1016/j.undsp.2021.10.005>
- Shabdirova, A., Kozhagulova, A., Samenov, Y., & others. (2024). Sand production prediction with machine learning using input variables from geological and operational conditions in the Karazhanbas oilfield, Kazakhstan. *Natural Resources Research*, 33, 2789–2805. <https://doi.org/10.1007/s11053-024-10389-3>
- Sheridan, R. P., Wang, W. M., Liaw, A., Mu, J. & Gifford, E. M. (2016). Extreme Gradient Boosting as a Method for Quantitative Structure-Activity Relationship. *Journal of Chemical Information and Modeling*, 56(12), 2353-2360. <https://doi.org/10.1021/acs.jcim.6b00591>
- Skjaerstein, A., Tronvoll, J., Santarelli, F. J., & Joranson, H. (1997). Effect of water breakthrough on sand production, experimental and field evidence. In *Proceedings - SPE Annual Technical Conference and Exhibition Pi*, 565–575.
- Song, J., Li, Y., Liu, S., Xiong, Y., Pang, W., He, Y. & Mu, Y. (2022). Comparison of Machine Learning Algorithms for Sand Production Prediction: An Example for a Gas-Hydrate-Bearing Sand Case. *Energies* 2022, 15(18), 6509. <https://doi.org/10.3390/en15186509>
- Speiser, J. L., Miller, M. E., Tooze, J. & Ip, E. (2019). A Comparison of Random Forest Variable Selection Methods for Classification Prediction Modeling. *Expert Systems with Applications*, 134, 93–101. <https://doi.org/10.1016/j.eswa.2019.05.028>
- Tiab, D. & Donaldson, E. C. (2004). Petrophysics: Theory and Practice of Measuring Reservoir Rock and Fluid Transport Properties, *Gulf Professional Publishing, Elsevier* (2nd. Ed., pp. 554-670)

Tixier, M.P. (1949) Evaluation of Permeability from Log Resistivity Gradients. *Oil and Gas Journal*, 48, 113-122.

Wang, Y., & Dusseault, M. B. (2010). Sand production potential near inclined perforated wellbores. In *47th Annual Technical Meeting of the Petroleum Society*. <https://doi.org/10.2118/9670>

Weingarten, J. S., & Perkins, T. K. (2007). Prediction of sand production in gas wells, methods and Gulf of Mexico case studies. *Journal of Petroleum Technology*, 47(7), 596–600.

Wu, B., Choi, S. K., Denke, R., Barton, T., Viswanathan, C., Lim, S., Zamberi, M., & Shaffee, S. (2016). A new and practical model for amount and rate of sand production. *Offshore Technology Conference*.



Some Applications of Berezin Radius Inequalities

Hamdullah BAŞARAN^{1*} 

¹ Antalya, Türkiye

Keywords	Abstract
Reproducing Kernel Hilbert Space Operator Norm Berezin Symbol Polar Decomposition Contraction Operator	We show a number of inequalities in the reproducing kernel Hilbert space (RKHS) in this study. Stronger boundaries between the Berezin radius and the numerical radius on the Berezin number of the bounded linear operator described in the RKHS than those found in the literature are obtained by using a few auxiliary theorems while analyzing the Berezin radius.
Cite	
Başaran, H. (2025). Some Applications of Berezin Radius Inequalities. <i>GU J Sci, Part A</i> , 12(2), 562-582. doi:10.54287/gujsa.1655950	
Author ID (ORCID Number)	Article Process
0000-0002-9864-9515 Hamdullah BAŞARAN	Submission Date 11.03.2025 Revision Date 12.05.2025 Accepted Date 28.05.2025 Published Date 30.06.2025

1. INTRODUCTION

We represent complex-valued functions on an existing set X in a Hilbert space $\mathcal{H} = \mathcal{H}(X)$ which means evaluation at any given point in X . As a reproducing kernel Hilbert space, we establish a Hilbert space $\mathcal{H} = \mathcal{H}(X)$ of complex-valued functions on a given set X such that evaluation at each point of X is a continuous functional on H . The Hilbert function space $\mathcal{H}$ has been shown to have a reproducing kernel according to the Riesz representation theorem. This means that for all $\rho \in X$, is present a function $k_\rho(z) \in \mathcal{H}$ which means $\langle f, k_\rho \rangle = f(\rho)$ for all $f \in \mathcal{H}$ and $\rho \in X$. The reproducing kernel of the space H is the term presented to this function. We demonstrate the normalized reproducing kernel by $\hat{k}_\rho = \frac{k_\rho}{\|k_\rho\|}$. The Dirichlet space $\mathcal{D}^2(\mathbb{D})$, Hardy space $\mathcal{H}^2(\mathbb{D})$, Bergman space $\mathcal{L}_\alpha^2(\mathbb{D})$, where $\mathbb{D} = \{z \in \mathbb{C}: |z| < 1\}$ is the unit disc, and the Fock space $\mathcal{F}(\mathbb{C})$ are the important RKHSs. For instance, Aronzajn (1950) extends a comprehensive account of the theory of RKHSs and reproducing kernels. Reproducing kernels plays an essential role in a number of pure and applied mathematics fields, which includes wavelets, frame theory, fractal theories, and signals (for example, see Jorgensen's book (Jorgensen, 2006) and its references).

The C^* -algebra of any bounded linear operator described on a complex Hilbert space $(\mathcal{H}, \langle \cdot, \cdot \rangle)$ is shown with symbol $\mathcal{L}(\mathcal{H})$. For each of the bounded linear operator N on $\mathcal{H}$ (that is, for $N \in \mathcal{L}(\mathcal{H})$) its Berezin symbol $\tilde{N}$

*Corresponding Author, e-mail: 07hamdullahbasaran@gmail.com

is denoted by (can observe, Berezin (1972) and Karaev (2006))

$$\tilde{N} = \langle N\hat{k}_\rho, \hat{k}_\rho \rangle, \rho \in X.$$

Because of the Cauchy-Schwarz inequality $|\tilde{N}(\rho)| \leq \|N\hat{k}_\rho\| \leq \|N\|$ for all $\rho \in X$, that scalar-valued function is bounded on X . An operator's Berezin symbol constitutes significant details with regard to the operator. As an example, has become commonly accepted that each of the operator within the RKHSs that is analytic functions (like the Dirichlet, Fock, Hardy, and Bergman spaces) is uniquely established by the Berezin symbol, i.e., $N_1 = N_2$ iff $\tilde{N}_1 = \tilde{N}_2$ (see, for example, Engliš (1995) and Zhu (1990)).

The Berezin set corresponding to the operator $N \in \mathcal{L}(\mathcal{H})$ is described

$$Ber(N) = Range(\tilde{N})$$

and Berezin radius of N is the number defined by

$$ber(N) = \sup\{|\xi|: \xi \in Ber(N)\}$$

(see Karaev (2006, 2013)). The Berezin norm of operators $N \in \mathcal{L}(\mathcal{H})$ is denoted by

$$\|N\|_{Ber} = \sup_{\rho \in X} \|N\hat{k}_\rho\|.$$

It becomes simplicity to demonstrate that actually $\|N\|_{Ber}$ chooses a new operator norm in $\mathcal{L}(\mathcal{H}(X))$ (since the set of reproducing kernels $\{k_\rho: \rho \in X\}$ span the space $\mathcal{H}(X)$). It is straightforward that $ber(N) \leq \|N\|_{Ber} \leq \|N\|$ (refer to Aronzajn (1950) and Berezin (1972) for further details on reproducing kernel Hilbert spaces and the Berezin symbol).

The following are given $W(N)$ and also $w(N)$ which stand for the numerical range and numerical radius of N :

$$W(N) = \{\langle Nf, f \rangle: f \in \mathcal{H} \text{ and } \|f\|_{\mathcal{H}} = 1\}$$

and

$$w(N) = \sup\{|\langle Nf, f \rangle|: f \in \mathcal{H} \text{ and } \|f\|_{\mathcal{H}} = 1\}.$$

It is clear that $Ber(N) \subseteq W(N)$ and $ber(N) \leq w(N)$ (the numerical radius of operator N). There are some intriguing qualities of an operator's numerical range. As an instance, it is widely acknowledged that the closure

of an operator's numerical range includes its spectrum. For essential characteristics of the numerical radius, we refer to Abu-Omar and Kittaneh (2015), Bhatia (2007), Bhunia et al. (2022), Bhunia et al. (2023a), Furuta, (2001), Kittaneh (1952), Kittaneh et al. (2015), Stojiljkovic and Gürdal (2025), Sababheh and Moradi (2023), Yamakazi (2007), Yang and Xu (2023).

It is often known that

$$\frac{1}{2}\|N\| \leq w(N) \leq \|N\| \quad (1)$$

and

$$ber(N) \leq w(N) \leq \|N\| \quad (2)$$

for any $N \in \mathcal{L}(\mathcal{H}(X))$. Hence, studying of the recently discovered numerical properties $Ber(N)$, $ber(N)$, and $\|N\|_{Ber}$, we will be essential for a detailed examination of both this numerical radius and range of operators based the RKHSs.

For $N \in \mathcal{L}(\mathcal{H})$, the Berezin norm of N is defined as:

$$\|N\|_{ber} = \sup\{|\langle N\hat{k}_\rho, \hat{k}_\varphi \rangle|, \rho, \varphi \in X\},$$

Here, $\hat{k}_\rho$ and $\hat{k}_\varphi$ represent normalized reproducing kernels of the space $\mathcal{H}$. It is essential to remember that the norm $\|\cdot\|_{ber}$ does not always imply the submultiplicative property. See Bhunia et al. (2023b), Karaev et al. (2016), Karaev et al. (2011) and Yamancı et al. (2020) for an examination of the fundamental properties and facts with respect to Berezin number.

The Cauchy-Schwarz inequality, which indicates that

$$|\langle x, y \rangle| \leq \|x\| \|y\|, \quad x, y \in \mathcal{H} \quad (3)$$

is belongs to the foremost significant and useful inequalities in operator theory. Buzano (1974) established an extension (3) involving it has been shown that

$$|\langle x, e \rangle \langle e, y \rangle| \leq \frac{1}{2} (\|x\| \|y\| + |\langle x, y \rangle|) \quad (4)$$

for any $x, y, e \in \mathcal{H}$ with $\|e\| = 1$. It is easy to see that (4) may be expressed for any $e \in \mathcal{H}$ in the form as follows:

$$|\langle x, e \rangle \langle e, y \rangle| \leq \frac{\|e\|^2}{2} (\|x\| \|y\| + |\langle x, y \rangle|). \quad (5)$$

Using the (3) yields

$$|\langle x, e \rangle \langle e, y \rangle| \leq \|e\|^2 \|x\| \|y\|,$$

which is important to remember here. As a consequence, (5) includes a more appealing bound than the one which ensues from using (3) twice. In close connection with Buzano's inequality (4), Dragomir (2016) presented that

$$|\langle Tx, Ty \rangle| \leq \frac{\|T\|^2}{2} (\|x\| \|y\| + |\langle x, y \rangle|).$$

for any $T \in \mathcal{L}(\mathcal{H})$.

In the present article, we prove several inequalities in RKHS. When analyzing the radius of Berezin we find stronger bounds on $ber(N)$ than those obtained in inequality (2). Following that, that is utilized Lemma 2.8 for demonstrating the final inequality.

2. MATERIAL AND METHOD

The following lemmas from the literature will be required in order to reach our conclusions. The first lemma is proven by Sababheh et al. (2024).

Lemma 2.1. Let $P, R, S \in \mathcal{L}(\mathcal{H})$, where P and R are positive. Then the next statements are equivalent to one another:

- (i) $\begin{bmatrix} P & S \\ S^* & R \end{bmatrix}$ is a positive operator in $\mathcal{L}(\mathcal{H} \oplus \mathcal{H})$.
- (ii) $\begin{bmatrix} P & S^* \\ S & R \end{bmatrix}$ is a positive operator in $\mathcal{L}(\mathcal{H} \oplus \mathcal{H})$.
- (iii) $|\langle Sx, y \rangle| \leq \langle Px, x \rangle \langle Ry, y \rangle$ for all $x, y \in \mathcal{H}$.
- (iv) There exists a contraction O (i.e. $\|O\| \leq 1$) such that $S = R^{\frac{1}{2}} O P^{\frac{1}{2}}$.

The following two lemmas are proven by Kittaneh (1997; 1988).

Lemma 2.2. Let $P, R \in \mathcal{L}(\mathcal{H})$ be a positive operators. Then

$$\|P + R\| \leq \max\{\|P\|, \|R\|\} + \left\|R^{\frac{1}{2}}P^{\frac{1}{2}}\right\|.$$

Lemma 2.3. Let $P, R, S \in \mathcal{L}(\mathcal{H})$, where P and R are positive and $RS = SP$. If $\begin{bmatrix} P & S^* \\ S & R \end{bmatrix}$ is positive in $\mathcal{L}(\mathcal{H} \oplus \mathcal{H})$, then $\begin{bmatrix} f^2(P) & S^* \\ S & g^2(R) \end{bmatrix}$ is also positive, where f, g are nonnegative continuous functions on $[0, \infty)$, such that $f(t)g(t) = t$ for every $t \in [0, \infty)$.

Since $P|P| = |P^*|P$ (see, (Pečarić et al., 2005)) Lemma 2.3, implies

$$\begin{bmatrix} f^2(|P|) & P^* \\ P & g^2(|P^*|) \end{bmatrix} \geq 0 \text{ for any } P \in \mathcal{L}(\mathcal{H}). \quad (6)$$

The following lemma is obtained by use to inequality (7) and a more elaborate version inequality (6) which is associated with the product of three operators.

Lemma 2.4. Let $P, R, S \in \mathcal{L}(\mathcal{H})$ and S positive. Then

$$|\langle PSRx, y \rangle| \leq \frac{\|S\|^2}{2} (|\langle PRx, y \rangle| + \|Rx\| \|P^*y\|) \quad (7)$$

for any $x, y \in \mathcal{H}$ (see, (Sababheh et al., 2024)).

We provide a new equivalent assertion about the positivity of a specific operator matrix in the following theorem, which may be added to the claims in Lemma 2.1.

Lemma 2.5. Let $P, R, S \in \mathcal{L}(\mathcal{H})$ and P and R positive and any $x, y \in \mathcal{H}$. Then $\begin{bmatrix} P & S^* \\ S & R \end{bmatrix}$ is a positive operator in $\mathcal{L}(\mathcal{H} \oplus \mathcal{H})$ iff

$$|\langle Sx, y \rangle| \leq \sqrt{\langle R^{\frac{1}{2}}|O|R^{\frac{1}{2}}x, x \rangle \langle P^{\frac{1}{2}}|O^*|P^{\frac{1}{2}}y, y \rangle} \quad (8)$$

for some contraction O (see, (Sababheh et al., 2024)).

This is the conclusion that follows directly from Lemma 2.5 and inequality (6).

Lemma 2.6. Let $N \in \mathcal{L}(\mathcal{H})$ and let $x, y \in \mathcal{H}$. If f, g are nonnegative continuous on $[0, \infty)$ obtaining

$f(t)g(t) = t$ ($t > 0$), then

$$|\langle Nx, y \rangle| \leq \sqrt{\langle g(|N^*|)|O|^{\frac{1}{2}}g(|N^*|)x, x \rangle \langle f(|N|)|O^*|^{\frac{1}{2}}f(|N|)y, y \rangle} \quad (9)$$

for some contraction O (see, (Sababheh et al., 2024)).

The following two lemmas are proven by using Lemma 2.1.

Lemma 2.7. Let $N \in \mathcal{L}(\mathcal{H})$ with the polar decomposition $N = U|N|$, where U is a partial isometry, with $0 \leq r \leq 1$. Then for any $x, y \in \mathcal{H}$,

$$|\langle Nx, y \rangle| \leq \frac{\|N\|^{\frac{1}{2}}}{2} \left(|\langle U|N|^{\frac{1}{2}}x, y \rangle| + \sqrt{\langle |N|^r x, x \rangle \langle |N^*|^{1-r} y, y \rangle} \right) \quad (10)$$

and

$$|\langle Nx, y \rangle| \leq \frac{\|N\|^{1-r}}{2} \left(|\langle U|N|^r x, y \rangle| + \sqrt{\langle |N|^r x, x \rangle \langle |N^*|^r y, y \rangle} \right) \quad (11)$$

(see, (Sababheh et al., 2024)).

Lemma 2.8. Let $N \in \mathcal{M}_n$. If f, g are nonnegative continuous based $[0, \infty)$ obtaining $f(t)g(t) = t$ ($t > 0$), then for all vectors $x, y \in \mathbb{C}$,

$$(i) \quad |\langle \Re Nx, y \rangle| \leq \frac{1}{2} \sqrt{\langle (f^2(|N|) + f^2(|N^*|))x, x \rangle \langle (g^2(|N|) + g^2(|N^*|))y, y \rangle} \quad (12)$$

and

$$(ii) \quad |\langle \Im Nx, y \rangle| \leq \frac{1}{2} \sqrt{\langle (f^2(|N|) + f^2(|N^*|))x, x \rangle \langle (g^2(|N|) + g^2(|N^*|))y, y \rangle} \quad (13)$$

(see, (Sababheh et al., 2024)).

3. MAIN RESULTS

In this part of the paper, we obtain Berezin radius inequality. We offer some bounds for the Berezin radius.

Let's present the first theorem.

Theorem 3.1. Assume $\mathcal{H} = \mathcal{H}(X)$ to be an RKHS. Given the polar decomposition $N = U|N|$, let $N \in \mathcal{H}$. Then for any $0 \leq r \leq 1$,

$$\text{ber}(N) \leq \frac{\|N\|^{\frac{1}{2}}}{2} \left(\text{ber}\left(U|N|^{\frac{1}{2}}\right) + \frac{1}{2} \| |N|^r + |N^*|^{1-r} \|_{\text{ber}} \right) \quad (14)$$

and

$$\text{ber}(N) \leq \frac{\|N\|^{1-r}}{2} \left(\text{ber}(U|N|^r) + \frac{1}{2} \| |N|^r + |N^*|^r \|_{\text{ber}} \right), \quad (15)$$

where U is a partial isometry.

Proof. Suppose $\hat{k}_\rho$ be a reproducing kernel that has been normalized. If we write $x = y = \hat{k}_\rho$ in the inequality (10), then we get

$$\begin{aligned} |\langle N\hat{k}_\rho, \hat{k}_\rho \rangle| &\leq \frac{\|N\|^{\frac{1}{2}}}{2} \left(\left| \langle U|N|^{\frac{1}{2}}\hat{k}_\rho, \hat{k}_\rho \rangle \right| + \sqrt{\langle |N|^r\hat{k}_\rho, \hat{k}_\rho \rangle \langle |N^*|^{1-r}\hat{k}_\rho, \hat{k}_\rho \rangle} \right) \\ &\leq \frac{\|N\|^{\frac{1}{2}}}{2} \left(\left| \langle U|N|^{\frac{1}{2}}\hat{k}_\rho, \hat{k}_\rho \rangle \right| + \frac{\langle |N|^r\hat{k}_\rho, \hat{k}_\rho \rangle + \langle |N^*|^{1-r}\hat{k}_\rho, \hat{k}_\rho \rangle}{2} \right) \\ &\leq \frac{\|N\|^{\frac{1}{2}}}{2} \left(\left| \langle U|N|^{\frac{1}{2}}\hat{k}_\rho, \hat{k}_\rho \rangle \right| + \frac{1}{2} \langle (|N|^r + |N^*|^{1-r})\hat{k}_\rho, \hat{k}_\rho \rangle \right) \end{aligned}$$

Thus, we obtain

$$|\langle N\hat{k}_\rho, \hat{k}_\rho \rangle| \leq \frac{\|N\|^{\frac{1}{2}}}{2} \left(\left| \langle U|N|^{\frac{1}{2}}\hat{k}_\rho, \hat{k}_\rho \rangle \right| + \frac{1}{2} \langle (|N|^r + |N^*|^{1-r})\hat{k}_\rho, \hat{k}_\rho \rangle \right).$$

Using the inequality above, by calculating the supremum over $\rho \in X$, we obtain

$$\text{ber}(N) \leq \frac{\|N\|^{\frac{1}{2}}}{2} \left(\text{ber}\left(U|N|^{\frac{1}{2}}\right) + \frac{1}{2} \| |N|^r + |N^*|^{1-r} \|_{\text{ber}} \right).$$

If we write $x = y = \hat{k}_\rho$ in the inequality (11), then we get

$$\begin{aligned}
|\langle N\hat{k}_\rho, \hat{k}_\rho \rangle| &\leq \frac{\|N\|^{1-r}}{2} \left(|\langle U|N|^r \hat{k}_\rho, \hat{k}_\rho \rangle| + \sqrt{\langle |N|^r \hat{k}_\rho, \hat{k}_\rho \rangle \langle |N^*|^r \hat{k}_\rho, \hat{k}_\rho \rangle} \right) \\
&\leq \frac{\|N\|^{1-r}}{2} \left(|\langle U|N|^r \hat{k}_\rho, \hat{k}_\rho \rangle| + \frac{\langle |N|^r \hat{k}_\rho, \hat{k}_\rho \rangle + \langle |N^*|^r \hat{k}_\rho, \hat{k}_\rho \rangle}{2} \right) \\
&\quad \text{(by arithmetic-geometric mean inequality)} \\
&\leq \frac{\|N\|^{1-r}}{2} \left(|\langle U|N|^r \hat{k}_\rho, \hat{k}_\rho \rangle| + \frac{1}{2} \langle (|N|^r + |N^*|^r) \hat{k}_\rho, \hat{k}_\rho \rangle \right).
\end{aligned} \tag{16}$$

Thus, we have

$$|\langle N\hat{k}_\rho, \hat{k}_\rho \rangle| \leq \frac{\|N\|^{1-r}}{2} \left(|\langle U|N|^r \hat{k}_\rho, \hat{k}_\rho \rangle| + \frac{1}{2} \langle (|N|^r + |N^*|^r) \hat{k}_\rho, \hat{k}_\rho \rangle \right).$$

Taking the supremum over $\rho \in X$ in the above inequality, we get

$$\text{ber}(N) \leq \frac{\|N\|^{1-r}}{2} \left(\text{ber}(U|N|^r) + \frac{1}{2} \| |N|^r + |N^*|^r \|_{\text{ber}} \right).$$

The proof has become completed.

Remark 3.2. If we put $r=1/2$ and apply Theorem 3.1 in Huban et al. (2022a), we get the following inequalities.

$$(i) \quad \text{ber}(N) \leq \frac{\|N\|^{\frac{1}{2}}}{2} \left(\text{ber}\left(|N|^{\frac{1}{2}}\right) + \frac{1}{2} \left\| |N|^{\frac{1}{2}} + |N^*|^{\frac{1}{2}} \right\|_{\text{ber}} \right)$$

$$(ii) \quad \text{ber}(N) \leq \frac{1}{2} \left(\text{ber}(|N|) + \frac{1}{2} \| |N| + |N^*| \|_{\text{ber}} \right).$$

Now, we will introduce theorem related to three operator.

Theorem 3.3. Assume $\mathcal{H} = \mathcal{H}(X)$ to be an RKHS. Let $P, R, S \in \mathcal{L}(\mathcal{H})$ and S be positive. Then

$$\text{ber}(PSR) \leq \frac{\|S\|}{2} \left(\text{ber}(PR) + \frac{1}{2} \| |P^*|^2 + |R|^2 \|_{\text{ber}} \right) \tag{17}$$

and

$$\text{ber}(PSR) \leq \frac{\|S\|}{2} \left(\text{ber}(PR) + \frac{1}{2} (\|P\|_{\text{ber}} \|R\|_{\text{ber}} + \|PR\|_{\text{ber}}) \right). \tag{18}$$

Proof. From inequality (6), we have

$$\begin{aligned}
|\langle PSR\hat{k}_\rho, \hat{k}_\rho \rangle| &\leq \frac{\|S\|}{2} (|\langle PR\hat{k}_\rho, \hat{k}_\rho \rangle| + \|R\hat{k}_\rho\| \|P^*\hat{k}_\rho\|) \\
&\leq \frac{\|S\|}{2} \left(|\langle PR\hat{k}_\rho, \hat{k}_\rho \rangle| + \sqrt{\langle |P^*|^2 \hat{k}_\rho, \hat{k}_\rho \rangle \langle |R|^2 \hat{k}_\rho, \hat{k}_\rho \rangle} \right) \\
&\leq \frac{\|S\|}{2} \left(|\langle PR\hat{k}_\rho, \hat{k}_\rho \rangle| + \frac{\langle |P^*|^2 \hat{k}_\rho, \hat{k}_\rho \rangle + \langle |R|^2 \hat{k}_\rho, \hat{k}_\rho \rangle}{2} \right) \\
&\quad (\text{by arithmetic-geometric mean inequality}) \\
&\leq \frac{\|S\|}{2} \left(|\langle PR\hat{k}_\rho, \hat{k}_\rho \rangle| + \frac{1}{2} \langle (|P^*|^2 + |R|^2) \hat{k}_\rho, \hat{k}_\rho \rangle \right)
\end{aligned} \tag{19}$$

Here, we get

$$|\langle PSR\hat{k}_\rho, \hat{k}_\rho \rangle| \leq \frac{\|S\|}{2} \left(|\langle PR\hat{k}_\rho, \hat{k}_\rho \rangle| + \frac{1}{2} \langle (|P^*|^2 + |R|^2) \hat{k}_\rho, \hat{k}_\rho \rangle \right).$$

Using the inequality above, by calculating the supremum over $\rho \in X$, we obtained

$$\text{ber}(PSR) \leq \frac{\|S\|}{2} \left(\text{ber}(PR) + \frac{1}{2} \| |P^*|^2 + |R|^2 \|_{\text{ber}} \right).$$

Now, let's prove the second inequality. If $P = 0$ or $R = 0$, the result follows trivially. Therefore, assume that P and R are nonzero. In inequality (15), if we replace P and R by $\sqrt{\frac{\|R\|_{\text{ber}}}{\|P\|_{\text{ber}}}} P$ and $\sqrt{\frac{\|P\|_{\text{ber}}}{\|R\|_{\text{ber}}}} R$, respectively, we reach

$$\begin{aligned}
\text{ber}(PSR) &\leq \frac{\|S\|}{2} \left(\text{ber}(PR) + \frac{1}{2} \left\| \frac{\|R\|_{\text{ber}}}{\|P\|_{\text{ber}}} |P^*|^2 + \frac{\|P\|_{\text{ber}}}{\|R\|_{\text{ber}}} |R|^2 \right\|_{\text{ber}} \right) \\
&\leq \frac{\|S\|}{2} \left(\text{ber}(PR) + \frac{1}{2} (\|P\|_{\text{ber}} \|R\|_{\text{ber}} + \| |P^*|^2 R \|_{\text{ber}}) \right) \\
&\quad (\text{by Lemma 2.2}) \\
&\leq \frac{\|S\|}{2} \left(\text{ber}(PR) + \frac{1}{2} (\|P\|_{\text{ber}} \|R\|_{\text{ber}} + \|RP\|_{\text{ber}}) \right),
\end{aligned}$$

in where

$$\| |P^*|^2 R \|_{\text{ber}}^2 = \| (|P^*|^2 R)^* (|P^*|^2 R) \|_{\text{ber}} = \| |R| |P^*|^2 |R| \|_{\text{ber}} = \| |R| |P P^*| |R| \|_{\text{ber}} = \| P^* |R| \|_{\text{ber}}^2$$

and

$$\|P^*|R|\|_{ber}^2 = \|P^*|R||R|P\|_{ber} = \|P^*R^*RP\|_{ber} = \| |RP| \|_{ber}^2 = \|RP\|_{ber}^2.$$

This completes the proof.

Corollary 3.4. *Let $P, R, S \in \mathcal{L}(\mathcal{H})$ and S be positive. Then*

$$\begin{aligned} (i) \quad & ber(PSR) \leq \frac{\|S\|}{2} \left(ber(PR) + \frac{1}{2} \| |P^*|^2 + |R|^2 \|_{ber} \right) \\ (ii) \quad & ber(PSR) \leq \frac{\|S\|}{2} \left(ber(PR) + \frac{1}{2} (\|P\|_{ber}\|R\|_{ber} + \|PR\|_{ber}) \right). \end{aligned}$$

Corollary 3.5. *Let $P, R, S \in \mathcal{L}(\mathcal{H})$ and S be positive. Then*

$$ber(PSR) \leq \frac{\|S\|}{2} (ber(PR) + \|R\|_{ber}\|P^*\|_{ber}).$$

Proof Taking the supremum over $\rho \in X$ in inequality (19), we have desired inequality.

Corollary 3.6. *Let $P, R \in \mathcal{L}(\mathcal{H})$. Then*

$$ber(PR) \leq \frac{1}{2} \| |P^*|^2 + |R|^2 \|_{ber}. \quad (20)$$

Proof. When we write $S = 1$ in inequality (17), we have

$$ber(PR) \leq \frac{1}{2} \left(ber(PR) + \frac{1}{2} \| |P^*|^2 + |R|^2 \|_{ber} \right).$$

Here, we get

$$ber(PR) \leq \frac{1}{2} \| |P^*|^2 + |R|^2 \|_{ber}.$$

Remark 3.7. *If we replace P by P^* in inequality (20), we have*

$$ber(P^*R) \leq \frac{1}{2} \| |P|^2 + |R|^2 \|_{ber}.$$

(see, (Huban et al., 2022b)).

Corollary 3.8. Let $N, M \in \mathcal{L}(\mathcal{H})$. Then

$$\text{ber}(NM) \leq \frac{\|N\|^{\frac{1}{2}}}{2} \| |N^*| + |M|^2 \|_{\text{ber}}$$

and

$$\text{ber}(NM) \leq \frac{\|N\|^{\frac{1}{2}}}{2} \left(\frac{1}{2} \| |N^*| + |M|^2 \|_{\text{ber}} + \frac{1}{2} \|M\|_{\text{ber}} \|N\|_{\text{ber}}^{\frac{1}{2}} \right).$$

Proof. Let $N = U|N|$ be the polar decomposition of N . $P = U|N|^{\frac{1}{2}}$, $S = |N|^{\frac{1}{2}}$ and $R = M$, in Theorem 3.4, we have

$$\text{ber}(NM) \leq \frac{\|N\|^{\frac{1}{2}}}{2} \left(\text{ber} \left(U|N|^{\frac{1}{2}}M \right) + \frac{1}{2} \| |N^*| + |M|^2 \|_{\text{ber}} \right) \quad (21)$$

and

$$\begin{aligned} \text{ber}(NM) &\leq \frac{\|N\|^{\frac{1}{2}}}{2} \left(\text{ber} \left(U|N|^{\frac{1}{2}}M \right) + \frac{1}{2} \left(\|U|N|^{\frac{1}{2}}\|_{\text{ber}} \|M\|_{\text{ber}} + \|MU|N|^{\frac{1}{2}}\|_{\text{ber}} \right) \right) \\ &\leq \frac{\|N\|^{\frac{1}{2}}}{2} \left(\text{ber} \left(U|N|^{\frac{1}{2}}M \right) + \frac{1}{2} \left(\|N\|_{\text{ber}}^{\frac{1}{2}} + \|M\|_{\text{ber}} \right) \right) \end{aligned} \quad (22)$$

In addition, we get

$$\begin{aligned} \left| \langle U|N|^{\frac{1}{2}}M\hat{k}_\rho, \hat{k}_\rho \rangle \right| &= \left| \langle M\hat{k}_\rho, |N|^{\frac{1}{2}}U\hat{k}_\rho \rangle \right| \\ &\leq \|M\hat{k}_\rho\| \left\| |N|^{\frac{1}{2}}U\hat{k}_\rho \right\| \\ &\quad \text{(by Cauchy-Schwarz inequality)} \\ &\leq \sqrt{\langle |M|^2\hat{k}_\rho, \hat{k}_\rho \rangle \langle U|N|U^*\hat{k}_\rho, \hat{k}_\rho \rangle} \\ &\leq \frac{\langle |M|^2\hat{k}_\rho, \hat{k}_\rho \rangle + \langle U|N|U^*\hat{k}_\rho, \hat{k}_\rho \rangle}{2} \\ &\quad \text{(by arithmetic-geometric mean inequality)} \\ &\leq \frac{1}{2} \langle (|M|^2 + |N^*|)\hat{k}_\rho, \hat{k}_\rho \rangle. \end{aligned} \quad (23)$$

Here, we reach

$$\left| \langle U | N |^{\frac{1}{2}} M \hat{k}_\rho, \hat{k}_\rho \rangle \right| \leq \frac{1}{2} \langle (|M|^2 + |N^*|) \hat{k}_\rho, \hat{k}_\rho \rangle.$$

Using the inequality above, by calculating the supremum over $\rho \in X$, we get

$$\text{ber}(NM) \leq \frac{1}{2} \| |N^*| + |M|^2 \|_{\text{ber}}.$$

We now come together inequality (21) and inequality (22).to present the desired outcome.

Corollary 3.9. *Let $N, M \in \mathcal{L}(\mathcal{H})$. Then*

$$\text{ber} \left(U | N |^{\frac{1}{2}} M \right) \leq \|M\|_{\text{Ber}} \left\| | N |^{\frac{1}{2}} U \right\|_{\text{Ber}} \leq \|M\|_{\text{Ber}} \left\| | N |^{\frac{1}{2}} U \right\|_{\text{Ber}} \leq \|M\|_{\text{Ber}} \|N\|_{\text{Ber}}.$$

Proof. From inequality (23), we have

$$\left| \langle U | N |^{\frac{1}{2}} M \hat{k}_\rho, \hat{k}_\rho \rangle \right| \leq \|M \hat{k}_\rho\| \left\| | N |^{\frac{1}{2}} U \hat{k}_\rho \right\|.$$

Given the inequality above, by calculating the supremum over $\rho \in X$, we get

$$\begin{aligned} \text{ber} \left(U | N |^{\frac{1}{2}} M \right) &\leq \|M\|_{\text{Ber}} \left\| | N |^{\frac{1}{2}} U \right\|_{\text{Ber}} \\ &\quad (\text{by } |U|^2 \leq 1) \\ &\leq \|M\|_{\text{Ber}} \left\| | N |^{\frac{1}{2}} U \right\|_{\text{Ber}} \\ &\leq \|M\|_{\text{Ber}} \|N\|_{\text{Ber}}. \end{aligned}$$

Corollary 3.10. *Let $N \in \mathcal{L}(\mathcal{H})$ with the polar decomposition $N = U | N |$. Then for any $0 \leq r \leq 1$,*

$$\text{ber}(N) \leq \frac{\|N\|^{1-r}}{2} (\text{ber}(U | N |^r) + \|N\|_{\text{ber}}^r). \quad (24)$$

Proof. Let $\hat{k}_\rho$ be a normalized reproducing kernel. If we write $x = y = \hat{k}_\rho$ in the inequality (11), then we have

$$|\langle N \hat{k}_\rho, \hat{k}_\rho \rangle| \leq \frac{\|N\|^{1-r}}{2} \left(|\langle U | N |^r \hat{k}_\rho, \hat{k}_\rho \rangle| + \sqrt{\langle |N|^r \hat{k}_\rho, \hat{k}_\rho \rangle \langle |N^*|^r \hat{k}_\rho, \hat{k}_\rho \rangle} \right).$$

In the inequality above, by computing the supremum over $\rho \in X$, we have

$$\begin{aligned} \text{ber}(N) &\leq \frac{\|N\|^{1-r}}{2} \left(\text{ber}(U|N|^r) + \sqrt{\| |N|^r \|_{\text{ber}} \| |N^*|^r \|_{\text{ber}}} \right) \\ &\leq \frac{\|N\|^{1-r}}{2} \left(\text{ber}(U|N|^r) + \sqrt{\|N\|_{\text{ber}}^r \|N\|_{\text{ber}}^r} \right) \\ &\quad (\text{by } \| |N| \|_{\text{ber}} = \| |N^*| \|_{\text{ber}} = \|N\|_{\text{ber}}) \\ &\leq \frac{\|N\|^{1-r}}{2} (\text{ber}(U|N|^r) + \|N\|_{\text{ber}}^r). \end{aligned}$$

Thus, we obtain

$$\text{ber}(N) \leq \frac{\|N\|^{1-r}}{2} (\text{ber}(U|N|^r) + \|N\|_{\text{ber}}^r).$$

Remark 3.11. Let $N \in \mathcal{L}(\mathcal{H})$ with the polar decomposition $N = U|N|$. When we write $r = \frac{1}{2}$ in inequality (24), it has been demonstrated by Başaran and Gürdal (2023) with the following inequality

$$\text{ber}(N) \leq \frac{\|N\|^{\frac{1}{2}}}{2} \left(\text{ber}\left(U|N|^{\frac{1}{2}}\right) + \|N\|_{\text{ber}}^{\frac{1}{2}} \right). \quad (25)$$

Corollary 3.12. Let $N \in \mathcal{L}(\mathcal{H})$. Then for any $0 \leq r \leq 1$,

$$\text{ber}(N) \leq \frac{\|N\|^{1-r}}{2} \| |N|^r + |N^*|^r \|_{\text{ber}}. \quad (26)$$

In particular

$$\text{ber}(N) \leq \frac{\|N\|^{\frac{1}{2}}}{2} \left\| |N|^{\frac{1}{2}} + |N^*|^{\frac{1}{2}} \right\|_{\text{ber}}.$$

Proof. Suppose $\hat{k}_\rho$ be a reproducing kernel that has been normalized. In the inequality (9), if we choose $f(t) = g(t) = t^{\frac{r}{2}}$ and $x = y = \hat{k}_\rho$, then we reach

$$\begin{aligned} |\langle N\hat{k}_\rho, \hat{k}_\rho \rangle| &\leq \sqrt{\langle |N^*|^{\frac{r}{2}} |O|^{\frac{1}{2}} |N^*|^{\frac{r}{2}} \hat{k}_\rho, \hat{k}_\rho \rangle \langle |N|^{\frac{r}{2}} |O^*|^{\frac{1}{2}} |N|^{\frac{r}{2}} \hat{k}_\rho, \hat{k}_\rho \rangle} \\ &\leq \sqrt{\langle |N^*|^{\frac{r}{2}} |N^*|^{\frac{r}{2}} \hat{k}_\rho, \hat{k}_\rho \rangle \langle |N|^{\frac{r}{2}} |N|^{\frac{r}{2}} \hat{k}_\rho, \hat{k}_\rho \rangle} \end{aligned}$$

$$\leq \sqrt{\langle |N|^r \hat{k}_\rho, \hat{k}_\rho \rangle \langle |N^*|^r \hat{k}_\rho, \hat{k}_\rho \rangle}.$$

Here, we get

$$|\langle N \hat{k}_\rho, \hat{k}_\rho \rangle| \leq \sqrt{\langle |N|^r \hat{k}_\rho, \hat{k}_\rho \rangle \langle |N^*|^r \hat{k}_\rho, \hat{k}_\rho \rangle}$$

From the above inequality, if $N = U|N|$ is the polar decomposition of $N \in \mathcal{L}(\mathcal{H})$, then we get

$$\begin{aligned} |\langle N \hat{k}_\rho, \hat{k}_\rho \rangle| &= |\langle U|N| \hat{k}_\rho, \hat{k}_\rho \rangle| \\ &\leq \sqrt{\langle |N|^r \hat{k}_\rho, \hat{k}_\rho \rangle \langle |N^*|^r \hat{k}_\rho, \hat{k}_\rho \rangle} \\ &\leq \frac{\langle |N|^r \hat{k}_\rho, \hat{k}_\rho \rangle + \langle |N^*|^r \hat{k}_\rho, \hat{k}_\rho \rangle}{2} \\ &\quad \text{(by arithmetic-geometric mean inequality)} \\ &\leq \frac{1}{2} \langle (|N|^r + |N^*|^r) \hat{k}_\rho, \hat{k}_\rho \rangle. \end{aligned}$$

Here, we obtain

$$|\langle U|N| \hat{k}_\rho, \hat{k}_\rho \rangle| \leq \frac{1}{2} \langle (|N|^r + |N^*|^r) \hat{k}_\rho, \hat{k}_\rho \rangle.$$

Given the inequality shown above, by computing the supremum $\rho \in X$, We obtain

$$\text{ber}(N) \leq \frac{1}{2} \| |N|^r + |N^*|^r \|_{\text{ber}}.$$

Given that we combine this with inequality (15), we find the desired result. Since $|N|^{\frac{1}{2}} \leq \|N\|_{\text{ber}}^{\frac{1}{2}} I$, we get $|N|^{\frac{1}{2}} \leq \|N\|^{\frac{1}{2}} |N|^{\frac{1}{2}}$. Similarly, $|N^*|^{\frac{1}{2}} \leq \|N\|^{\frac{1}{2}} |N^*|^{\frac{1}{2}}$. Hence,

$$\langle |N| \hat{k}_\rho, \hat{k}_\rho \rangle \leq \|N\|^{\frac{1}{2}} \langle |N|^{\frac{1}{2}} \hat{k}_\rho, \hat{k}_\rho \rangle \text{ and } \langle |N^*| \hat{k}_\rho, \hat{k}_\rho \rangle \leq \|N\|^{\frac{1}{2}} \langle |N^*|^{\frac{1}{2}} \hat{k}_\rho, \hat{k}_\rho \rangle.$$

Here, we have

$$\begin{aligned} |\langle N \hat{k}_\rho, \hat{k}_\rho \rangle| &\leq \sqrt{\langle |N| \hat{k}_\rho, \hat{k}_\rho \rangle \langle |N^*| \hat{k}_\rho, \hat{k}_\rho \rangle} \\ &\quad \text{(by mixed Cauchy-Schwarz inequality of Kato (1952))} \\ &\leq \sqrt{\langle |N|^{\frac{1}{2}} \hat{k}_\rho, \hat{k}_\rho \rangle \langle |N^*|^{\frac{1}{2}} \hat{k}_\rho, \hat{k}_\rho \rangle} \end{aligned}$$

$$\begin{aligned}
&\leq \|N\|^{\frac{1}{2}} \sqrt{\langle |N|^{\frac{1}{2}} \hat{k}_\rho, \hat{k}_\rho \rangle \langle |N^*|^{\frac{1}{2}} \hat{k}_\rho, \hat{k}_\rho \rangle} \\
&\leq \|N\|^{\frac{1}{2}} \frac{\langle |N|^{\frac{1}{2}} \hat{k}_\rho, \hat{k}_\rho \rangle + \langle |N^*|^{\frac{1}{2}} \hat{k}_\rho, \hat{k}_\rho \rangle}{2} \\
&\leq \|N\|^{\frac{1}{2}} \frac{1}{2} \langle (|N|^{\frac{1}{2}} + |N^*|^{\frac{1}{2}}) \hat{k}_\rho, \hat{k}_\rho \rangle.
\end{aligned}$$

Thus, we get

$$|\langle N \hat{k}_\rho, \hat{k}_\rho \rangle| \leq \|N\|^{\frac{1}{2}} \frac{1}{2} \langle (|N|^{\frac{1}{2}} + |N^*|^{\frac{1}{2}}) \hat{k}_\rho, \hat{k}_\rho \rangle.$$

Using the inequality shown above, by computing the supremum $\rho \in X$, We are arriving at

$$\text{ber}(N) \leq \frac{\|N\|^{\frac{1}{2}}}{2} \left\| |N|^{\frac{1}{2}} + |N^*|^{\frac{1}{2}} \right\|_{\text{ber}}.$$

Remark 3.13. Let $N \in \mathcal{L}(\mathcal{H})$ with the polar decomposition $N = U|N|$. When we write $r = \frac{1}{2}$ in inequality (24), it has been demonstrated by Başaran and Gürdal (2023) with the following inequality

$$\text{ber}(N) \leq \frac{\|N\|^{\frac{1}{2}}}{2} \left(\left\| |N|^{\frac{1}{2}} + |N^*|^{\frac{1}{2}} \right\|_{\text{ber}} \right).$$

The next statement has been given by Theorem 3.1 in the case $T = \frac{1}{2}$.

Corollary 3.14. Let $N \in \mathcal{L}(\mathcal{H})$ with the polar decomposition $N = U|N|$. Then

$$\text{ber}(N) \leq \frac{\|N\|^{\frac{1}{2}}}{2} \left(\text{ber} \left(U|N|^{\frac{1}{2}} \right) + \left\| |N|^{\frac{1}{2}} + |N^*|^{\frac{1}{2}} \right\|_{\text{ber}} \right). \quad (27)$$

The inequality (25) is refined by the inequality (27). Because $\left\| |N|^{\frac{1}{2}} + |N^*|^{\frac{1}{2}} \right\|_{\text{ber}} \leq 2\|N\|^{\frac{1}{2}}_{\text{ber}}$.

Theorem 3.15. Assume $\mathcal{H} = \mathcal{H}(X)$ to be an RKHS. Suppose that $N \in \mathcal{L}(\mathcal{H})$. If f, g are nonnegative continuous function on $[0, \infty)$ obtaining $f(t)g(t) = t$ ($t > 0$), then

$$ber(N) \leq \frac{1}{2} \left\| g(|N^*|)|O|^{\frac{1}{2}}g(|N^*|) + f(|N|)|O^*|^{\frac{1}{2}}f(|N|) \right\|_{ber}$$

for some contraction O .

Proof From the inequality (9), we get

$$\begin{aligned} |\langle N\hat{k}_\rho, \hat{k}_\rho \rangle| &\leq \sqrt{\langle g(|N^*|)|O|^{\frac{1}{2}}g(|N^*|)\hat{k}_\rho, \hat{k}_\rho \rangle \langle f(|N|)|O^*|^{\frac{1}{2}}f(|N|)\hat{k}_\rho, \hat{k}_\rho \rangle} \\ &\leq \frac{\langle g(|N^*|)|O|^{\frac{1}{2}}g(|N^*|)\hat{k}_\rho, \hat{k}_\rho \rangle + \langle f(|N|)|O^*|^{\frac{1}{2}}f(|N|)\hat{k}_\rho, \hat{k}_\rho \rangle}{2} \\ &\quad \text{(by arithmetic-geometric mean inequality)} \\ &\leq \frac{1}{2} \langle [g(|N^*|)|O|^{\frac{1}{2}}g(|N^*|) + f(|N|)|O^*|^{\frac{1}{2}}f(|N|)] \hat{k}_\rho, \hat{k}_\rho \rangle. \end{aligned}$$

Thus, we have

$$|\langle N\hat{k}_\rho, \hat{k}_\rho \rangle| \leq \frac{1}{2} \langle [g(|N^*|)|O|^{\frac{1}{2}}g(|N^*|) + f(|N|)|O^*|^{\frac{1}{2}}f(|N|)] \hat{k}_\rho, \hat{k}_\rho \rangle.$$

Taking the supremum over $\rho \in X$ in the above inequality, we have

$$ber(N) \leq \frac{1}{2} \left\| g(|N^*|)|O|^{\frac{1}{2}}g(|N^*|) + f(|N|)|O^*|^{\frac{1}{2}}f(|N|) \right\|_{ber}.$$

Recall that $|O|^{\frac{1}{2}} \leq \left\| |O|^{\frac{1}{2}} \right\|_{ber} \leq I$. Hence

$$g(|N^*|)|O|^{\frac{1}{2}}g(|N^*|) \leq g^2(|N^*|). \quad (28)$$

Similarly, we get

$$f(|N|)|O^*|^{\frac{1}{2}}f(|N|) \leq f^2(|N|). \quad (29)$$

There, we takes utilization of the fact that O^* becomes a contraction also if $|N|$ becomes contraction operator. The combination of inequality (28) with inequality (29), The next are the results we obtain.

Corollary 3.16. Let $N \in \mathcal{L}(\mathcal{H})$. If f, g are nonnegative continuous function on $[0, \infty)$ satisfying $f(t)g(t) = t$ ($t > 0$), then

$$\begin{aligned} \text{ber}(N) &\leq \frac{1}{2} \left\| g(|N^*|) |O|^{\frac{1}{2}} g(|N^*|) + f(|N|) |O^*|^{\frac{1}{2}} f(|N|) \right\|_{\text{ber}} \\ &\leq \frac{1}{2} \|g^2(|N^*|) + f^2(|N|)\|_{\text{ber}}. \end{aligned} \quad (30)$$

for some contraction O .

Remark 3.17. The inequality (30) provides a improvement of (Huban et al., 2022a).

Remark 3.18. If, $N, M \in \mathcal{L}(\mathcal{H})$, then

$$\begin{aligned} \text{ber}(N) &\leq \frac{1}{2} \left\| |N| |O|^{\frac{1}{2}} |N| + |M| |O^*|^{\frac{1}{2}} |M| \right\|_{\text{ber}} \\ &\leq \frac{1}{2} \| |M|^2 + |N|^2 \|_{\text{ber}} \end{aligned}$$

for some contraction O . The above inequality from Lemma 2.7 and positivity $\begin{bmatrix} |N|^2 & N^*M \\ M^*N & |M|^2 \end{bmatrix}$. Recall that the above result refines (Huban et al, 2022b).

Now, we prove the following theorem by using The Lemma 2.8.

Theorem 3.19. Assume $\mathcal{H} = \mathcal{H}(X)$ to be an RKHS. Suppose that $N \in \mathcal{L}(\mathcal{H})$. If f, g are nonnegative continuous function on $[0, \infty)$ obtaining $f(t)g(t) = t$ ($t > 0$), then

$$\text{ber}(N) \leq \frac{1}{2} \sqrt{\|f^2(|N|) + f^2(|N^*|)\|_{\text{ber}} \|g^2(|N|) + g^2(|N^*|)\|_{\text{ber}}} \quad (31)$$

and

$$\text{ber}(N) \leq \frac{1}{4} \|f^2(|N|) + f^2(|N^*|) + g^2(|N|) + g^2(|N^*|)\|_{\text{ber}}.$$

Proof. Depending on Lemma 8(i), we have

$$|\langle \Re N x, y \rangle| \leq \frac{1}{2} \sqrt{\langle (f^2(|N|) + f^2(|N^*|))x, x \rangle \langle (g^2(|N|) + g^2(|N^*|))y, y \rangle}$$

Putting $x = y = \hat{k}_\rho$ and $N = e^{i\theta}N$ in the above inequality, we reach

$$|\langle \Re(e^{i\theta}N)\hat{k}_\rho, \hat{k}_\rho \rangle| \leq \frac{1}{2} \sqrt{\langle (f^2(|N|) + f^2(|N^*|))\hat{k}_\rho, \hat{k}_\rho \rangle \langle (g^2(|N|) + g^2(|N^*|))\hat{k}_\rho, \hat{k}_\rho \rangle}.$$

In the inequality above, by computing the supremum over $\rho \in X$, we have

$$\|\Re(e^{i\theta}N)\| \leq \frac{1}{2} \sqrt{\|f^2(|N|) + f^2(|N^*|)\|_{ber} \|g^2(|N|) + g^2(|N^*|)\|_{ber}}.$$

If we take supremum over $\theta \in \mathbb{R}$ and use the fact that $ber(N) = \sup_{\theta \in \mathbb{R}} \|\Re(e^{i\theta}N)\|$, we reach

$$ber(N) \leq \frac{1}{2} \sqrt{\|f^2(|N|) + f^2(|N^*|)\|_{ber} \|g^2(|N|) + g^2(|N^*|)\|_{ber}}$$

as desired.

Let's prove the second inequality. If we use utilize the arithmetic-geometric mean inequality for norm inequality (31), we have

$$ber(N) \leq \frac{1}{4} \|f^2(|N|) + f^2(|N^*|) + g^2(|N|) + g^2(|N^*|)\|_{ber}.$$

This completes the proof.

Remark 3.20. Among the celebrated upper bounds for $ber(N)$, we get Huban et al. (2022a)

$$ber(N) \leq \frac{1}{2} \| |N| + |N^*| \|_{ber}. \quad (32)$$

In Theorem 3.19, if we $f(t) = g(t) = t^{\frac{1}{2}}$, then the inequality (31) follows as a particular case of Theorem 3.19. Thus, this last theorem presents a generalizations of (31).

Başaran (2024), Başaran et al. (2019), Başaran and Gürdal (2023), Başaran et al. (2022), Gürdal et al. (2025), Gürdal and Başaran (2023), Gürdal et al. (2024), Garayev et al. (2021), and Huban et al. (2021) contain the fundamental characteristics and background knowledge on these creative concepts.

CONFLICT OF INTEREST

The author declares no conflict of interest.

ACKNOWLEDGMENT

I would like to thank the referees who contributed to the preparation of this article.

REFERENCES

- Abu-Omar, A., & Kittaneh F. (2015). Upper and lower bounds for the numerical radius with an application to involution operators. *The Rocky Mountain Journal of Mathematics*, 45(4), 1055-065. <https://doi.org/10.1216/RMJ-2015-45-4-1055>
- Aronzajn, N. (1950). Theory of reproducing kernels. *Transactions of the American Mathematical Society*, 68, 337-404.
- Başaran H. (2024). Berezin radius inequalities related to the contraction operator. *Montes Taurus Journal of Pure and Applied Mathematics*, 6 (1), 102-109. Article ID:MTJPAM-D-23-00044.
- Başaran, H., Gürdal, M., & Güncan, A. N. (2019). Some operator inequalities associated with Kantorovich and Hölder-McCarthy inequalities and their applications. *Turkish Journal of Mathematics*, 43(1) (2019), 523-532. <http://doi.org/10.3906/mat-1811-10>
- Başaran, H., & Gürdal, V. (2023). Berezin radius and Cauchy-Schwarz inequality. *Montes Taurus Journal of Pure and Applied Mathematics*, 5(3), 16-22. Article ID: MTJPAM-D-21-00068
- Başaran H., Huban M. B., Gürdal M. (2022). Inequalities related to Berezin norm and Berezin number of operators. *Bulletin of Mathematical Analysis and Applications*, 14(2), 1-11. <https://doi.org/10.54671/bmaa-2022-2-1>
- Berezin, F.A. (1972). Covariant and contravariant symbols for operators. *Mathematics of the USSR-Izvestiya*, 6(5), 1117-1151. <http://dx.doi.org/10.1070/IM1972v006n05ABEH001913>
- Bhunia, P., Dragomir, S. S., Moslehian, M. S., & Paul, K. (2022). *Lectures on Numerical Radius Inequalities*. Springer Cham.
- Bhatia, R. (2007). *Positive definite matrices*. Princeton, NJ: Princeton University Press.
- Bhunia, P., Jana, S., & Paul, K. (2023a). Numerical radius inequalities and estimation of zeros of polynomials. *Georgian Mathematical Journal*, 30(5), 671-682. <https://doi.org/10.1515/gmj-2023-2037>
- Bhunia, P., Gürdal, M., Paul, K., Sen, A., & Tapdigoglu, R. (2023b). On a new norm on the space of reproducing kernel Hilbert space operators and Berezin radius inequalities. *Numerical Functional Analysis and Optimization*, 44(9):970-986. <https://doi.org/10.1080/01630563.2023.2221857>
- Buzano, M. L. (1974). *Generalizzazione della disuguaglianza di Cauchy-Schwarz*. Rendiconti del Seminario Matematico. Università e Politecnico Torino, 31 (1971/73), 405-409.
- Dragomir, S. S. (2016). Buzano's inequality holds for any projection. *Bulletin of the Australian Mathematical Society*, 3, 504-510. <https://doi.org/10.1017/S0004972715001525>
- Engliš, M. (1995). Toeplitz operators and the Berezin transform on H^2 . *Linear Algebra and Applications*, 223-224, 171-204. [https://doi.org/10.1016/0024-3795\(94\)00056-J](https://doi.org/10.1016/0024-3795(94)00056-J)
- Furuta, T. (2001). *Invitation to linear operators*, London: Taylor and Francis.

- Garayev, M. T., Guedri, H., Gürdal M., & Alsahli, G. M. (2021). On some problems for operators on the reproducing kernel Hilbert space. *Linear and Multilinear Algebra*, 69(11), 2059-2077. <https://doi.org/10.1080/03081087.2019.1659220>
- Gürdal, M., Alomari, M.W., Başaran, H. (2025). Berezin radius inequalities via Orlicz function. *Palestine Journal of Mathematics*, 14(1), 928-945.
- Gürdal, M., Erkan, G. G., & Garayev, M. (2024). Berezin norm and Berezin radius inequalities of product and sums with Selberg inequality. *Proceedings of the Institute of Mathematics and Mechanics*, 50(2), 258-273. <https://doi.org/10.30546/2409-4994.2024.50.2.258>
- Gürdal, V., & Başaran H. (2023). On Berezin radius inequalities via Cauchy-Schwarz type inequalities. *Malaya Journal of Matematik*, 11(02) (2023), 127-141. <https://doi.org/10.26637/mjm1102/002>
- Huban, M.B., Başaran, H. & Gürdal, M. (2021). New upper bounds related to the Berezin number inequalities. *Journal of Inequalities and Special Functions*, 12(3), 1-12.
- Huban, M.B., Başaran, H. & Gürdal, M. (2022a). Some new inequalities via Berezin numbers. *Turkish Journal of Mathematics and Computer Science*, 14(1), 129-137. <https://doi.org/10.47000/tjmcs.1014841>
- Huban, M.B., Başaran, H. & Gürdal, M. (2022b). Berezin number inequalities via convex functions. *Filomat*, 36(7), 2333-2344. <https://doi.org/10.2298/FIL2207333H>
- Jorgensen, P. (2006). *Analysis and probability: wavelets, signals, fractals*, Springer.
- Karaev, M.T. (2006). Berezin symbol and invertibility of operators on the functional Hilbert spaces. *Journal of Functional Analysis*, 238, 181-192. <https://doi.org/10.1016/j.jfa.2006.04.030>
- Karaev, M.T. (2013). Reproducing kernels and Berezin symbols techniques in various questions of operator theory. *Complex Analysis and Operator Theory*, 7, 983-1018. <https://doi.org/10.1007/s11785-012-0232-z>
- Karaev, M.T., Gürdal, M., & Huban, M.B. (2016). Reproducing kernels, Englis algebras and some applications. *Studia Mathematica*, 232(2), 113-141. <https://doi.org/10.4064/sm8385-4-2016>
- Karaev, M., Gürdal, M., & Saltan, S. (2011). Some applications of Banach algebra techniques. *Mathematische Nachrichten*, 284(13):1678-1689. <https://doi.org/10.1002/mana.200910129>
- Kato, T. (1952). Notes on some inequalities for linear operators. *Mathematische Annalen*, 125, 208-212. <https://doi.org/10.1007/BF01343117>
- Kittaneh, F. (1988). Notes on some inequalities for Hilbert space operators. *Publications of the Research Institute for Mathematical Sciences*, 24, 283-293. <https://doi.org/10.2977/PRIMS/1195175202>
- Kittaneh, F. (1997). Norm inequalities for certain operators sums. *Journal of Functional Analysis*, 143, 337-348. <https://doi.org/10.1006/jfan.1996.2957>
- Kittaneh, F., Moslehian, M. S., & Yamakazi, T. (2015). Carthesian decomposition and numerical radius inequalities. *Linear Algebra and Applications*, 471, 46-53. <https://doi.org/10.1016/j.laa.2014.12.016>
- Pečarić, J., Furuta, T., Hot, M.H., & Seo, Y. (2005). *Mond-Pečarić method in operator inequalities*. Zagreb; Elements.

- Sababheh, M., Conde, C., & Moradi, H. R. (2023). On the Cauchy-Schwarz inequality. *Operators and Matrices*, 17(2) (2023), 525-538. <https://doi.org/10.7153/oam-2023-17-34>
- Sababheh, M., & Moradi, H. R. (2023). Numerical radius of kronecker product of matrices. *Journal of Applied Analysis & Computation*, 16(3), 2943-2954. <https://www.jaac-online.com/article/doi/10.11948/20230064>
- Sababheh, M., Moradi, H. R., & Sahoo, S. (2024). Inner product inequalities with application. *Linear and Multilinear Algebra*, (2024), 1-14. <https://doi.org/10.1080/03081087.2024.2312444>
- Stojiljkovic, V., & Gürdal, M. (2025). Estimates of the numerical radius utilizing various function properties. *Turkish Journal of Mathematics*, 49(1), 49-64. <https://doi.org/10.55730/1300-0098.3573>
- Yamakazi, T. (2007). On upper and lower bounds of the numerical radius and an equality condition. *Studia Mathematica*, 178, 83-89. <https://doi.org/10.4064/sm178-1-5>
- Yamancı, U., Tunç, R., & Gürdal, M. (2020). Berezin number, Grüss-type inequalities and their applications. *Bulletin of the Malaysian Mathematical Sciences Society*, 43(3), 2287-2296. <https://doi.org/10.1007/s40840-019-00804-x>
- Yang, C., & Xu, M. (2023). Some new numerical radius and Hilbert-Schmidt numerical radius inequalities for Hilbert space operators. *Journal of Mathematical Inequalities*, 17(1), 269-282. <https://doi.org/10.7153/jmi-2023-17-19>
- Zhu, K., (1990). *Operator Theory in Function Spaces*, Marcel Dekker Inc., New York, USA Series: Pure and Applied Mathematicsa Series of Monographs and Textbooks, Volume 139.



Gazi University

Journal of Science

PART A: ENGINEERING AND INNOVATION

<http://dergipark.org.tr/guj.1647353>

Strength and Microstructural Analysis of Nano Powder and Basalt Fibre Reinforced RCC Pavement Concrete

Şeyma SÜNBÜL^{1*} Ahmet TORTUM² ¹ Erzincan Binali Yıldırım University, Erzincan, Türkiye² Atatürk University, Erzurum, Türkiye

Keywords	Abstract
Roller Compacted Concrete SEM and BET Analyses Nanopowders Basalt Fibre	Nowadays, durability and longevity represent the main issues in the field of road concrete. Roller Compacted Concrete (RCC) is a sustainable option for road pavements due to its low binder content, fast application and high strength. Nevertheless, there is a necessity for the improvement of RCC road concretes through the incorporation of various additives to enhance their performance. The present study investigates the mechanical and microstructural properties of RCC road concretes containing Fe ₂ O ₃ and TiO ₂ nano powders and basalt fiber additives. Three distinct mixtures were formulated, and a series of compressive strength, flexural strength and abrasion resistance tests were conducted. The results were verified by BET (Brunauer-Emmett-Teller) and SEM (Scanning Electron Microscope) analyses. The experimental findings demonstrated that the incorporation of nano powder and basalt fiber additives was effective in increasing the durability of concrete. The third mixture demonstrated optimal performance, exhibiting a compressive strength of 61.20 MPa, a flexural strength of 8.80 MPa, and a mass loss of 2.5 g. BET analysis revealed that the incorporation of nanopowders enhanced the matrix density by increasing the surface area. Furthermore, the SEM images demonstrated that the nanopowders were uniformly dispersed within the matrix, and the basalt fibres exhibited robust interfacial interactions. In conclusion, the combined use of Fe ₂ O ₃ and TiO ₂ nanopowders with basalt fibre additive is considered as an effective approach to improve the performance of RCC road concretes.

Cite

Sünbül, Ş. & Tortum, A. (2025). Strength and Microstructural Analysis of Nano Powder and Basalt Fibre Reinforced RCC Pavement Concrete. *GU J Sci, Part A, 12(2)*, 583-607. doi:[10.54287/guj.1647353](https://doi.org/10.54287/guj.1647353)

Author ID (ORCID Number)	Article Process
0000-0003-2571-1709 Şeyma SÜNBÜL	Submission Date 26.02.2025
0000-0002-5770-766X Ahmet TORTUM	Revision Date 24.03.2025
	Accepted Date 28.05.2025
	Published Date 30.06.2025

1. INTRODUCTION

The incorporation of nano powder and fibre additives into RCC road concrete has emerged as a pivotal subject in contemporary construction research, particularly regarding enhancing strength and durability. RCC has garnered significant attention as a sustainable alternative due to its reduced binder content and energy requirements in comparison to conventional concrete. The preference for RCC is particularly pronounced in the domain of road construction, owing to its cost-effectiveness, rapid construction time, and environmental benefits. The water-cement ratio employed in the production of RCC differs from that of traditional concrete, resulting in a drier concrete with zero slump rate. This concrete can be opened to light traffic without any damage. Additionally, RCC coatings offer minimize construction time and labour costs. The present study aims to investigate the effects of nano powder and basalt fiber additives used in RCC road concrete on the mechanical properties of concrete. The incorporation of nano powder and fibre into RCC road concretes

*Corresponding Author, e-mail: seymaercansunbul@gmail.com

represents a significant development in contemporary construction industry, particularly regarding the enhancement of strength and durability. RCC technology is distinguished by its ability to withstand substantial traffic loads and the deterioration that results from environmental effects. The material's distinctiveness is evidenced by its enhanced sustainability, as evidenced by its production with reduced binder content and energy consumption when compared to conventional road concretes. The utilizations of RCC, particularly in the context of road construction, is experiencing a notable increase in prevalence, primarily due to its distinct advantages, namely its capacity to expedite construction processes while maintaining a low cost (Choi & Groom, 2001). Rigid roller compactable pavements are essentially like conventional concrete pavements, but the cement content used is less than that used in conventional concrete. The utilizations of a reduced cement content in the concrete mixture has been demonstrated to engender a concomitant increase in the economic efficiency of the material. The addition of various additives to the concrete can further enhance its strength. These types of coatings are characterized by a higher level of dryness in comparison to traditional coatings, exhibiting minimal propensity for collapse (Mardani et al, 2020). This issue, which impacts workability, is addressed by utilizing high compression energy. The transportation, placement and compaction of RCC is analogous to that of bituminous pavements. Consequently, RCC can be applied expeditiously and economically. RCC coatings, which were previously preferred in industrial field floors where heavy loads were carried and travelled at low speeds due to their high surface roughness, have also started to be used in urban roads and intercity highways in recent years. The production and construction of roller-compacted concrete is a more practical approach when compared with traditional concrete methods and can be utilised following the compaction process (Menekşe, 2020). The material is laid and compacted in layers. It is imperative that adequate and effective compaction is provided in each layer. The term used for this process is 'vibratory rolling', which involves the utilization of vibratory rollers for the compaction of the concrete (Öztürk, 2018). The production of roller-compacted concrete roads involves the integration of several characteristics found in both conventional concrete roads and asphalt roads. These common features can be classified in accordance with the following criteria: the utilizations of materials analogous to those employed in the construction of traditional concrete roads; discrepancies in mixing ratios; and variations in curing requirements. The similarities with asphalt roads are that the equipment types such as aggregate grading, placement, compaction, etc. are the same. Furthermore, RCC pavements are typically utilised in contexts characterized by low-speed heavy traffic, and in areas where factors such as strength, durability and cost-effectiveness are of paramount importance, such as in the case of airport runways and taxiways (Huang, 2004; Mallick & Tahar, (2018). Roller compactable concrete is a superstructure construction method for rigid roads that is characterized by its zero-slump concrete properties, which are derived from the process of consolidation. This new type of pavement is distinguished from traditional concrete pavements by its unique construction techniques, which draw parallels with those employed in the fabrication of flexible pavements. The materials utilized in the construction of RCC are analogous to those employed in the construction of traditional roads, namely water, binder, aggregate, and so forth. Nevertheless, the water-cement ratio utilized in the construction of RCC pavements deviates from that employed in conventional road design. Due to its drier composition,

RCC exhibits no tendency to slump. The water/cement and air void ratio of RCC pavements is lower than that of conventional concrete pavements. This is due to the compaction method. Furthermore, it has been demonstrated that high load-bearing capacity can be attained immediately following the compaction process. Consequently, the RCC pavement can be opened to light traffic without incurring any damage after the compaction process. Moreover, the use of RCC pavements eliminates the necessity for formwork, reinforcement, dowels or joints present in conventional concrete pavement, consequently leading to a reduction in labour costs and construction time for pavement construction of similar length (ACI Committee, 2014). Furthermore, given that the cementitious content of RCC is lower than that of conventional concrete, it is more environmentally friendly (LaHucik et al, 2017). As the application of RCC coatings is analogous to that of asphalt coatings, the construction time and labour costs are comparable. However, it has been demonstrated that RCC pavements exhibit superior strength and durability in comparison to asphalt pavements. Furthermore, RCC pavements exhibit a high resistance to rutting, and do not typically exhibit significant cracking under the turning or braking forces of vehicles. It is notable for its stability at elevated temperatures, a property that renders it resistant to degradation by several chemicals, including fuels (Prusinski, 1997). A thorough examination of these situations reveals that, despite certain disadvantages associated with RCC, it offers numerous benefits for rigid pavements. Moreover, it is more environmentally friendly than other coating materials and is regarded as one of the ideal building materials for road construction. RCC pavements are produced by utilizing the same materials in the traditional concrete pavement mixture, albeit in different proportions. They are produced by using coarse and fine aggregates, binding materials (cement, fly ash, slag, etc.), water and chemical additives when necessary. In the present study, the focus is on the utilizations of nano powder additives in the context of RCC design, with particular emphasis on their impact on the microstructure and the effectiveness of crack prevention in fibres. The incorporation of nano-additives into concrete has been demonstrated to enhance the mechanical and durability properties of the material. This is achieved by creating a dense matrix through the filling of micropores that form during the process of cement hydration. A significant amount of research is currently being conducted on the various applications of nanomaterials in the field of building materials, with a particular focus on cement-based materials (Kawashima et al., 2013). Nanotechnology is an area of research that is currently undergoing rapid development and expansion. It has been identified as one of the most active research domains in contemporary academia, with a wide range of potential applications across a broad spectrum of disciplines. These include, but are not limited to, the following areas: energy, defence, information and technology, agriculture, environmental protection and health (Sanchez & Sobolev, 2010). The applications of nanoscience and nanotechnology in cement and concrete systems can be categorised under the following four headings. These are as follows:

- Research and modification of the nanostructure of calcium silicate-hydrate (C-S-H) gel,
- Controlled release of chemical additives into the concrete mixture,
- Cement-based composite materials containing carbon nanotubes,

- Cement systems containing nanoparticles (Raki, 2010).

The utilizations of nanoparticles in the field of building materials is a subject that is garnering increasing attention daily. The common application areas of nanotechnology in the construction sector can be classified as follows: new production techniques, devices and controls; nanostructured arranged materials; high-performance structural materials; smart structures and use of micro/nano sensors; special coatings, paints and thin films; multifunctional materials and components; energy-saving lights; fuel cells; and communication and computer devices (Zhu et al, 2010). Advancements in the domain of nanotechnology have facilitated enhanced resilience of concrete, nanomaterial, to environmental stress and chemical agents (e.g., deicing salts) (Öcal et al, 2018). This progress has led to the emergence of self-healing smart concretes, which exhibit remarkable resilience to external damage. The following three points summarise the problems related to the applications of nanotechnology in road engineering: The determination of current needs that cannot be effectively addressed using existing technology, the identification of applicable nanotechnology solutions in the field of road engineering, and the determination of nanotechnology solutions that will achieve the highest benefit/cost ratios by combining two concepts (Steyn, 2011). In recent years, a significant body of research has been dedicated to investigating the diverse applications of nanotechnology in the domain of building materials, with a particular focus on those containing nano-SiO₂, nano-Al₂O₃, nano-TiO₂, nano-Fe₂O₃, nano-CaCO₃, nanoparticles, carbon nanotubes, and nano-clay, among others. The behaviour of these materials is primarily influenced by the chemical reactions that occur at the interface. Consequently, the employment of nanoparticles in cement-based materials has been demonstrated to yield superior outcomes by modifying their behaviour in both fresh and hardened states (Sanchez & Sobolev, 2010; Senff et al, 2012). The fibres utilised in concrete serve the primary function of preventing the propagation of stress within the concrete matrix by means of a crack bridging effect, thereby enhancing the flexural strength and mitigating the fragility of the concrete. Among these fibres, basalt fibres are distinguished by their strength and environmental resistance properties. As demonstrated in the extant literature, basalt fibres, in addition to steel, glass and polypropylene fibres, have been shown to be effective in increasing the durability of concrete (Bolat, 2009). However, the environmental compatibility and thermal resistance properties of basalt fibres render this material more attractive. A plethora of studies have been conducted on the utilisation of fibres as road pavement. Bolat (2009) conducted a large-scale study in which he determined the performance of polyester (PYTB), and polypropylene fibre-reinforced concrete (PPTB) compared to steel fibre-reinforced (ÇTB) and non-fibre reinforced concrete. This was achieved by conducting fresh and hardened concrete tests and investigating the usability of the mixtures as highway superstructure. The study encompassed a comprehensive array of tests, extending beyond the conventional physical assessments of concretes. This encompassed plate splitting tensile strength, approximate compressive strength ascertained through surface hardness, beam flexural, compressive and elastic modulus, ultrasonic transmission speed, abrasion, capillary water absorption, and carbonation depth. As a result, when the wire-reinforced concretes were compared to the unreinforced concretes; it was determined that the collapse was reduced, the resistance to cracking was increased, the flexural strength was

increased, the tensile strength in splitting was increased, and the compressive strength was relatively low. Consequently, it was deduced that the produced wire-reinforced concrete could be utilised in highway superstructures (Bolat, 2009). In a separate study, an attempt was made to enhance the mechanical properties of concrete utilised in road superstructure through the reinforcement of the material with polyamide and steel fibres. The study concluded that the most effective parameter in compressive strength and flexural strength was the use of polyamide fibre. It was observed that as the percentage of polyamide fibre incorporated into the mixture increased, a concomitant decrease in strength was recorded. It has been determined that material ratios have a significant effect on the design of road concrete using polyamide and steel fibres. Furthermore, the applicability of the Taguchi method has been envisaged to determine the optimum amount (Asatekin, 2019). In a separate study, the utilisation of basalt fibres in concrete roads was examined. The objective of this study was to enhance the flexural and compressive strength of concrete by incorporating basalt fibres into the concrete mixture. Furthermore, improvements in the mechanical properties of concrete using different types and amounts of fibres were comparatively studied. Consequently, the mechanical properties of the concrete samples containing fibre additives exhibited an enhancement in comparison with the unreinforced control samples. A comparative analysis was conducted to assess the usability of these materials in concrete roads, and it was observed that the addition of steel fibre, basalt fibre and polypropylene fibre resulted in optimal performance (Çevik, 2014).

The object of this study is to examine the microstructure of RCC road concrete with nano powder and basalt fibre additives by BET and SEM analysis, with a view to revealing the effects of these additives on road concrete in detail. BET analysis is a widely utilised technique in the field of materials science, with applications including the comparison of pore structures across diverse material types. It serves as a valuable tool for examining the impact of additives on enhancing the specific surface area of materials. Additionally, it facilitates the evaluation of the influence of these structural modifications on the mechanical properties of the material, offering insights into the relationship between material design and performance. For instance, studies evaluating the impact of nano-additives on the microstructure of concrete have underscored the significance of BET analysis in the filling of micropores [20]. Additionally, it is extensively utilised in the investigation of the effects of innovative nano-additives, such as carbon nanotubes (Seis et al, 2022). On the other hand, SEM analysis facilitates the visualisation of the microstructure of the material, thereby enabling a more detailed examination of cracks, fibre distributions and filling effects. SEM analysis is a frequently employed technique in the evaluation of microstructural features, including the distribution of fibres and the observation of cement hydration products. In academic literature, the SEM images provide valuable insights into the distribution of nanomaterials within the cement matrix, and the synergistic effects they create with fibres. The present study focuses on the pore-filling ability of nano powder and the role of basalt fibres in crack control. The aim is to reveal their contributions to the strength and longevity performance of the material. In this context, the connections between microstructure and material performance will be discussed in depth, and inferences that will guide future applications will be made in the light of the findings. The results of the BET and SEM

analyses obtained in the present study will be compared with those of similar studies in literature by correlating them with compressive, flexural and abrasion tests. This approach will facilitate a more comprehensive evaluation of the impact of nanoadditives and basalt fibres on microstructure and macro performance.

2. MATERIAL AND METHOD

2.1. Material

In this study, CEM I 42.5R type cement produced by Aşkale Çimento company, which complies with TS EN 197-1 standards, was used as the binder (TS EN 197-1, 2012). The mixing water was obtained from the municipal drinking water supply, complying with TS EN 1008 standards (TS EN 1008, 2003). Aggregates of three different sizes (0-5 mm, 5-12 mm, and 12-22 mm) were sourced from Mazlumoğlu Construction in accordance with TS 3530 EN 933-1 (TS 3530 EN 933-1, 2012). Granulometry curves of aggregates are shown in Figure 1.

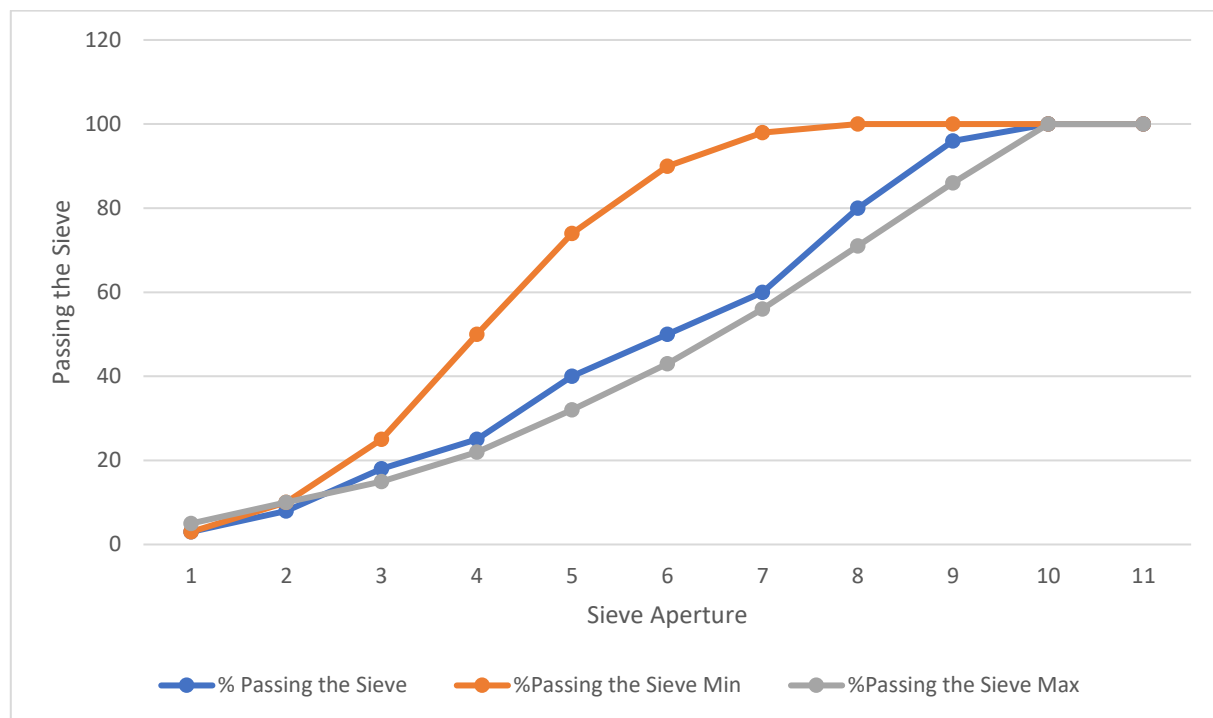


Figure 1. *Granulometry Curves*

Basalt fibre, procured from Dost Chemistry Industry, was utilized to enhance the mechanical properties of the concrete. Nano Fe_2O_3 and nano TiO_2 powders, supplied by Nanografi Nano Technology, were incorporated into the mix to improve the microstructure and durability characteristics. Fly ash obtained from Yatağan Thermal Energy Production Corporation was used as a supplementary cementitious material in accordance with TS EN 450-1 (TS EN 450-1, 2013). Additionally, a plasticizer (Master Pozzolith-515) from Master Builders Solutions, complying with the TS EN 934-2+A1 standard, was included to increase the workability of the concrete (TS EN 934-2+A1, 2013).

2.2. Method

2.2.1. Mixture Calculation

Three mixtures were prepared using the optimum mixture calculation determined in the studies of Sünbül and Tortum (2024), and the proportions of these mixtures are given in Table 1 (Sünbül & Tortum, 2024). Following a 28-day period of curing, the prepared mixtures were subjected to a series of mechanical tests, including compressive, flexural and abrasion. The optimum values were thus evaluated separately, and the most ideal mixture was determined.

Table 1. Optimum Mixture Proportions

Mixture No.	Cement (kg)	Fly Ash (kg)	Nano Fe ₂ O ₃ (kg)	Nano TiO ₂ (kg)	Basalt Fibre(kg/m ³)
Mixture-1	322	17.5	0	10.5	0
Mixture-2	339.5	0	10.5	0	0
Mixture-3	329	0	7	14	2.5

2.2.2. Slump Test

The slump test is typically conducted using a truncated cone, known as a slump cone (Clayton et al, 2003; Roussel & Leroy, 2005; Bayar, 2019). It is expected that the slump value will be within certain limits. Concrete with slumps greater than the limit value demonstrate a decrease in strength, whilst those with slumps less than the limit value exhibit a decrease in workability.

Given that the concrete designed in this study is RCC, it is anticipated that the slump value will be approximately zero. The utilisation of a roller for the placement of the concrete is expected to mitigate any potential issues arising from the low slump value, ensuring optimal workability. The slump test was conducted independently on three distinct mixtures, in accordance with the provisions of the TS EN 12350-2 standard (Figure 2) (TS EN 12350-2, 2019). The slump values were found to be in the range of 0.3-0.5 mm.



Figure 2. Applied Slump Test

2.2.3. Abrasion (Bohme) Test

The phenomenon of abrasion is associated with the aggregate's resistance to abrasion. The aggregate demonstrating the greatest resistance to abrasion also exhibits a similar effect in concrete and exhibits a high level of compressive strength (Bayar, 2019). The addition of nano or fibrous materials to concrete has been demonstrated to result in a variation in the abrasion resistance, depending on the ratio of the materials. Consequently, under conditions where the appropriate proportion and uniform distribution of these materials can be accomplished, an enhancement in abrasion resistance may be observed (Seis et al, 2022). A variety of test methods are available for the purpose of determining abrasion. To ascertain the abrasion resistance of RCC, the subject of this study, the Böhme abrasion test was selected as the abrasion test. The samples prepared for the test were then subjected to the test using the Baz Makine brand Böhme (abrasion) test device in accordance with the TS 2824 EN 1338 standard (TS 2824 EN 1338, 2005). In the Böhme test, a device with a horizontally placed rotating abrasive disk of approximately 750 mm diameter, rotating at a speed of $30 \text{ rpm} \pm 1 \text{ rpm}$, was utilised on cube samples with side lengths of 70 mm (Figure 3). A quantity of $20 \text{ g} \pm 0.5 \text{ g}$ of abrasive powder was dispensed onto the friction strip, which was then loaded with $294 \pm 3 \text{ N}$ by means of a steel lever. The test, which consists of 22 cycles for each sample, was applied to sample 16 times.



Figure 3. Abrasion Application to The Sample

After the meticulous cleansing of the test sample, its dimensions were measured with a precision caliper and weighed on a precision balance. Consequently, the initial and final weights of the sample were measured, and volume losses were calculated.

Total volume loss was calculated according to Equation 1 below:

$$\Delta v = \Delta m / d \quad (1)$$

Here;

Δv : Total volume loss after 16 cycles (mm^3)

Δm : Total mass loss after 16 cycles (g)

d : Sample density (g/mm^3)

The visual appearance of the test samples prior to and following abrasion is illustrated in Figure 4.



Figure 4. Condition of the Sample Before and After Abrasion

2.2.4. Compressive Strength Test

In the context of RCC roads, the compressive strength of concrete assumes significant importance. To ensure that the road can safely and efficiently accommodate the anticipated traffic loads, with minimal deterioration, it is imperative that the compressive strength of the concrete is sufficiently high. The findings of the studies indicated that the mean of the 7- and 28-day compressive strength test results were met for RCC mixtures (Öztürk, 2018; Abut, 2017; Yağtu, 2019). The compressive strength test was conducted utilising the Besmak brand compressive strength test device, in accordance with the requirements of the TS EN 12390-5 standard (TS EN 12390-5, 2019). The samples upon which the test was applied have dimensions of 150x150x150 mm. The cube samples were positioned in a manner that ensured the load application direction was perpendicular to the concrete casting direction. The experiment was conducted by transferring the load to the sample without impact effect with a constant loading rate of 0.2 MPa/s (N/mm².s). In accordance with the established protocol, the load was applied at a constant velocity until the maximum load that could be applied to the sample was attained (Figure 5). At this point, the reading of the load from the indicator was documented.



Figure 5. Applying Compressive Strength Test to Concrete

2.2.5. Flexural Strength Test

In the cross-section of samples in flexural state, compressive stresses occur in the region proximate to the inner surface and tensile stresses occur in the region proximate to the outer surface (Luther-Davies, 2014). The samples that were produced for the flexural strength test were then subjected to three-point flexural testing using a Besmak brand sample flexural testing device in accordance with the requirements of the TS EN 12390-5 standard (TS EN 12390, 2019). The test was applied to 40x100x100 mm beam samples from the midpoint of the opening using the simple beam method on concrete samples (Figure 6).



Figure 6. 3-Point Flexural Strength Test Application

2.2.6. Scanning Electron Microscopy Method (SEM)

The first electron microscope to be developed was the Transmission Electron Microscope (TEM). The scanning electron microscope (SEM) was first commercialised in the 1960s and began to spread with the development of technology (Karaduman, 2017). In the last 20-30 years, the use of high-tech instruments such as SEM and TEM, which are used to determine the properties of nanomaterials and their use, has opened the doors of the nano world to researchers (Arivalagan et al. 2017). Scanning electron microscopy (SEM) is a type of electron microscope that acquires images by scanning the sample surface with a focused beam of electrons. Scanning electron microscopy (SEM) is a technique used to visualise the morphological, topographic, component structures and dimensions of the sample and its surface. The utility of SEM extends to the examination of a wide range of non-conductive samples, provided they do not possess liquid properties or conduct electricity. The use of the scanning electron microscope (SEM) allows for the examination of a wide range of materials, including metals, textiles, fibres, plastics, polymers, and particles such as sand, gravel, and pollen. The examination of non-conductive samples can be facilitated through the application of a conductive coating. The EDX (Energy Dispersive X-ray Spectrometer) detector enables the relative amounts of elements and compounds to be obtained as percentages. The mapping method facilitates the identification of elements concentrated in the target region on the SEM image. It is possible to obtain crystallographic information from the sample using the electron backscatter diffraction detector. The utilisation of multiple detectors in conjunction with SEM has been demonstrated to enhance the comprehensiveness of the resulting data. The SEM setup comprises an electron gun, condenser lenses that function to collect and direct the electron beam,

mechanical apertures that serve to control the diameter of the electron beam and scanning coils that deflect electrons to scan the sample surface. The sample chamber, which is maintained under vacuum, contains holders and detectors into which the samples are placed (Karaduman, 2017; Acharya, 2023).

2.2.7. Gas Adsorption (Brunauer and Emmett) Method

The Brunauer and Emmett methods, predicated on the adsorption-desorption process, are the most common methods employed in practice. The BET method was developed by three researchers who, in a seminal paper, employed multilayer adsorption theory to determine the specific surface area of solid materials. In this method, two adsorbents (nitrogen and water vapour) are utilised, and a drying process is a prerequisite for the preparation of samples (Skalny & Hearn, 2001; Beaudoin & Marchand, 2001; Odler, 2003). However, in samples containing CSH phase, BET N₂ value is generally lower than BET H₂O value. This phenomenon can be attributed to the smaller size of water vapour, which consequently enters smaller voids and between CSH layers, resulting in elevated values (Odler., 2003). BET analysis, an adsorption isotherm model, consists of the incremental addition of a new adsorbed molecular layer to the layer beneath it. The BET gas adsorption theory forms the basis for the measurement of surface area in materials with high specific surface area. Remarkably, the BET method gives a higher specific surface area for a fine-grained sample than for a coarse-grained sample (Brunauer, 1938). It has been established that, due to the inability of N₂ gas to reach the inner surfaces in the BET method, the measurement of N₂ gas adsorption is limited to the outer surfaces (Su, 2019). A typical graph obtained from gas adsorption in general is shown in Figure 7. This representation facilitates the determination of the volumes of the voids in the sample according to their diameters (Lange & Grasley, 2007).

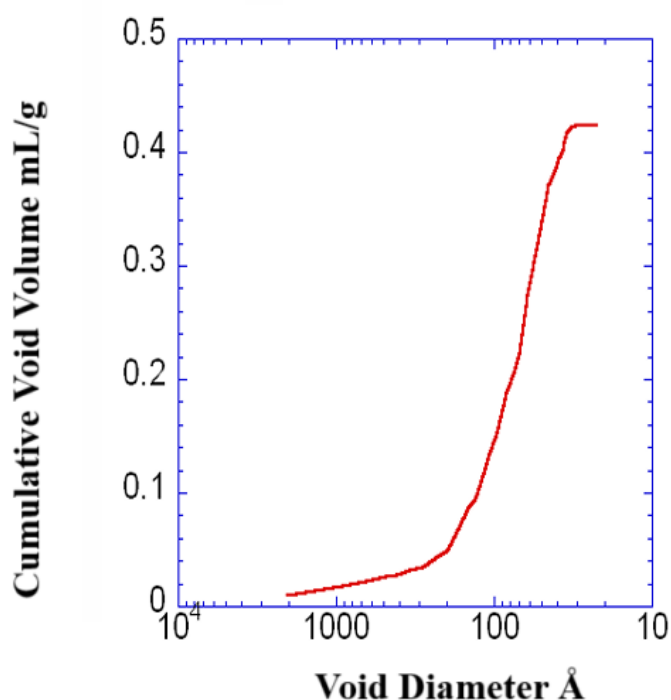


Figure 7. The Graph Obtained as A Result of Gas Adsorption

The equation developed for the BET isotherm is given in equation 2 (Noorimotlagh et al, 2014).

$$\frac{c_e}{(c_s - c_e)} = \frac{1}{BQ^0} + \left[\frac{B - 1}{BQ^0} \right] \left(\frac{c_e}{c_s} \right) \quad (2)$$

Here;

Cs: Saturation concentration of adsorbate (mg/L)

Qo: Amount of adsorbate adsorbed on all layers on the surface of the adsorbent (mg/g)

B: BET constant (gives the energy relationship between the adsorbent surface and the adsorbate).

3. RESULTS AND DISCUSSION

3.1. Results

3.1.1. Compression, Flexural and Abrasion Test Results

To conduct a mixture calculation, three test samples were prepared for each calculation. These samples were then subjected to a series of tests, including compressive, flexural and abrasion tests. The mean values of the test results were calculated to obtain the average compressive, flexural strength and abrasion strength values. The compressive strength of mixture 1 was determined to be 35.46 MPa, that of Mixture 2 was found to be 45.88 MPa, and mixture 3 was found to be 61.20 MPa. It is evident from the data presented in Figure 8 that mixture 3 exhibited the highest compressive strength. The primary cause of this increase is the enhancement of pozzolanic reactions by nano Fe₂O₃ and TiO₂, the improvement of matrix density, and the limitation of crack propagation of basalt fibres. The presence of nanoscale materials has been demonstrated to support the formation of C-S-H (calcium silicate hydrate), thereby providing a more compact and durable structure.

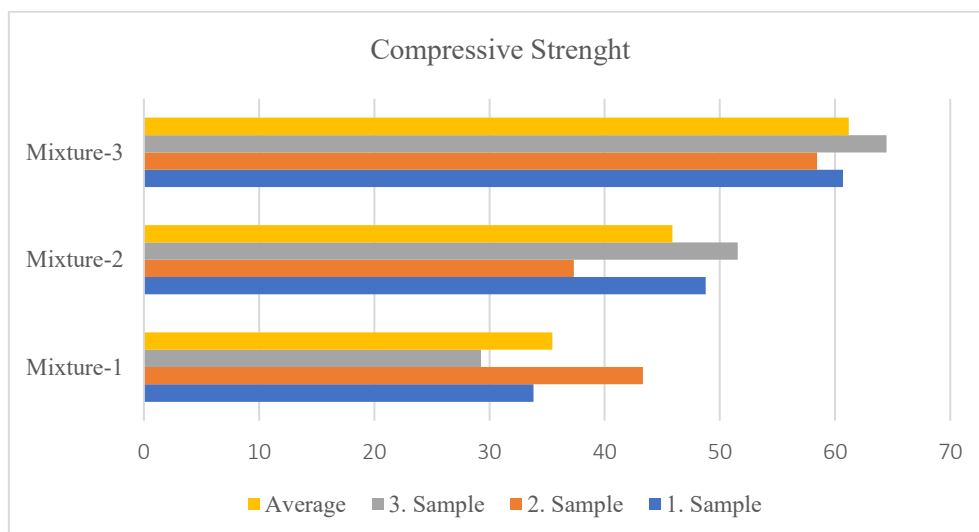


Figure 8. Compressive Strength Data

During the flexural strength test, the mean flexural strength of Mixture 1 was determined to be 7.25 MPa, that of mixture 2 was found to be 6.95 MPa and that of mixture 3 was established to be 8.80 MPa. It was determined that mixture 3 exhibited the highest flexural strength (Figure 9). This result is associated with an increase in crack resistance capacity of the basalt fibre additive and a strengthening of the bond structure of the nano

powders within the matrix. The increase in bond strength, attributable to the filling of the voids within the nano particles, and the contribution of the basalt fibres to the load carrying capacity, resulted in an enhancement of the flexural strength.

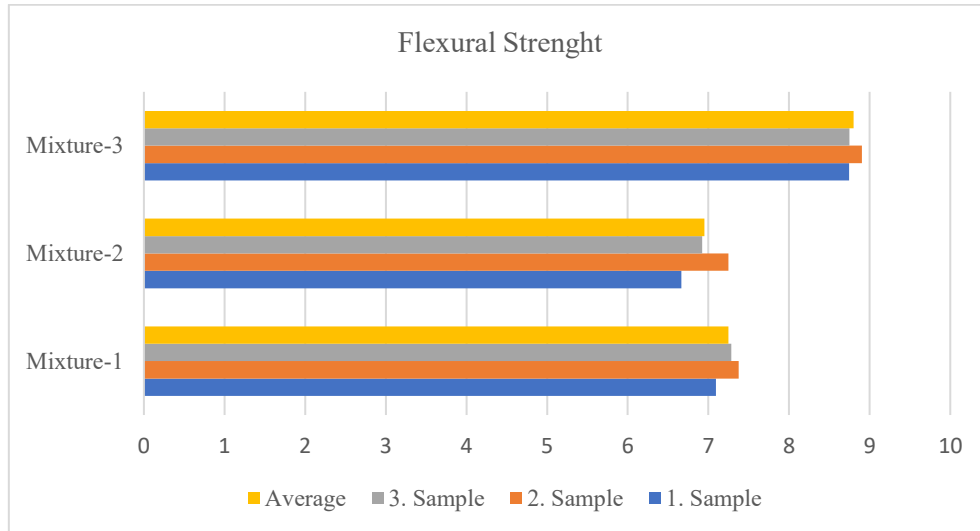


Figure 9. Flexural Strength Data

The weights of the samples that were subjected to the abrasion test were measured prior to and following the test, and the average mass and volume losses were calculated at the conclusion of the test. The average mass loss of mixture 1 was determined to be 3.9 g, the volume loss was 1.81 g, the average mass loss of mixture 2 was determined to be 3.2 g, the volume loss was 1.48 g, and the average mass loss of mixture 3 was determined to be 2.5 g, the volume loss was 1.12 g. Consequently, the mixture that exhibited minimal abrasion loss was identified as the third mixture (Figure 10). This phenomenon can be attributed to the presence of nano particles, which fill the interstices within the matrix, thereby enhancing its resistance to abrasion. The basalt fibres, in turn, contribute to the maintenance of mechanical integrity. The matrix, characterised by a denser and more compact bond structure, exhibits reduced loss when subjected to abrasive loads.

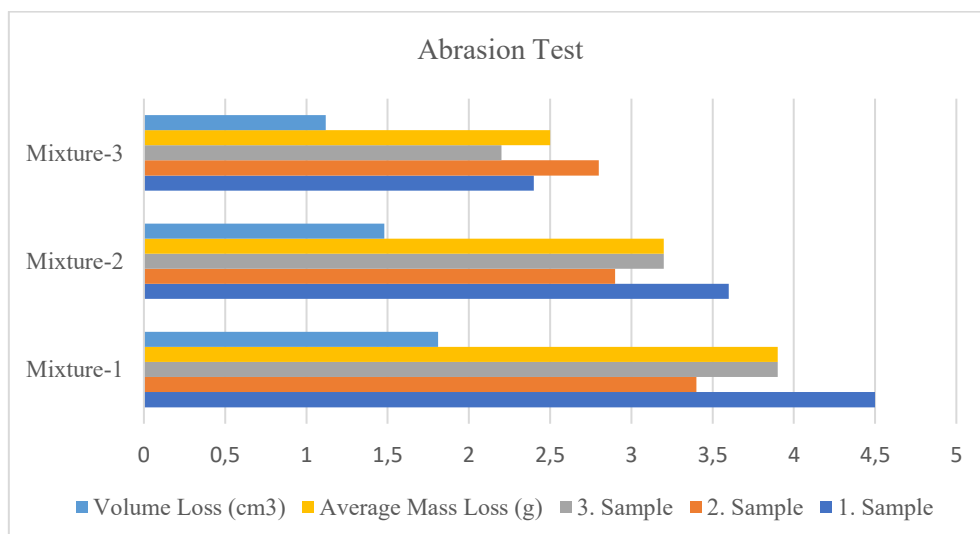


Figure 10. Abrasion Test Data

3.1.2. SEM-EDX Analysis Results

SEM imaging was performed on the optimum calculation samples that had been produced, with three growth amounts — 5k, 10k and 20k — to obtain a high-resolution internal structure view and elemental analysis with EDX. The samples were different, so there were big elemental analysis results differences. The samples were found to contain CH (calcium hydroxide) crystals, which exhibited a distinct plate-like surface morphology. These crystals are the result of the hydration process of cement, and the presence of C-S-H (calcium silicate hydrate) gels was also identified. These gels have been identified as a primary contributing factor to the strength of concrete.

Figure 11 presents the 5k, 10k and 20k magnification SEM images of the samples produced with the 1st mixture calculation. Using fly ash and nano-TiO₂ in concrete makes the structure smaller than with cement, as the particles fill the voids. A more thorough examination of the image revealed the presence of smaller voids. This finding indicates that fly ash and nano TiO₂ are inadequate in filling the voids. Consequently, it can be deduced that the compressive strength of the initial verification specimen will be lower than that of the other mixtures due to the presence of voids. The compressive strength tests also corroborate this situation. The presence of CH crystals and C-S-H gels in the mixture was also observed. Furthermore, fly ash was found to be present within the microvoids that had originated from the C-S-H or other microvoids.

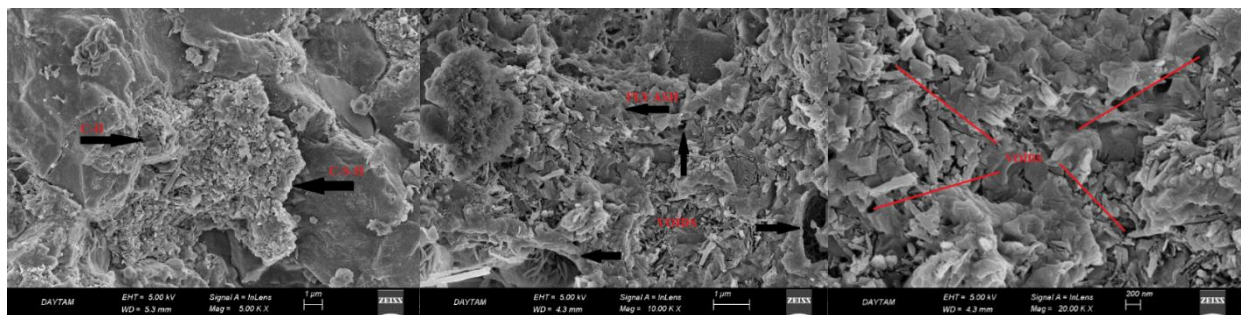


Figure 11. SEM Image of Mixture-1

Figure 12 presents the SEM images of the samples produced with the second mixture calculation at 5k, 10k and 20k magnifications. A comparison of the two mixtures reveals that the void ratio in the latter is less. The presence of a small, porous structure was attributed to the incorporation of nano-Fe₂O₃ material within the concrete. The presence of CH products has been observed at the interfaces of pores in various materials. Moreover, CH crystals and C-S-H gels have also been observed in concrete. The nanomaterial utilised in the mixture contributed to the strength of the material by forming additional C-S-H gel. Furthermore, an increase in reactions with the seeding (nucleation) effect was observed, which also resulted in an increase in strength with the filling effect. This hypothesis was corroborated by the findings of compressive strength tests.

Figure 13 presents the 5k, 10k and 20k magnification SEM images of the samples produced with the 3rd mixture calculation. Upon examination of the images depicting the mixture in which basalt fibre was incorporated, it was observed that the void structure was significantly reduced, and the C-S-H gel formations

were more intense. It is evident that, under typical conditions, basalt fibre exhibits a porous structure, thereby hindering the anticipated level of adhesion between the fibre and cement. However, the incorporation of nanomaterials into concrete has been shown to enhance adhesion. In 20k images, the presence of materials and the formation of C-S-H gels on the basalt fibre is more clearly discernible. A homogeneous structure is observed in the mixture upon examination in general. It is understood that a good adherence relationship is formed in the aggregate-cement-water reaction. Furthermore, analysis of the images indicates that the addition of basalt fibre to concrete does not result in the formation of a new phase. While there are minor voids present, the fibres appear to be well amalgamated and distributed homogeneously within the mixture. The reinforcement effect created by the homogeneous distribution of basalt fibres in the concrete has been demonstrated. Consequently, the addition of fibre to concrete is understood to enhance its mechanical properties. It has been determined through the implementation of flexural strength tests that this structure increases strength.

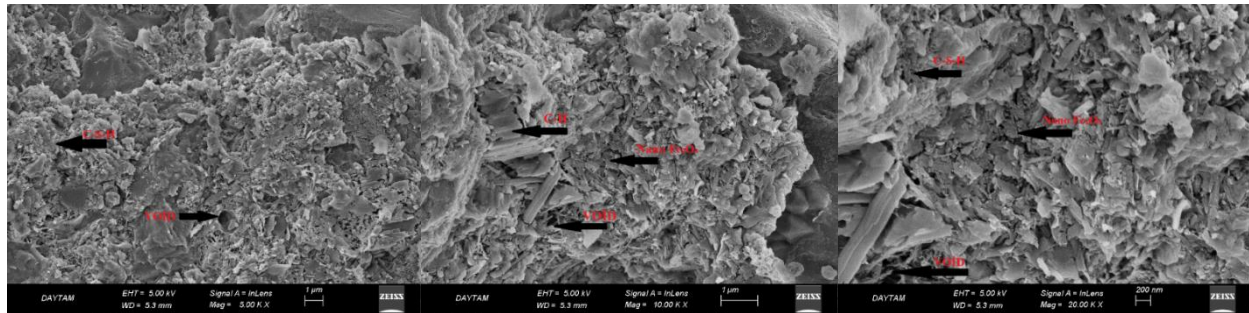


Figure 12. SEM Image of Mixture-2

The incorporation of two distinct nanopowder types in the third mixture resulted in a substantial augmentation in the amount of C-S-H formed in the concrete and the seeding effect. The primary distinction between the samples devoid of nanomaterials or comprising a solitary nanomaterial and those containing two types of nanomaterials in the SEM images pertained to the dough matrices of the samples containing nanomaterials, which exhibited increased density and a higher proportion of crystalline structure in comparison to the samples devoid of nanomaterials. Therefore, it is anticipated that the mechanical properties of the sample will undergo enhancement. The superior mechanical properties of the third sample were further substantiated by compressive strength tests. Concurrent studies by other researchers have yielded analogous results (Adak et al, 2014; Deb et al, 2015; Ramachandran & Beaudoin, 2000). The findings of the analyses support these researchers.

In order to facilitate the investigation of the C, O, Fe and Ti densities during the EDX analysis, the ratios of these elements are presented. SEM images of the third mixture with the highest strength values were examined and compared with EDX images, which provided information about the elemental composition of the sample (Figure 14). Accordingly, it can be concluded that the SEM images agree with the EDX results. A thorough examination of the EDX spectra of the third mixture reveals that the levels of carbon and oxygen in the environment exceed those observed in the other mixtures. According to the resultant data, most of the phases

formed in the third verification sample, that is, as a percentage by weight in the sample, consist of compounds of these elements. The carbon density of the material under investigation is attributable to the presence of cement and fly ash. The data obtained in this study are consistent with those reported in similar studies (Ramachandran & Beaudoin, 2000; Erdoğan, 2011). Consequently, a comparison of the three prepared verification samples revealed that the mechanical properties of the third mixture, the basalt fibre and nano powder concrete sample, would exceed those of the other concrete samples. This is attributable to the low amount of voids, the good adhesion relationship between the phases, and the formation of a homogeneous structure. The outcomes of the compressive and flexural strength assessments further substantiate this phenomenon.

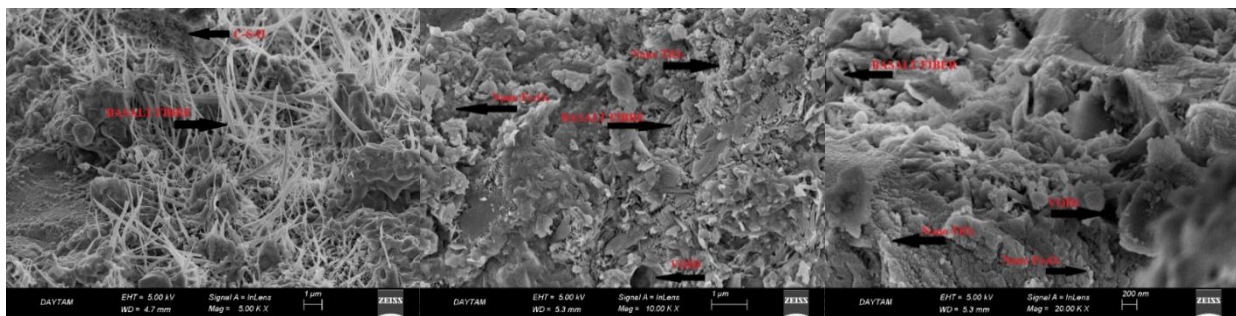


Figure 13. SEM Image of Mixture-2

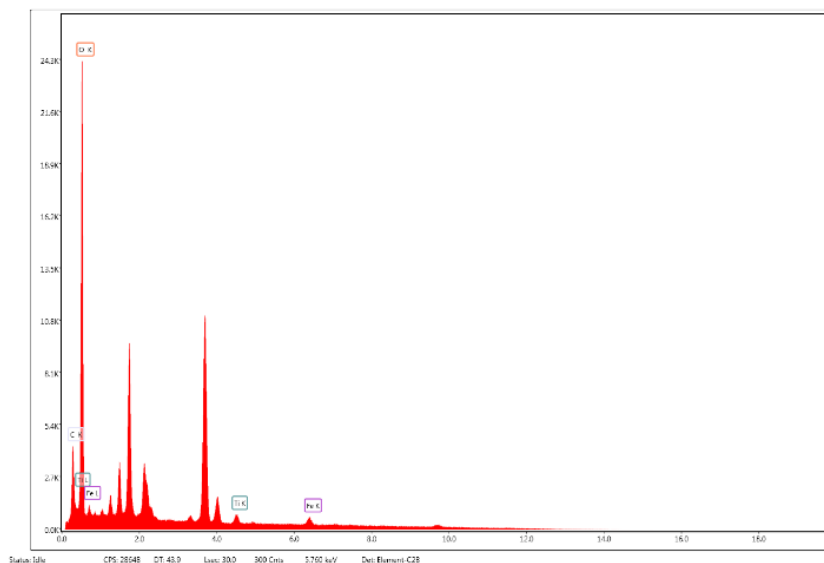


Figure 14. Third Mixture EDX Chart

3.1.3. BET Analysis Results

The BET method was applied to determine the surface area and the width of the pores of the material. The adsorption process was facilitated by the utilisation of nitrogen gas. The Barrett, Joyner and Halenda (BJH) method was utilised to ascertain the pore sizes. The BJH method, a computational procedure developed by Barrett, Joyner and Halenda in 1951, has emerged as the most prevalent technique for determining the mesopore (multipore) size distribution (Rouquerol et al, 1999; Rouquerol et al, 2002). The specific surface

area and adsorption-desorption pore sizes, calculated by the BET method, and the cumulative pore width, obtained by the BJH method, are presented in Tables 2 and 3, respectively.

Table 2. Material Surface Areas

Surface Area (m ² /g)			
Mixture No	Mixture-1	Mixture -2	Mixture -3
Single Point Surface Area (p/p° = 0,35)	11.42	10.59	5.68
BET Surface Area	11.51	11.25	6.72
Cumulative Surface Area of BJH Adsorption Pores	14.90	15.91	8.36
Cumulative Surface Area of BJH Desorption Pores	18.18	19.08	9.91
Cumulative Surface Area of D-H Adsorption Pores	19.59	20.34	12.44
Cumulative Surface Area of D-H Desorption Pores	18.88	20.18	15.11

It is anticipated that the surface area and pore size will be minimal due to the elevated density of nanomaterials in the sample derived from the third mixture, which was ascertained to be optimal. The analysis results also support this situation. Hysteresis loops are utilised in the characterisation of nanoporous materials. The International Union of Pure and Applied Chemistry (IUPAC) has determined the types of hysteresis loops, of which the first appeared in 1984 (Collet et al, 2008). Subsequently, the International Union of Pure and Applied Chemistry (IUPAC) categorised hysteresis loops into four distinct types and correlated them with pore structures (Figure 15) (Sing et al, 2004).

The analysis of the distribution of pore widths according to the hysteresis between the adsorption-desorption isotherms was conducted using hysteresis loop tables for the three mixture samples. In accordance with the UNAC classification, the hysteresis loops in all three mixtures are of the H3 type (Figure 16). As demonstrated in the tabular data, there is an absence of a discernible peak in the BJH desorption pore volume curves of the optimum mixture samples. This finding indicates that the pore size distribution of the sample is characterised by a high content, suggesting the absence of a uniform pore structure. Isotherms exhibiting H3 type hysteresis are indicative of the potential presence of interconnected micropores or mesopores within the samples (Kruk & Jaroniec, 2001).

Table 3. Material Adsorption-Desorption Pore Sizes

Pore Size (nm)			
Mixture No	Mixture-1	Mixture -2	Mixture -3
Adsorption Average Pore Diameter	8.43	9.93	6.06
Desorption Average Pore Diameter	11.47	13.60	7.83
BJH Adsorption Average Pore Width	8.75	10.34	4.51
BJH Desorption Average Pore Width	7.48	8.83	5.47
BJH Desorption Average Pore Width	7.57	9.06	4.27
D-H Desorption Average Pore Width	7.67	8.97	4.74

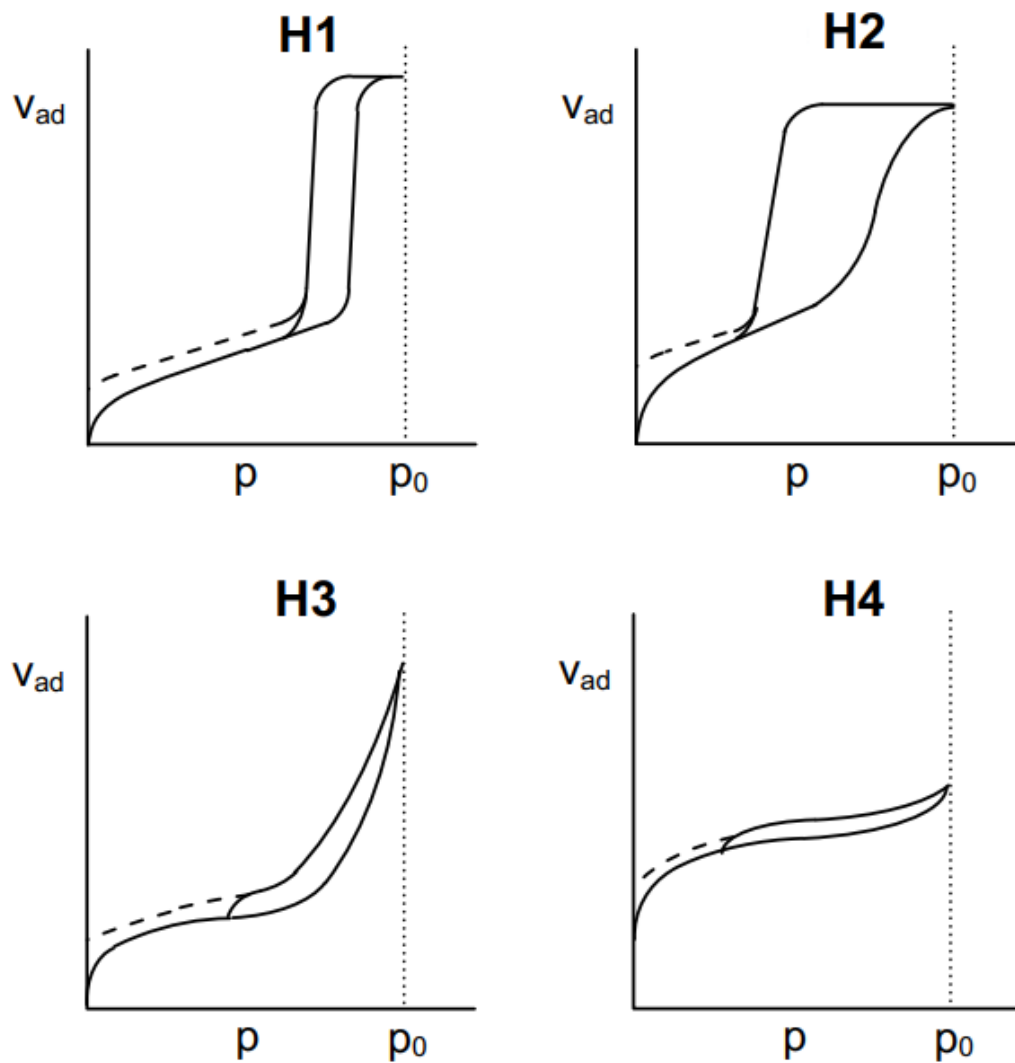


Figure 15. IUPAC Hysteresis Loop Tables

Following a thorough evaluation of the analyses, it can be concluded that the volume values of the pores in the samples are low. However, the pore type is not uniform. The findings from the analysis provide a robust rationale for the objective of reducing the void ratio, with the aim of enhancing the impermeability of concrete. Since changes in pore type depend on compression, good compression in practice will ensure that the void ratio is distributed evenly, and the pore type is not too diverse.

4. DISCUSSION

Roller compacted concrete (RCC) is characterised by a lower water/binder ratio, higher compaction requirements and a reduced void structure in comparison with conventional road concretes. The present study is an in-depth investigation into the microstructural and mechanical effects of nanomaterials and basalt fibres on RCC road concretes. SEM analysis demonstrated that nano-additive mixtures significantly reduced the voids and formed a more compact structure. It was observed that fly ash and nano-TiO₂ were present within

the voids of the first mixture, but that they did not provide sufficient density. Research has demonstrated that nano-TiO₂ promotes the formation of calcium silicate hydrate (C-S-H) during cement hydration, thereby enhancing compressive strength. This finding was corroborated in the study by Wang et al. (2020), which demonstrated that nano-TiO₂ increased the density of cement paste and enhanced the pore structure (Wang et al, 2020). The utilisation of nano Fe₂O₃ in the second mixture resulted in the partial filling of the voids, thereby enhancing the concrete's compactness and increasing its strength through the acceleration of hydration reactions. Furthermore, the process resulted in a filling effect by demonstrating a nucleation effect, thereby enhancing the strength of the material. As stated in the extant literature, the presence of nano Fe₂O₃ in cement systems has been shown to induce a nucleation effect, thereby promoting the formation of the C-S-H phase and resulting in an increase in strength (Xing et al, 2018). In the third mixture, optimum results were obtained by using nano Fe₂O₃, nano TiO₂ and basalt fibre together. As demonstrated by SEM images, the void ratio is minimised in this mixture, resulting in the formation of a denser matrix. Furthermore, it was observed that basalt fibres were distributed homogeneously in concrete without creating a new phase and increased the flexural strength by bridging micro-cracks. The contribution of basalt fibres is a frequently emphasised benefit in the extant literature, and it is well-documented that these fibres positively affect the long-term performance of concrete due to their high thermal resistance and resistance to alkalis (Ulusoy, 2016). Moreover, in a separate study in which nano TiO₂ and nano Fe₂O₃ were utilised in conjunction, SEM images were referenced, and the beneficial impact of nano powders on the compressive strength of concrete was highlighted (Mutuk, 2013).

BET analyses were conducted to reveal the detailed changes that nano-additives and basalt fibres effect on the pore structure of concrete. In the third mixture, a significant decrease in surface area and pore volume was observed in comparison to the other mixtures. This finding indicates that the impermeability of RCC concrete increased, thereby rendering the road surface more resistant to the effects of water and chemicals. Various scholars have shown that adding nanomaterials to concrete can reduce how much water it absorbs and how permeable it is. This is because the nanomaterials change the concrete's pore structure (Senff et al, 2012). Increasing the impermeability of RCC concretes is a factor that improves the long-term performance of road concretes, especially by increasing their resistance to freeze-thaw cycles. Concretes with lower void ratios have been demonstrated to impede surface erosion and chemical deterioration by preventing water from seeping into the subsurface. In addition, studies have been conducted in which nano Fe₂O₃ and nano TiO₂ have been utilised in binary combinations with other nano powders. Oltulu (2014) obtained binary combinations by using nano Fe₂O₃ with nano Al₂O₃ and Nano SiO₂ and examined the void structure of the material with BET analysis. The study concluded that it was not possible to reach a definitive conclusion regarding the effect of increasing the nanopowder ratio on the variation in the samples' properties (Oltulu and Şahin, 2014). The findings of this study demonstrate that the amalgamation of nano Fe₂O₃ and nano TiO₂ yields highly favourable outcomes in this regard, thereby ensuring a substantial enhancement in the performance of RCC concretes.

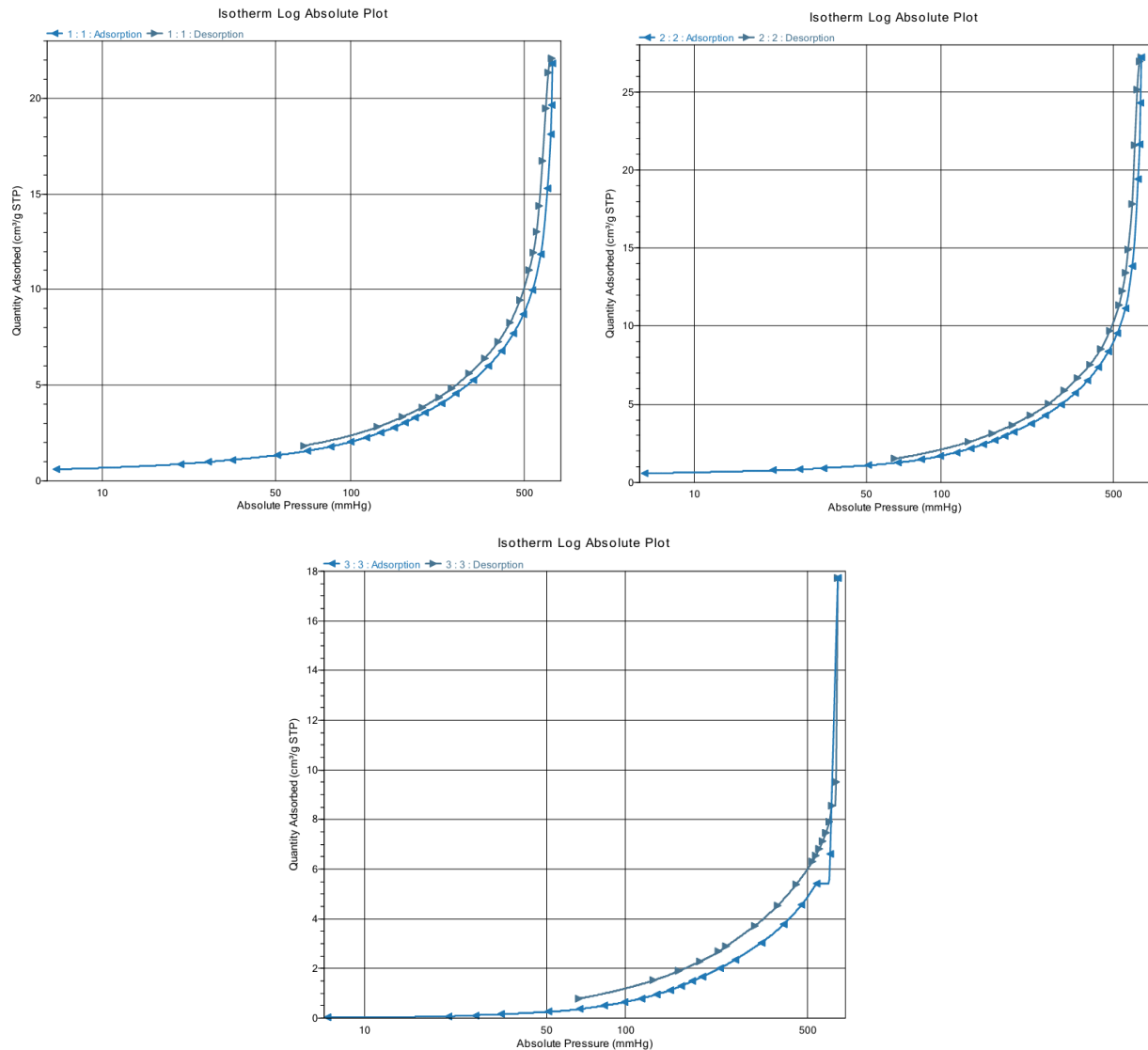


Figure 16. IUPAC Hysteresis Loop Tables

5. CONCLUSION

The compaction method employed for RCC road concrete is distinct from that of traditional vibrated concretes, a distinction that arises from the former's comparatively low water content. In this context, the present study has demonstrated that nanomaterial and fibre additives have a decisive role in determining the mechanical and durability properties of RCC. The utilisation of nanomaterials resulted in the formation of a more compact structure, whilst the incorporation of basalt fibres led to an augmentation in flexural strength due to the retardation of crack formation. This combination provides a significant advantage for increasing the resistance of RCC to traffic loads. The dense microstructure of the concrete surface resulted in reduced deformation under abrasive loads. This is of particular significance in the context of heavy road-bearing roads. The results of BET analyses demonstrate that nanomaterials reduce voids and increase impermeability. This feature is a critical parameter for longevity in road concretes. Furthermore, the finding that the slump value measured in the study was close to zero indicates that the appropriate consistency for RCC has been achieved and that the

compressibility of the concrete is at an optimum level. This outcome aligns with the recommended consistency values for RCC as outlined in the extant literature (Öztürk, 2018).

The present study investigates the effects of nanomaterials and basalt fibres on RCC road concretes. The findings of the study have led to the formulation of the following recommendations:

- **Further Elaboration of Microstructural Analysis:** The effects of nano-additives on hydration processes can be examined in more detail by using additional analysis methods such as XRD (X-ray Diffraction) and MIP (Mercury Porosimetry).
- **Long-Term Performance Tests:** To evaluate the effects of nanomaterial additives on performance over time, freeze-thaw, sulfate resistance and abrasion tests should be performed.
- **Different Nanomaterial Combinations:** The microstructure optimisations of diverse nanomaterials can be investigated by examining the combination of other nanomaterials, including nano SiO₂, nano Al₂O₃ with Fe₂O₃ and TiO₂.
- **Field Applications:** To verify the laboratory-scale results, field tests should be carried out to evaluate performance under different traffic and climate conditions.

In conclusion, this study provides significant data on the microstructural enhancement of nano-additive RCC road concretes. The extant literature contains only a limited number of studies on the use of RCC using nano Fe₂O₃, nano TiO₂ and basalt fibre together, and these are supported by both SEM and BET analyses. Consequently, the findings of this study offer significant contributions to both the academic field and practical applications in road engineering.

AUTHOR CONTRIBUTIONS

Conceptualization, Ş.S. and A.T.; methodology, A.T.; fieldwork, Ş.S.; software, Ş.S.; title, A.T. and Ş.S.; validation, Ş.S.; laboratory work, Ş.S.; formal analysis, Ş.S.; research, A.T. and Ş.S.; sources, A.T. and Ş.S.; data curation, A.T. and Ş.S.; manuscript-original draft, Ş.S.; manuscript-review and editing, A.T.; visualization, Ş.S.; supervision, A.T.; project management, A.T.; funding, A.T. All authors have read and legally accepted the final version of the article published in the journal.

ACKNOWLEDGEMENT

This work is supported by Atatürk University BAP office with the research project numbers 11012.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Abut, Y., 2017. Roller Compacted Concrete Road: A Case Study in Kocaeli Province. PhD Thesis, Kocaeli University, Institute of Science, Kocaeli.
- Acharya, A. (2023). Development And Application Of SEM/EDS In Biological, Biomedical & Nanotechnological Research. arXiv preprint arXiv:2311.00667. <https://doi.org/10.48550/arXiv.2311.00667>
- ACI Committee 327 (2014). Guide to Roller-Compacted Concrete Pavements (ACI 327R-14), Farmington Hills, MI: American Concrete Institute.
- Adak, D., Sarkar, M., Mandal, S., 2014. Effect Of Nano-Silica on Strength and Durability of Fly Ash Based Geopolymer Mortar, Construction and Building Materials, vol. 70, pp. 453–459. <https://doi.org/10.1016/j.conbuildmat.2014.07.093>
- Arivalagan, K., Ravichandran, S., Rangasamy, K., & Karthikeyan, E. (2011). Nanomaterials and Its Potential Applications. <https://doi.org/10.1002/9781394167227.ch5>
- Asatekin, A., 2019. Investigation of the Usability of Concretes Produced with Polyamide and Steel Fibre in Road Superstructure. Master Thesis, Atatürk University, Institute of Science, Erzurum.
- Bayar, E. (2019). Investigation Of Abrasion Problems in Dam Concrete. Master Thesis, Bilecik Seyh Edebali University, Institute of Science, Bilecik.
- Beaudoin, J.J., Marchand J., 2001. Pore Structure. Chapter 14, Handbook of Analytical Techniques in Concrete Science and Technology, Ed: V.S., Ramachandran, J.J., Beaudoin, Noyes.
- Bolat, H., 2009. Investigation of the Usability of Polyester and Polypropylene Fibre Concrete as Road Pavement. PhD Thesis, Gazi University, Institute of Science, Ankara.
- Brunauer, S., Emmett, P. H., Teller, E. 1938. Adsorption Of Gases in Multimolecular Layers. Journal of The American Chemical Society, 60(2), 309-319. <https://doi.org/10.1021/ja01269a023>
- Choi, Y., and Groom, J. L. (2001). RCC Mix Design – Soils Approach, Journal of Materials in Civil Engineering, 13(1), pp 71-76. [https://doi.org/10.1061/\(ASCE\)0899-1561\(2001\)13:1\(71\)](https://doi.org/10.1061/(ASCE)0899-1561(2001)13:1(71))
- Clayton, S., Grice, T. G. and Boger, D. V., Analysis of The Slump Test For On-Site Yield Stress Measurement Of Mineral Suspensions. Int. J. Miner. Proc., 2003, 70, 3–21. 18. [https://doi.org/10.1016/S0301-7516\(02\)00148-5](https://doi.org/10.1016/S0301-7516(02)00148-5)
- Çevik, N., 2014. Usability of Basalt Fibres in Concrete Roads. Master Thesis, Pamukkale University, Institute of Science, Denizli.
- Deb, P.S., Sarker, P.K., Barbhuiya, S., 2015. Effects of Nano-Silica on the Strength Development of Geopolymer Cured at Room Temperature, Construction and Building Materials, vol. 101, pp. 675–683. <https://doi.org/10.1016/j.conbuildmat.2015.10.044>
- Erdoğan D, 2011. Production and Characterization of Sodium-Potassium Silicate Based Concrete Mixtures with Glass Frit Addition. Master Thesis, Gebze Technical University, Institute of Science, Kocaeli.

- F. Collet, M. Bart, L. Serres, J. Miriel, Porous Structure and Water Vapour Sorption of Hemp-Based Materials, *Constr. Build. Mater.* 22 (2008) 1271–1280. <https://doi.org/10.1016/j.conbuildmat.2007.01.018>
- Huang, Y. H., (2004), Pavement Analysis and Design, Pearson Education, Inc (Second Edi).
- Sing, K.S.W., Williams, R.T. (2004). "Physisorption Hysteresis Loops and The Characterization Of Nanoporous Materials," *Adsorption Science & Technology*, 22, 773–782. <https://doi.org/10.1260/0263617053499032>
- Karaduman, B., 2017. Criminological Hair Analysis with Scanning Electron Microscope. Master Thesis, Bursa Uludağ University, Institute of Science, Bursa.
- Kawashima, S., Hou, P., Corr, D. J., & Shah, S. P. (2013). Modification Of Cement-Based Materials with Nanoparticles. *Cement and Concrete Composites*, 36, 8-15. <https://doi.org/10.1016/j.cemconcomp.2012.06.012>
- LaHucik, J., Dahal, S., Roesler, J., & Amirkhanian, A. N. (2017). Mechanical Properties of Roller-Compacted Concrete with Macro-Fibres. *Construction and Building Materials*, 135, 440-446. <https://doi.org/10.1016/j.conbuildmat.2016.12>.
- Lange D.A., Grasley Z.C., 2007. Pore Structure of Cement-Based Materials and A New Method to Measure Permeability.
- Luther-Davies, B. (2014). Flexible Chalcogenide Photonics. *Nature Photonics*, 8(8), 591-593.
- M. Kruk, M. Jaroniec, Gas Adsorption Characterization of Ordered Organic Inorganic Nanocomposite Materials, *Chem. Mater.* 13 (2001) 3169–3183. <https://doi.org/10.1021/cm0101069>
- Mallick, R. B., & Tahar, E.-K., (2018), Pavement Engineering Principles and Practice (Third Edit).
- Mardani-Aghabaglou, A., Bayqra, S. H., Özen, S., Altun, M. G., Faqırı, Z. A., & Ramyar, K. (2020). Design Methods of Roller Compacted Concrete Mixtures and Studies. *Pamukkale University Journal of Engineering Sciences*, 26(3), 419-431. <https://doi.org/10.5505/pajes.2019.93530>
- Menekşe, F., 2020. A Comparative Study on the Different Rigid Pavement Design Methodologies Used for Plain and Fiber Reinforced Cement-Based Materials. Master's Thesis, Boğaziçi University, Institute of Science, Istanbul.
- utuk, T. (2013). Investigation Of Physical-Mechanical and Microstructural Properties of Cement With Nano-Sized Powder Additives. Master's Thesis, Institute of Science.
- Noorimotlagh, Z., Soltani, R. D. C., Khataee, A. R., Shahriyar, S., & Nourmoradi, H. (2014). Adsorption of A Textile Dye in Aqueous Phase Using Mesoporous Activated Carbon Prepared from Iranian Milk Vetch. *Journal of the Taiwan Institute of Chemical Engineers*, 45(4), 1783-1791. <https://doi.org/10.1016/j.jtice.2014.02.017>
- Odler, I., 2003. The BET-Specific Surface Area of Hydrated Portland Cement And Related Materials. *Cement and Concrete Research*, 33, 2049–2056. [https://doi.org/10.1016/S0008-8846\(03\)00225-4](https://doi.org/10.1016/S0008-8846(03)00225-4)
- Oktay, D. Effect of Different Rates of Nano Silica Addition On Strength And Durability Properties Of Cement Based Mortars. *Nigde Omer Halisdemir University Journal of Engineering Sciences*, 13(3). <https://doi.org/10.28948/ngumuh.1464677>

- Oltulu, M., and Şahin, R. (2014). Pore Structure Analysis of Hardened Cement Mortars Containing Silica Fume and Different Nano-Powders. *Construction and Building Materials*, 53, 658-664. <https://doi.org/10.1016/j.conbuildmat.2013.11.105>
- Öcal, A., Gürü, M., & Karacasu, M. (2018). Improvement of Bitumen Performance Properties with Nano Magnesium Spinel and Colemanite. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 33(3). <https://doi.org/10.17341/gazimmfd.416397>
- Öztürk, O., 2018. Mixture Design and Mechanical Properties of Synthetic Fiber Reinforced Roller Compacted Concrete for Pavements. Master's Thesis, Istanbul Technical University, Institute of Science, Istanbul.
- Prusinski, J. (1997). Roller-Compacted Concrete Carries a Heavy Load, *Roads & Bridges*, 35(7).
- Sanchez, F., & Sobolev, K. (2010). Nanotechnology in Concrete—A Review. *Construction and building materials*, 24(11), 2060-2071. <https://doi.org/10.1016/j.conbuildmat.2010.03.014>
- Raki, L., Beaudoin, J., Alizadeh, R., Makar, J., Sato, T., 2010. Cement and Concrete Nanoscience and Nanotechnology, *Materials*. <https://doi.org/10.3390/ma3020918>
- Ramachandran, V. S., & Beaudoin, J. J. (2000). *Handbook of Analytical Techniques in Concrete Science and Technology: Principles, Techniques and Applications*. Elsevier.
- Rouquerol, F., Rouquerol, J. and Sing, K. (1999) *Adsorption by Powders and Porous Solids*, Academic Press, London.
- Rouquerol, F., Rouquerol, J. and Sing, K.S.W. (2002) *Handbook of Porous Solids*, Schuth, F., Sing, K.S.W., Weitkamp, J., Eds, Wiley–VCH, Weinheim, p. 236.
- Roussel, N. S. and Leroy, R., From Mini Cone Test to Abrams Cone Test: Measurement of Cement Based Materials Yield Stress Using Slump Tests. *Cem. Concr. Res.*, 2005, 35, 817–822. <https://doi.org/10.1016/j.cemconres.2004.07.032>
- Seis, M., Subaşı, S., Maraşlı, M., & Dehghanpour, H. (2022). Investigation of Mechanical and Electrical Properties of UHPCs with Ultra Low SWCNT Addition. *Gazi University Faculty of Engineering and Architecture Journal*, 38(1), 509-520. <https://doi.org/10.17341/gazimmfd.929690>
- Senff, L., Hotza, D., Lucas, S., Ferreira, V. M., & Labrincha, J. A., 2012. Effect of nano-SiO₂ and nano-TiO₂ Addition on the Rheological Behavior and the Hardened Properties of Cement Mortars. *Materials Science and Engineering: A*, 532, 354- 361. <https://doi.org/10.1016/j.msea.2011.10.102>
- Skalny, J. and Hearn N., 2001. Surface Area Measurements. Chapter 13, *Handbook of Analytical Techniques in Concrete Science and Technology*, Ed: V.S., Ramachandran, J.J., Beaudoin, Noyes.
- Steyn, W.J., 2011, *Applications of Nanotechnology in Road Pavement Engineering*, *Nanotechnology in Civil Infrastructure*, K. Gopalakrishnan, B. Birgisson, P. Taylor, N.A. Okine (Eds.), Springer, 49-83.
- Su, L. 2019. Research Progress in Organosilane Modification of Clay Minerals. *Conservation and Utilization of Mineral Resources*, 39(1), 124-130. <https://doi.org/10.13779/j.cnki.issn1001-0076.2019.01.023>
- Sünbül, Ş., & Tortum, A. (2024). Investigation of The Use of Nano Powder on Roller Compressible Fibrous Concrete Roads. *Case Studies in Construction Materials*, 21, e03572. <https://doi.org/10.1016/j.cscm.2024.e03572>

- TS EN 12390-5- Beton- Testing Hardened Concrete - Part 5: Determination of Flexural Strength of Test Specimens, 2002.
- TS EN 1008, Mixing Water for Concrete - Specification for Sampling, Testing and Evaluation of Suitability of Water as Mixing Water for Concrete, Including Water Recovered from Processes in the Concrete Industry, 2003.
- TS 2824 EN 1338- Concrete Covering Blocks for Floor Slabs - Required Conditions and Test Methods, 2005.
- TS EN 197-1, Cement – Part 1: Composition, Properties and Suitability Criteria of Common Cements, 2012.
- TS 3530 EN 933-1, Tests for Geometric Properties of Aggregates- Part 1: Determination of Particle Size Distribution- Sieving Method, 2012.
- TS EN 450-1, Fly Ash for Concrete- Part 1: Definition, Specifications and Suitability Criteria, 2013.
- TS EN 934-2+A1, Concrete, Mortar and Mortar Admixtures - Part 2: Concrete Admixtures - Definitions, Requirements, Compliance, Marking and Labelling, 2013.
- TS EN 12350-2- Concrete - Fresh concrete tests - Part 2: Slump Test, 2019.
- Ulusoy, U. (2016). Preparation Of Basalt and Glass Fibre Reinforced Bio-Based Polyamide Composite Materials And Investigation Of Their Properties Master's Thesis, Institute Of Science.
- Wang, Z., Yu, Q., Gauvin, F., Feng, P., Qianping, R., & Brouwers, H. J. H. (2020). Nanodispersed TiO₂ Hydrosol Modified Portland Cement Paste: The Underlying Role Of Hydration On Self-Cleaning Mechanisms. *Cement and Concrete Research*, 136, 106156. <https://doi.org/10.1016/j.cemconres.2020.106156>
- Xing, X., Xu, J., Bai, E., Zhu, J., & Wang, Y. (2018). Response surface research of the preparation of nano-Fe₂O₃ cement-based composite. *Mater. Rep*, 32, 1367-1372. <https://doi.org/10.11896/j.issn.1005-023X.2018.08.030>
- Yağtu, İ., 2019. Applicability of Rigid Pavements in Village Roads: Eskipazar Example. Master's Thesis, Karabuk University, Institute of Science, Karabuk.
- Zhu, W., Bartos, P. J., & Porro, A., 2004. Application of Nanotechnology in Construction. *Materials and Structures*, 37(9), 649-658. <https://doi.org/10.1007/BF02483294>



Analysis and Design of a W-Band MMIC LNA with 4.2 dB Noise Figure in GaAs-based 100 nm pHEMT Technology

Galip Orkun ARICAN^{1*} Burak Alptuğ YILMAZ²

¹ ASELSAN Inc, Communication and Information Technologies Division, Ankara, Türkiye

Keywords	Abstract
THz	A W-band Monolithic Microwave Integrated Circuit (MMIC) Low Noise Amplifier (LNA) is presented in this paper. The UMS PH10 process, which is the GaAs/InGaAs based pseudomorphic High Electron Mobility Transistors (pHEMTs) technology, is utilized to design the proposed W-band MMIC LNA. The proposed LNA has a simulated noise figure (NF) of 4.2 dB in the operating frequency range from 94 to 104 GHz while the simulated minimum noise figure (NF _{min}) of 3.9 dB at the center frequency. Besides, proposed W-band MMIC LNA has very good reflection loss performance, well below -10 dB and high small signal gain (S ₂₁), above 16.3 dB. Moreover, MMIC LNA is unconditionally stable up to 160 GHz. Furthermore, the proposed 3-stage MMIC LNA has a total DC power dissipation of 120 mW DC while drain voltage is 2 V. The proposed W-band LNA has a small size of 2.2 mm x 1.2 mm which yields a total chip size of 2.64 mm ² .
W-band	
MMIC	
GaAs pHEMT	
mm-Wave	

Cite

Arican, G.O., & Yilmaz, B.A. (2025). Analysis and Design of a W-Band MMIC LNA with 4.2 dB Noise Figure in GaAs-based 100 nm pHEMT Technology. GU J Sci, Part A, 12(2), 608-618. doi:10.54287/guj.1662033

Author ID (ORCID Number)	Article Process
0000-0002-9375-886X	Galip Orkun ARICAN
0000-0003-1871-7272	Burak Alptuğ YILMAZ
	Submission Date 20.03.2025
	Revision Date 08.04.2025
	Accepted Date 22.05.2025
	Published Date 30.06.2025

1. INTRODUCTION

The improvements in technology increase functionality and flexibility of communication platforms by increasing operating frequency band and reducing system dimensions (Arican et al., 2019). Radar architecture, especially spaceborne cloud profiling radars and automotive radars, fifth generation (5G) base stations and high frequency active electronically scanning arrays (AESAs) have benefited from these technological advances (Arican & Yilmaz, 2024). Therefore, the necessity of millimeter-Wave (mm-Wave) radio frequency (RF) components has been increased to deploy in modern communication systems (Murthy et al., 2024). In the meantime, upward trend in the use of modern communication systems has emerged a great demand for high-data-rate and ultra-wide bandwidth communication. So as to meet this necessity, W-band (75 GHz – 110 GHz) has a huge potential for new generation communication systems with its low-latency, high data rate and wide bandwidth advantageous (Tong et al., 2020). Hence, the benefits and drawbacks of W-band communication systems have been investigated in literature for many years (Leuther et al., 2009; Liu et al., 2019; Longhi et al., 2025; Thome et al., 2020; Thome et al., 2022; Wang et al., 2024). Besides, internet of things (IoT) technology has become popular, gradually. It gives a chance to enhance human beings' lives because any of objects can communicate with each other with high speed data rates. Low atmospheric loss,

*Corresponding Author, e-mail: goarican@aselsan.com

wide operating frequency bandwidth and short wavelength advantageous of W-band frequencies provide to design versatile and long-range communication systems especially in space/satellite platforms as a main carrier. Moreover, the most general digital modulation techniques such as binary phase shift keying (BPSK) are well suited in W-band satellite systems which have high data rate capability. Furthermore, while the number of devices that are linked to one network like one base station has been rising, the demand for high data rates and wide bandwidth has been essential. Finally, W-band with profits of high temperature sensitivity and advanced resolution is very critical for medical imaging systems and sensors (Cuadrado-Calle et al., 2024; Gao et al., 2020).

In addition, the frequency allocation system of W-band with fixed 5 GHz bandwidth results in designing simple and compact transmitter and receiver devices in comparison with Q-band and V-band which are commonly utilized for communication systems from past to present. Therefore, the necessity of W-band RF components such as LNAs, power amplifiers (PAs), couplers, power dividers/combiners and circulators, has gone up. These RF components must be compatible with space and airborne platforms, if their operating frequencies are in W-band spectrum. So as to develop W-band systems, there are a couple of MMIC process technologies such as GaAs, Indium Phosphate (InP), Silicon Germanium (SiGe), Silicon on Insulator (SOI) and Gallium Nitride (GaN) which are the most known substrate materials to fabricate MMICs. Although GaN based MMICs come to the forefront in case of their high-power density and compact design advantageous, GaAs based devices have given chance to achieve cost effective solutions (Kobayashi & Kumar, 2021).

In this paper, a W-band MMIC LNA is presented with simulation results and design constraints. The paper is organized as an introduction, MMIC process and LNA design, simulation results and discussion and conclusion parts, respectively. During the design, proper HEMT has been chosen from PH10 GaAs pHEMT process design kit (PDK) of United Monolithic Semiconductors (UMS) company according to small and large signal measurements and DC-IV characterizations of pHEMTs included in its library. HEMTs in the library have been fabricated in 0.1 μm very low-noise high gain GaAs pHEMT based technology considering RF performances. The proposed MMIC LNA operates with small signal gain (S_{21}) of 16.3 dB and noise figure (NF) of 4.2 dB. However, it is important that $\text{NF}_{\min}$ is 3.9 dB at 99 GHz which is the center of the operating frequency. Moreover, the LNA operates in the frequency range from 94 GHz to 104 GHz with 10 GHz bandwidth which yields a fractional bandwidth of approximately 10%. The device operates with very good input and output (I/O) reflection losses performance which are lower than -10 dB. Thus, it is convenient to deploy any system that required maximum 10% power reflection. Additionally, the proposed design is unconditionally stable up to 160 GHz. Thus, it can be said that there are not any oscillation mechanisms, and the stability factor is always higher than 1 up to 160 GHz. Additionally, it consumes 120 mW DC power while drain to source voltage (V_{ds}) is 2 V. Finally, the proposed device has very competitive dimensions of 2.2 x 1.2 mm^2 .

2. MMIC PROCESS AND LNA DESIGN

In the W-band MMIC LNA design, the PH10 process, which is provided by United Monolithic Semiconductors (UMS), was utilized. The PH10 is a GaAs/InGaAs pseudomorphic high electron mobility transistor (pHEMT) technology with a short 100 nm gate length which optimized by UMS to have high linear gain and low noise performance up to 110 GHz. In addition, the PH10 process has consisted of two interconnect metallization layers, metal-insulator-metal (MIM) capacitors, Tantalum Nitride (TaN) resistor, Titanium Tungsten Silisium (TiWSi) resistors, airbridges and via holes. Besides, the schematic and layout design of the proposed W-band LNA was performed with utilizing Keysight's commercially available Advanced Design System (ADS) software and the UMS PH10 PDK was introduced to ADS. In the electromagnetic (EM) simulations ADS software's Momentum tool was utilized to solve the EM numerical analysis with utilizing Method of Moments.

In the first step of the design, a numerous Direct Current (DC) and Radio Frequency (RF) simulations were performed for different size of gate peripheries. In the simulations, a variety of different gate periphery was tried to achieve the minimum noise figure (NF_{min}) and optimum gain performance. So as to achieve this, the gate periphery of 4x50 μm was chosen because of its 280 mA/mm current density and 725 mS/mm transconductance. In addition, the gate and drain voltages were swept to check the both noise and maximum available gain (MAG) performance of the 4x50 μm pHEMT. According to the simulation results, it was observed that NF_{min} of the 4x50 μm pHEMT was at around 3.6 dB when the gate to source voltage (V_{gs}) was -0.3 V and the drain to source voltage (V_{ds}) was 2V for 100 GHz at ambient temperature as seen in the Figure 1. Moreover, it was also observed that the MAG is greater than 6.3 dB for 100 GHz at ambient temperature when the V_{gs} and V_{ds} were respectively -0.3 V and 2 V as seen in the Figure 2.

In the LNA design process, the noise factor (F) and NF are the crucial figure of merits (FoM) of the distortion in the signal to noise ratio (SNR), which is the ratio of signal power density to noise power density caused by the signal chain of the network (Arican et al., 2019; Arican et al., 2021; Arican & Akcam, 2023). As given in the Equation 1, F and NF are related to each other, while the F is a unitless FoM, NF can be demonstrated as the logarithm of F in unit of decibels [dB].

$$NF = 10 \log_{10} F \text{ [dB]} \quad (1)$$

The F is expressed as the ratio of the output noise power of a device to the portion of the thermal noise at the input port attributable at the standard noise temperature of 290 K. In other words, the F also can be described as the ratio of the input to output SNR as given in the Equation 2.

$$F = \frac{SNR_i}{SNR_o} \quad (2)$$

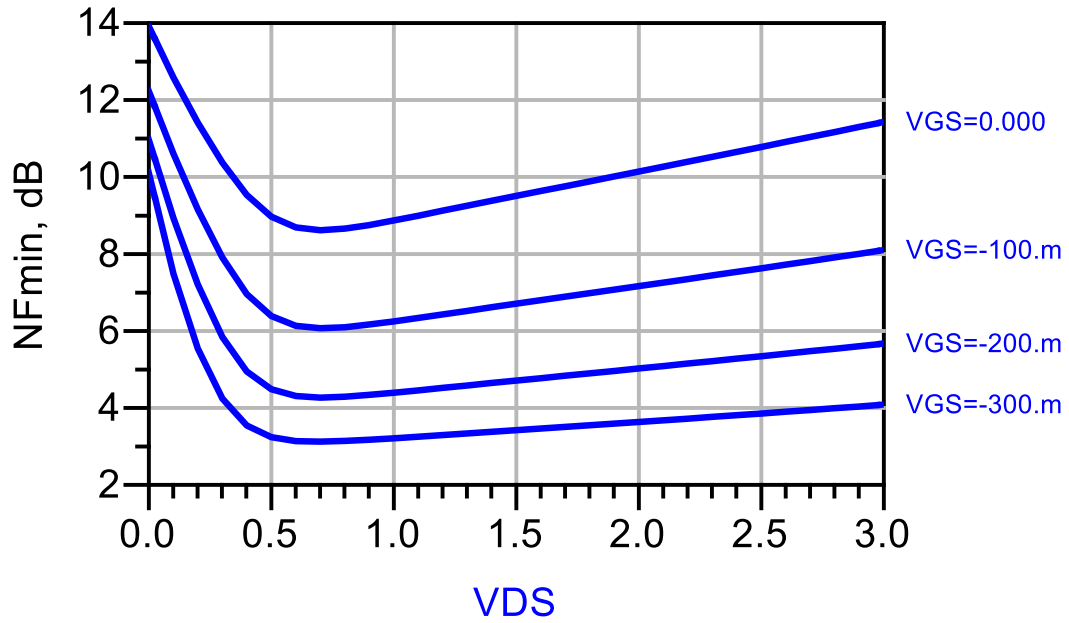


Figure 1. Simulation Results of NF_{min} at 100 GHz for $4 \times 50 \mu m$ pHEMT

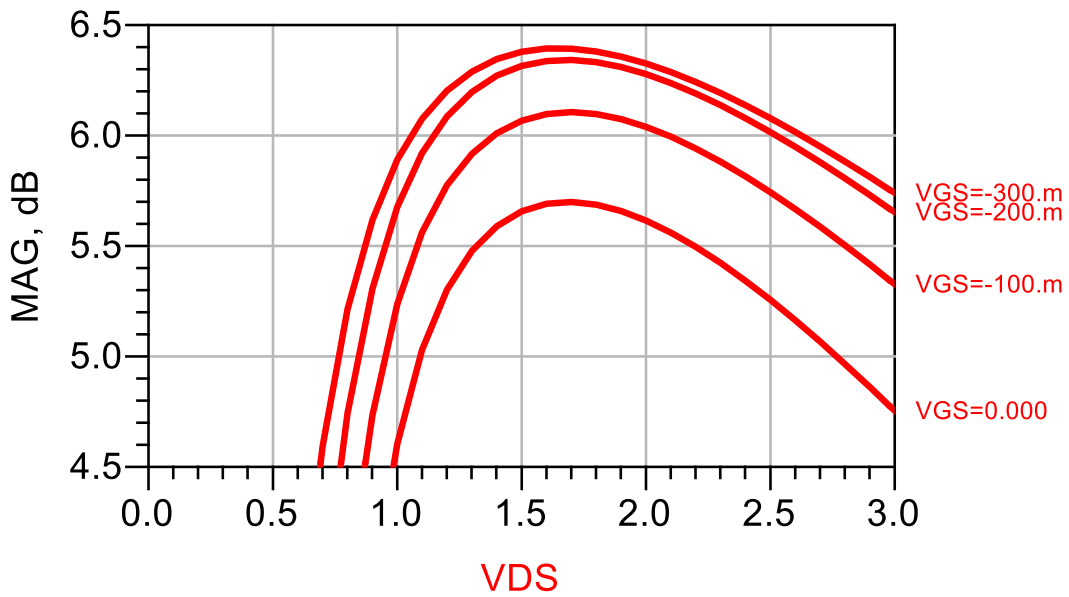


Figure 2. Simulation Results of MAG at 100 GHz for $4 \times 50 \mu m$ pHEMT

In addition, the F can be expressed as a function of reflection coefficient as given in the Equation 3.

$$F = F_{min} + \frac{R_n}{G_s} |Y_s - Y_{opt}|^2 \quad (3)$$

where Y_s and Y_{opt} indicates the source admittance ($Y_s = G_s + jB_s$ & $Y_s = 1/Z_s$) and optimum input admittance ($Y_{opt} = G_{opt} + jB_{opt}$), Z_0 is the characteristic impedance, R_n indicates the equivalent noise resistance and G_s indicates the real part of Y_s .

On the other hand, the Lange invariant parameter of N , which is the sensitivity of the noise factor between Y_s and Y_{opt} , is related to R_n via $N = R_n \cdot G_{opt}$. The F can be expressed as given in the Equation 4.

$$F = F_{min} + \frac{N}{G_{opt}G_s} |Y_s - Y_{opt}|^2 \quad (4)$$

By utilizing $F_{min}^{-1} \approx 2N$, the Equation 3 can be simplified as given Equation 5.

$$F \approx F_{min} + \frac{NF_{min} - 1}{2G_{opt}G_s} |Y_s - Y_{opt}|^2 \quad (5)$$

After choosing proper pHEMT, input and output impedance values with respect to gain and power targets, stability factor and noise figure specifications of MMIC LNA were defined. It is aimed that the device has worked between 94 GHz to 104 GHz with the gain of higher than 16 dB and better than 5 dB NF performances. In the design specifications, the minimum noise performance can be achieved with matching the source impedance (Z_s) to the optimum input impedance (Z_{opt}) as observed in the Equation 5. In the light of this fact, the load-pull and source-pull simulations was performed to define input and output impedance values of 4x50 μm pHEMTs with respect to gain and power targets, stability factor and noise figure specifications of MMIC LNA. In addition, the Z_s was matched the Z_{opt} ($12.3-j*7.4$) for the 4x50 μm pHEMTs (Q_1 , Q_2 & Q_3). Besides, minimum number of RF components have been deployed in I/O matching networks to decrease their additive effects on NF . Moreover, I/O matching networks of each transistor have been matched to 50 Ω impedance to avoid the need for an additional external matching circuit in the module design. On the other hand, in the design, 3-stages LNA has built with connecting 3 identical pHEMT in serial way for meeting S_{21} requirement. Additionally, the gate bias (TL_3 , TL_{15} & TL_{24}) and drain bias (TL_7 , TL_{12} & TL_{21}) feedlines have been consisted of the transmission line of $\lambda/4$ length to protect the DC power supply from undesirable RF signal. Furthermore, radial stub ($Stub_1$, $Stub_2$, $Stub_3$, $Stub_4$, $Stub_5$, $Stub_6$) and shunt MIM capacitors (C_2 , C_3 , C_5 , C_6 , C_7 , C_8 , C_{10} , C_{11} , C_{13} , C_{14} , C_{16} & C_{17}) have been positioned in both input and output terminals to block oscillation mechanisms while and series resistors (R_1 , R_2 , R_5 , R_6 , R_8 & R_9) have been placed to increase the stability of the circuit. By the same token, a shunt RC network (R_3C_4 , R_4C_9 & R_7C_{15}) have put in input terminals of each stage of LNA as a stabilization circuit. In the stabilization network, a resistor and MIM capacitor have been serially connected to each other and in this way, the gate current cannot flow through the ground while the circuit become stable. Hence, the device works up to 160 GHz with constant current density and without any oscillation. Additionally, serial connected capacitors (C_1 , C_{12} , C_{18} & C_{19}) have been utilized between each stage to prevent the DC flow through the RF terminals. The schematic view of MMIC LNA can be seen in Figure 3.

During the design of LNA, a couple of optimizations were performed considering initial circuit schematic. After modeling of RF components such as capacitors, inductors, and resistors, real-like behaviors of these components have been used in electromagnetic (EM) simulations of the proposed device and layout design rules. Besides, the proposed device's 2.5D layout simulations have been performed via Method of Moments (MoM) technique. Additionally, The DC bias pad has a dimension of 100 x 100 μm to be able to have at least 2 pieces of gold (Au) wire bond connections. Furthermore, gate and drain bias points have placed to north and south of the device, respectively for easy of packing and on-wafer measuring. The dimension of the device is

$2.2 \times 1.2 \text{ mm}^2$ ($0.73 \lambda_0 \times 0.4 \lambda_0$ at 100 GHz). The layout of the proposed W-band MMIC LNA can be seen in Figure 4.

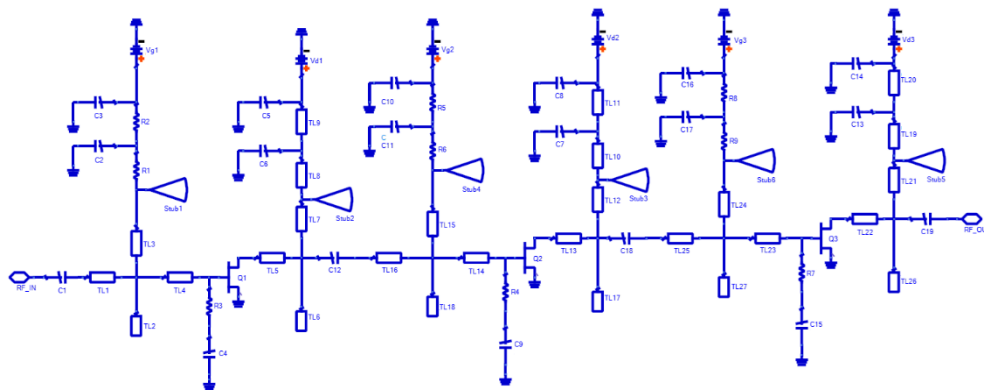


Figure 3. The Proposed MMIC LNA Schematic View

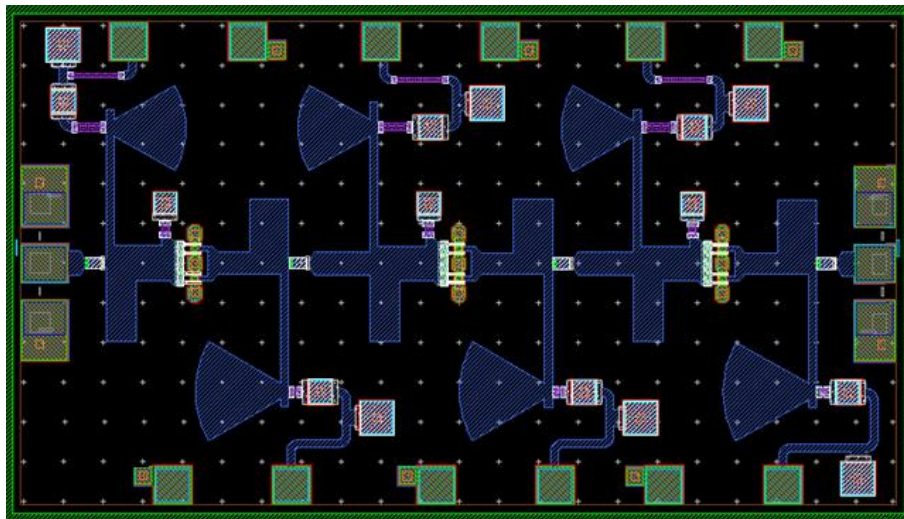


Figure 4. The layout of the proposed W-band MMIC LNA

3. SIMULATION RESULTS AND DISCUSSION

The summary of electrical and RF performances of the proposed MMIC LNA can be found in Table 1. Performances of I/O matching (S_{11} & S_{22}), gain (S_{21}), stability and NF of MMIC LNA are shown in Figure 5, 6, 7 and 8, respectively. According to simulation results, S_{11} and S_{22} are lower than -10 dB while S_{21} is higher than 16.3 dB in the range of operating frequency of 94 GHz – 104 GHz. Moreover, the gain flatness is approximately 0.13 dB/GHz. Secondly, stability performance of the device is higher than 1 up to 160 GHz with delta (Δ), lower than 1. It is crystal clear that proposed LNA is unconditionally stable. Finally, NF of the proposed W-band LNA is lower than 4.2 dB over the range of 94 GHz – 104 GHz and NF_{min} is 3.9 dB at the center frequency of 99 GHz.

The state of the art W-band MMIC LNAs is presented in Table 2. The proposed MMIC LNA has better noise and gain performances compared to study that is fabricated 3-stage with GaAs technology (Tong et al., 2020). It is stated that when 50 nm GaAs technology is preferred, NF performance gets better, and bandwidth

increases (Thome et al., 2017). Moreover, if the technology is chosen as above 100 nm, the operating frequency decreases (Tanahashi et al., 2003). According to other studies that are presented in Table 2, if the number of HEMTs used increases, gain performance increases but the proposed device has been already preferable in case of NF performance (Li et al., 2020; Sharma et al., 2024; Yawei & Wihua, 2013; Zang et al., 2018). Finally, the proposed MMIC LNA is also very competitive compared to literature in terms of its area of $2.2 \times 1.2 \text{ mm}^2$ and power consumption of 120 mW as can be seen in Table 2.

Table 1. Summary of Electrical and RF Performances

	Value	Unit
Frequency range	94 – 104	GHz
I/O matching	< -10	dB
Gain	> 16.3	dB
Gain Flatness	< 0.13	dB/GHz
NF	< 4.2	dB
Drain voltage	2	V
DC current	60	mA
Stability	Unconditional Stable	-

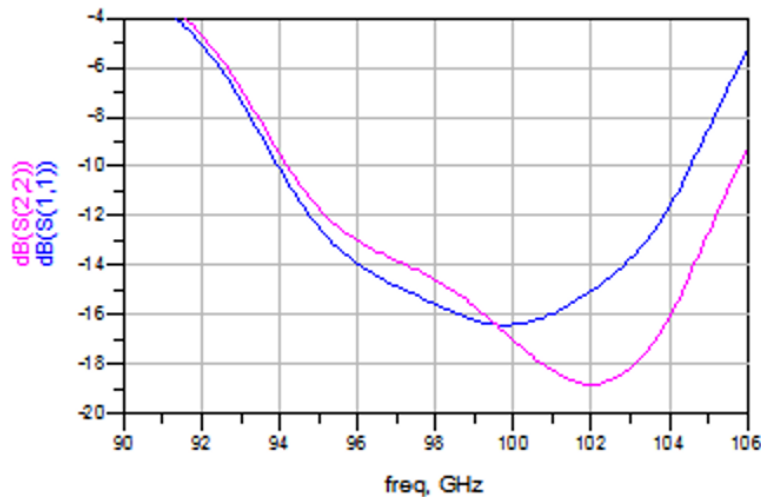


Figure 5. Simulation Results of S_{11} and S_{22}

4. CONCLUSION

A W-band MMIC LNA is brought forward in this paper. The presented device is designed with using PH10 PDK of UMS Company and the PDK provides in GaAs Technology HEMTs. According to simulation results, the simulated NF of the proposed LNA is across the 3.9-4.2 dB in the operating frequency bandwidth. Additionally, the device operates from 94 GHz to 104 GHz with reflection loss performance better than -10 dB and small signal gain performance better than 16 dB. Furthermore, the LNA works unconditionally stable, and the stability factor is always higher than 1 for whole operating frequency band and up to 160 GHz. Moreover, its P_{dc} is equal to 120 mW and layout dimensions are $2.2 \times 1.2 \text{ mm}^2$. As a result, with its prominent features, it is perfect nominee for applications such as high frequency AESAs, 5G communication systems,

IOT applications, medical imaging and sensors that require high data rates, wide bandwidth, high temperature sensitivity and resolution.

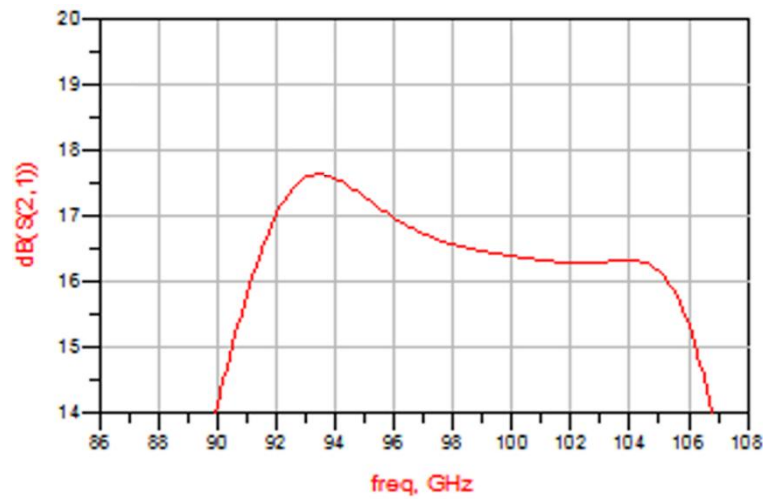


Figure 6. Simulation Result of S_{21}

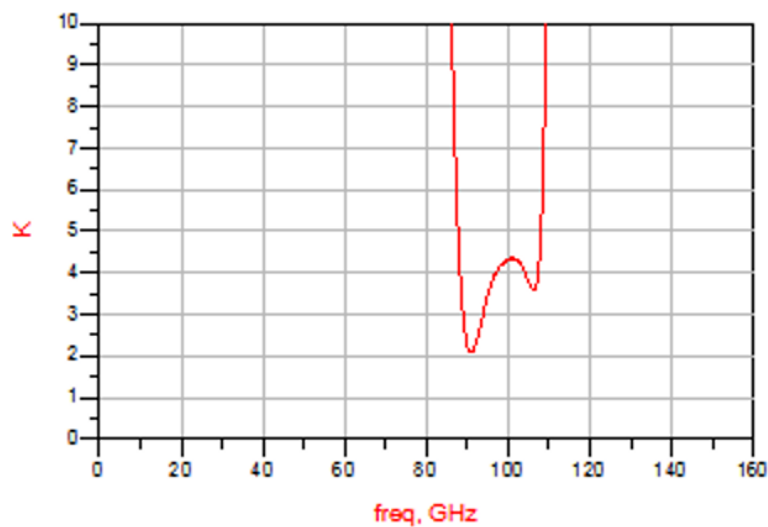


Figure 7. Simulation Result of Stability Factor

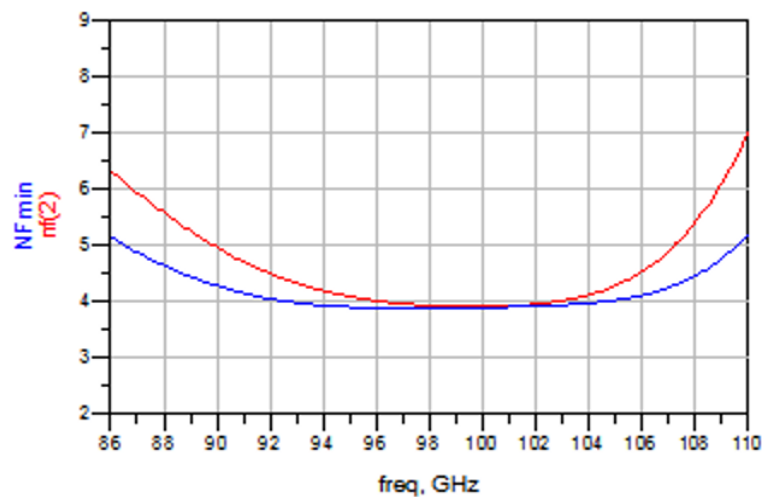


Figure 8. Simulation Result of NF

Table 2. *State of the Art W-Band MMIC LNAs*

Ref	Freq (GHz)	Gain (dB)	NF (dB)	Techn.	Area (mm ²)	# of Stage	Pdc (mW)
Tong et al., 2020	78.5 – 90	14.5	5.2	GaAs (100 nm)	3 x 1.4	3	190
Thome et al., 2017	75 – 110	20	2.6	GaAs (50 nm)	1.5 x 1.0	4	40.8
Tanahashi et al., 2003	76 – 77	15	3.5	GaAs (190 nm)	1.88 x 1.2	3	N/A
Zhang et al., 2018	75 – 110	20	5	GaAs (100 nm)	1.2 x 1.0	4	92
Li et al., 2020	80 – 100	22	5	GaAs (100 nm)	1.48 x 0.66	4	96
Yawei & Weihua, 2013	90 – 113	19	3.2	GaAs (N/A)	1.3 x 0.51	4	N/A
Sharma et al., 2024	92 – 115	19	4.6	GaAs (100 nm)	4.0 x 2.5	4	80
This Work	94 – 104	16.3	4.2	GaAs (100 nm)	2.2 x 1.2	3	120

AUTHOR CONTRIBUTIONS

All authors have read and legally accepted the final version of the article published in the journal.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Arıcan, G. O., & Yılmaz, B. A. (2024). A 10-W GaN on SiC CPW MMIC High-Power Amplifier With 44.53% PAE for X-Band AESA Radar Applications. *Electrica*, 24(3), 780-788. <https://doi.org/10.5152/electrica.2024.24090>
- Arıcan, G. O., Akcam, N., & Yazgan, E. (2019, April). Ku-band MMIC LNA design for space applications. In *2019 6th International Conference on Electrical and Electronics Engineering (ICEEE)* (pp. 274-278). IEEE. <https://doi.org/10.1109/ICEEE2019.2019.00059>
- Arıcan, G. O., Dokmetas, B., Akcam, N., & Yazgan, E. (2019, November). 28-36 GHz MMIC LNA Design for Satellite Applications. In *2019 11th International Conference on Electrical and Electronics Engineering (ELECO)* (pp. 726-729). IEEE. <https://doi.org/10.23919/ELECO47770.2019.8990444>
- Arıcan, G. O., Akcam, N., & Yazgan, E. (2021). Ku-band GaAs mHEMT MMIC and RF front-end module for space applications. *Microwave and Optical Technology Letters*, 63(2), 417-425. <https://doi.org/10.1002/mop.32613>
- Arıcan, G. O., & Akcam, N. (2022). Design of a Low Cost X-Band LNA with Sub-1-dB NF for SATCOM Applications. *Gazi University Journal of Science*, 36(1), 208-2018. <https://doi.org/10.35378/gujs.998008>

- Cuadrado-Calle, D., Kantanen, M., Valenta, V., & Ayllón, N. (2024). A GaN-on-Si MMIC LNA for Spaceborne Cloud Profiling Radars and W-Band Telecom Links. *IEEE Microwave and Wireless Technology Letters*, 34(12), 1359-1362. <https://doi.org/10.1109/LMWT.2024.3469276>
- Gao, L., Wagner, E., & Rebeiz, G. M. (2019). Design of E-and W-band low-noise amplifiers in 22-nm CMOS FD-SOI. *IEEE Transactions on Microwave Theory and Techniques*, 68(1), 132-143. <https://doi.org/10.1109/TMTT.2019.2944820>
- Kobayashi, K. W., & Kumar, V. (2021). A broadband 70–110-GHz E-/W-band LNA using a 90-nm T-gate GaN HEMT technology. *IEEE Microwave and Wireless Components Letters*, 31(7), 885-888. <https://doi.org/10.1109/LMWC.2021.3076360>
- Leuther, A., Tessmann, A., Kallfass, I., Losch, R., Seelmann-Eggebert, M., Wadefalk, N., ... & Ambacher, O. (2009, May). Metamorphic HEMT technology for low-noise applications. In *2009 IEEE International Conference on Indium Phosphide & Related Materials* (pp. 188-191). IEEE. <https://doi.org/10.1109/ICIPRM.2009.5012475>
- Li, Z., Yan, P., Chen, J., & Hou, D. (2020, September). A wide-bandwidth W-band LNA in GaAs 0.1 μm pHEMT technology. In *2020 IEEE MTT-S International Wireless Symposium (IWS)* (pp. 1-3). IEEE. <https://doi.org/10.1109/IWS49314.2020.9360082>
- Liu, C., Ren, Q., & Yang, F. (2019, March). A W-band LNA for Passive millimeter wave imaging Application. In *2019 China Semiconductor Technology International Conference (CSTIC)* (pp. 1-3). IEEE. <https://doi.org/10.1109/CSTIC.2019.8755738>
- Longhi, P. E., Colangeli, S., Ciccognani, W., Parand, P., Serino, A., & Limiti, E. (2024). 4.1-dB Noise Figure and 20-dB Gain 92–115-GHz GaAs LNA With Hot Via Interconnections. *IEEE Microwave and Wireless Technology Letters*. <https://doi.org/10.1109/LMWT.2024.3520229>
- Murthy, B. V., Manjushree, T. M., Meghana, N., Ramya, H. G., & Vaishnavi, V. (2024, May). A millimeter Wave MMIC Power Amplifier using 0.15 μm GaN for Space Communication Systems. In *2024 International Conference on Smart Systems for applications in Electrical Sciences (ICSSES)* (pp. 1-6). IEEE. <https://doi.org/10.1109/ICSSES62373.2024.10561334>
- Sharma, S. S., Sharma, S., Longhi, P., Colangeli, S., Ciccognani, W., & Limiti, E. (2024, June). Mismatch based implementation of W band LNA using GaAs pHEMTs. In *2024 19th Conference on Ph. D Research in Microelectronics and Electronics (PRIME)* (pp. 1-3). IEEE. <https://doi.org/10.1109/PRIME61930.2024.10559742>
- Tanahashi, N., Kanaya, K., Matsuzuka, T., Katoh, I., Notani, Y., Ishida, T., ... & Matsuda, Y. (2003, June). A W-band ultra low noise amplifier MMIC using GaAs pHEMT. In *IEEE MTT-S International Microwave Symposium Digest*, 2003 (Vol. 3, pp. 2225-2228). IEEE. <https://doi.org/10.1109/MWSYM.2003.1210607>
- Thome, F., Leuther, A., Massler, H., Schlechtweg, M., & Ambacher, O. (2017, June). Comparison of a 35-nm and a 50-nm gate-length metamorphic HEMT technology for millimeter-wave low-noise amplifier

- MMICs. In *2017 IEEE MTT-S International Microwave Symposium (IMS)* (pp. 752-755). IEEE. <https://doi.org/10.1109/MWSYM.2017.8058685>
- Thome, F., Heinz, F., & Leuther, A. (2020). InGaAs MOSHEMT W-band LNAs on silicon and gallium arsenide substrates. *IEEE Microwave and Wireless Components Letters*, 30(11), 1089-1092. <https://doi.org/10.1109/LMWC.2020.3025674>
- Thome, F., Brückner, P., Leone, S., & Quay, R. (2021). A wideband E/W-band low-noise amplifier MMIC in a 70-nm gate-length GaN HEMT technology. *IEEE Transactions on Microwave Theory and Techniques*, 70(2), 1367-1376. <https://doi.org/10.1109/TMTT.2021.3134645>
- Tong, X., Zheng, P., & Zhang, L. (2020). Low-noise amplifiers using 100-nm gate length GaN-on-silicon process in W-band. *IEEE Microwave and Wireless Components Letters*, 30(10), 957-960. <https://doi.org/10.1109/LMWC.2020.3019816>
- Wang, Y., Lin, F., Sun, H., Wu, H., Xu, C., Fang, Y., ... & Zeng, Z. (2023). W-band GaN T/R single chip with 1-W output power and 6.4-dB noise figure for AESA applications. *IEEE Transactions on Microwave Theory and Techniques*, 72(2), 1056-1069. <https://doi.org/10.1109/TMTT.2023.3304987>
- Yawei, W., & Weihua, Y. (2013, August). Design and simulation of a W-band broadband Low Noise Amplifier. In *2013 IEEE International conference on microwave technology & computational electromagnetics* (pp. 119-122). IEEE. <https://doi.org/10.1109/ICMTCE.2013.6812435>
- Zhang, J., Lu, Y., & Sun, P. (2018, December). W-band Broadband Low Noise Amplifier Using 0.1- μm GaAs pHEMT Process. In *2018 12th International Symposium on Antennas, Propagation and EM Theory (ISAPE)* (pp. 1-3). IEEE. <https://doi.org/10.1109/ISAPE.2018.8634085>



Gazi University

Journal of Science

PART A: ENGINEERING AND INNOVATION

<http://dergipark.org.tr/guj.1646889>

On the Structural, Electric and Optic Properties of $B_{1-x}Al_xN$ Alloys Using Ab Initio Calculation

Özlem BAYAL^{1*} Ali GÜLTEKİN² A. Kürşat BİLGİLİ² M. Kemal ÖZTÜRK²

¹ Photonics Application and Research Center, Gazi University, 06560, Ankara, Türkiye

² Department of Physics, Faculty of Science, Gazi University, 06500, Ankara, Türkiye

Keywords	Abstract
$B_{1-x}Al_xN$ Castep Code Structural Properties Elastic Properties Optical Properties	In this study, electronic, optical, elastic, dynamical and structural properties of $B_{1-x}Al_xN$ structure are investigated. During this investigation, density functional theory (DFT) in CASTEP code, local density approximation (LDA) and generalized gradient approximation (GGA) are employed. B/G ratio, Bulk module (B), Scherer module (G), Poisson's ratio (ν), Kellman parameters (ξ), Cauchy pressure (P) and compressibility are gained by using elastic constants for the mentioned structure. $B_{1-x}Al_xN$ structure showed epitaxial semiconductor behaviour with well crystallized and brittle characteristics. It also showed cubic structure behaviour for 0.25 and 0.75 Al content and tetragonal structure for 0.50 Al content. With increasing Al ratio, band gap values showed decreasing behaviour as 4.06, 3.56 and 3.34 eV respectively. It is noticed that compressibility is affected by increasing Al contents. It is also found out that real part of refraction constant (RPRF) is in accordance with real part of Dielectric function (RPDF). Peak center of losing function is gained as 27.75 eV for %25 Al content. This value corresponds with maximum Plasmon frequency.
Cite	
Bayal, Ö., Gültekin, A., Bilgili, A. K., & Öztürk, M. K. (2025). On the Structural, Electric and Optic Properties of $B_{1-x}Al_xN$ Alloys Using Ab Initio Calculation. <i>GU J Sci, Part A, 12</i> (1), 619-631. doi:10.54287/guj.1646889	
Author ID (ORCID Number)	Article Process
0000-0003-0718-9734	Özlem BAYAL
0009-0004-2381-6368	Ali GÜLTEKİN
0000-0003-3420-4936	A. Kürşat BİLGİLİ
0000-0002-8508-5714	M. Kemal ÖZTÜRK
	Submission Date 26.02.2025 Revision Date 24.03.2025 Accepted Date 10.06.2025 Published Date 30.06.2025

1. INTRODUCTION

In recent years, semiconductor devices have a rising trend because of their unique electrical and optical properties. For this reason, II-VI and III-V group compounds are begun to be investigated in detail (Troullier & Martins, 1991; Larbi et al., 2016). Well crystallized triple or quadruple alloys are begun to be used in optoelectronic device production by using different proportions of Boron nitride (BN), Gallium nitride (GaN), Aluminium nitride (AlN), Gallium Phosphate (GaP), Indium nitride (IN), Gallium arsenide (GaAs) (Teles et al., 2002; Riane et al., 2008). Due to their good properties such as wide band gap, mentioned semiconductor devices are commonly used in electronics and optoelectronics (Zheng et al., 2001; Ilyasov et al., 2005; Wu et al., 2007). Improvement of III-nitride technology can contribute constructing light emitting diodes (LEDs), laser diodes (LDs) (Watanabe et al., 2003) and modulation doped field effect transistors (MODFETs) (Zhai et al., 2012; Bouarissa et al., 2013). BN has zinc-blend cubic structure and it is constructed like diamond (Ilyasov

et al., 2005). In terms of technological properties III-Nitride semiconductors are important but they are still not much available for constructing electronic devices and crystallite growth.

Abid et al measured refraction indices and constructed $B_x Al_{1-x} N$ films which have different Boron concentration by using metal organic vapor phase epitaxy technique (MOVPE) (Matori et al., 2010). Djoidi et al calculated structural and electronic properties of $B_x Al_{1-x} N$ supercell alloys (Ponomareva et al., 2012).

In this manuscript, main purpose is to improve and adjust critical properties of semi-conductors. For this aim, optical, electronic and dynamic properties of $B_{1-x} Al_x N$ (x is Al content) crystallites are examined by using CASTEP package program that is based on DFT.

2. MATERIAL AND METHOD

Abid et al measured refraction indices and constructed $B_x Al_{1-x} N$ films which have different Boron concentration by using metal organic vapor phase epitaxy technique (MOVPE) (Matori et al., 2010). Djoidi et al calculated structural and electronic properties of $B_x Al_{1-x} N$ supercell alloys (Ponomareva et al., 2012).

In this manuscript, main purpose is to improve and adjust critical properties of semi-conductors. For this aim, optical, electronic and dynamic properties of $B_{1-x} Al_x N$ crystallites are examined by using CASTEP package program that is based on Discrete Fourier Transform (DFT).

In this study, during calculations DFT is employed by using in-plane waves (Rai et al., 2014). Pseudopotentials for B, Al and N are gained from (LDA) that is parametrized by Perdew, Burke, Ernzerhof and Troullier simulation (Shen & Zhou, 2008). Valence electron configurations of B, Al and N atoms are $2s^2 2p^1$, $3s^2 3p^1$ and $2s^2 2p^3$, respectively. In CASTEP package program, plane oscillation functions of valence electrons are described in the base constant of a plane wave. Ecut et al found out that Troullier-Martins norm-pseudopotentials curves should be used to calculate plane-wave energy (Kleinman, 1962). Interrelated spaces between atoms resembles Brillouin zone. Monkhorst and Pack Mesh used a large number of samples to investigate this situation. Wave functions extend in the plane until cut-off kinetic energy value reaches 0.6 keV. Lattice parameters in Brillouin zone are $6 \times 6 \times 6$ for these samples used. DFT charge densities are calculated numerically as 1 μC per atom.

This study is important because, during calculations both LDA and GGA are used. In calculations of electronic and optical properties LDA is used. Elastic coefficients are found by using GGA. If Al content is considered in the structure the methods mentioned above are more convenient.

In these sensitive calculations an optimization should be made. During this optimization, change in energy is taken as 1×10^{-5} eV/A, stress is 0.06 GPa, maximum displacement is 0.011 Å and SCF tolerance is 2×10^{-6} eV per atom. Structural properties of BN and AlN compounds are calculated by assuming that these compounds are in simple cubic structure.

There are eight atoms (four atoms B and four atoms N) in a unit cell of fc cubic structure belonging to BN compound. When Al is doped instead of B, crystallite structure becomes simple cubic for $B_{0.75}Al_{0.25}N$ and $B_{0.25}Al_{0.75}N$. For $B_{0.5}Al_{0.5}N$ unit cell becomes tetragonal structure. Crystallite structures of $B_{1-x}Al_xN$ are shown in Figure 1 and in Table 1 with bond length and with increasing x values. If Al content is 0.25 and 0.75, alloys are in simple cubic structure and bond lengths are the same for B-B, N-N, Al-N and B-N compounds. If Al content is 0.5, because sp^3 hybridization is predominant, alloy becomes simple tetragonal. In Figure 1 bond structure can be seen.

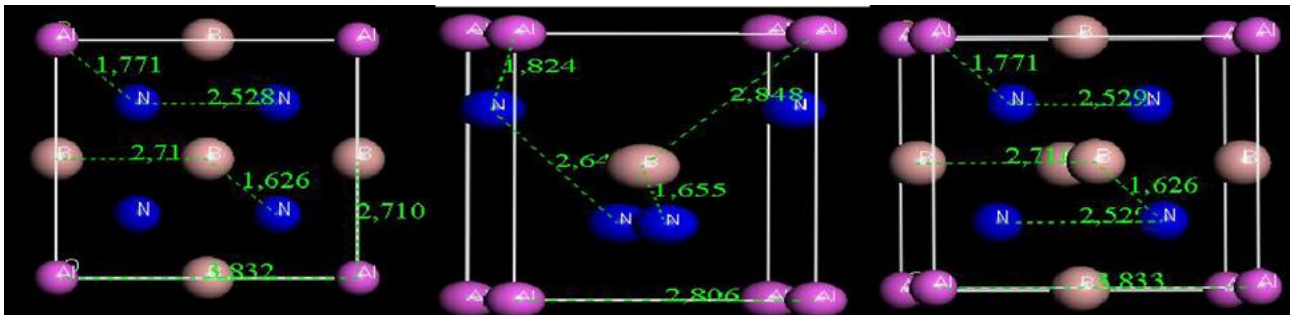


Figure 1. Bond length for cubic ($B_{0.75}Al_{0.25}N$ and $B_{0.25}Al_{0.75}N$) and tetragonal ($B_{0.50}Al_{0.50}N$) of $B_{1-x}Al_xN$ alloys.

Table 1. Crystallite structures and bond length for cubic ($B_{0.75}Al_{0.25}N$ and $B_{0.25}Al_{0.75}N$) and tetragonal ($B_{0.50}Al_{0.50}N$) of $B_{1-x}Al_xN$ alloys.

$B_{1-x}Al_xN$	Space group-Structure	B-B(Å)	N-N(Å)	Al-N(Å)	B-N(Å)
$B_{0.75}Al_{0.25}N$	P-43M Primitive cubic	2.48	2.61	4.26	4.1
$B_{0.5}Al_{0.5}N$	P-42M Primitive Tetragonal	2.64	2.64	4.31	4.31
$B_{0.25}Al_{0.75}N$	P-43M Primitive cubic	2.5	2.67	4.4	4.28

By using DFT, lattice constant, cell volume and band gaps are calculated dependent on increasing x in $B_{1-x}Al_xN$ structure. Experimental and theoretical values are shown in Table 2. Lattice constants for $B_{1-x}Al_xN$ are calculated as follows dependent on increasing x values: $x=0.25$, $a=3.79$ Å, $x=0.50$, $a=2.78$ Å, $x=0.75$, $a=4.15$ Å. Lattice parameters and Bulk modulus values are in good accordance with previously performed values. Band structures and band gap values are also alike with theoretic values (Wooten, 1972; Godlewski et al., 1995).

As shown in Figure 2, Density of states (DOS) and band structures in first Brilloune zone are calculated in equilibrium state for $B_{1-x}Al_xN$ structure using lattice coefficients. Results show that DOS and band structures correspond with each other. As seen in Figure 2, energy values show continuous behaviour. DOS curves in bands have sharp peaks. All structures have direct band transition and exhibit semi-conductor behaviour (Khenata et al., 2006; Hosseini, 2008).

Quadratic elastic coefficients are calculated numerically by using “volume conserving” method. 6 independent constants for tetragonal structure are “ $C_{ij}(C_{12}, C_{13}, C_{33}, C_{11}, C_{44}$ and $C_{66})$ ”. These coefficients must maintain “Born-Huang” criteria for optimization of calculations.

Table 2. Lattice constants, bulk modulus and forbidden band gap values for $B_{1-x}Al_xN$ alloys.

$B_{1-x}Al_xN$	a_0 (Å)	b_0 (Å)	c_0 (Å)	B (GPa)	E_g (eV)
$B_{0.75}Al_{0.25}N$	3.89 ^a ,			317 ^a ,	4.06 ^a ,
(cubic)	3.78	3.75 ^b 3.84 ^c ,	3.78	318 ^c ,	4.9 ^e ,
		3.86 ^d		296 ^e	4.05 ^f
$B_{0.5}Al_{0.5}N$	4.04 ^a ,			266 ^a ,	3.63 ^a ,
(tetragonal)	2.7	3.97 ^b , 4.04 ^c ,	2.7	266 ^c ,	4.4 ^e ,
		4.06 ^d	4.06	245 ^e	3.64 ^f
$B_{0.25}Al_{0.75}N$	4.20 ^a ,			230 ^a ,	3.48 ^a ,
(cubic)	4.14	4.18 ^b , 4.20 ^c ,	4.14	235 ^c , 214 ^e	3.7 ^e , 3.4 ^f

[a, b, c, d, e, f; Troullier, Watanabe, Larbi, Khenata, Woonten, Godlewski]

3 independent coefficients for cubic crystallites are “ C_{ij} (C_{12} , C_{11} , and C_{33})”. These must maintain Born-Huang criteria for stability “(($C_{11}>0$, $C_{11}-C_{12}>0$, $C_{44}>0$ and $2(C_{11}+C_{12})>0$)” (Dulong & Petit, 1819) found elastic coefficients for $B_{1-x}Al_xN$ alloys are calculated numerically by using DFT code. Six number of elastic coefficients are gained for $B_{1-x}Al_xN$ tetragonal and cubic structures. Calculated elastic coefficients provide each stability conditions in Table 3 (Frantsevich, 1982).

Sherar, Bulk, Young modules, compressibility, B/G, Poisson’s ratio (ν) and cauchy pressure values are calculated by using numerical elastic coefficients. Numerically calculated results for cubic $B_{0.5}Al_{0.5}N$ are shown in Table 4. B/G and Poisson’s ratio values are important for gaining knowledge about stiffness and durability of material (Born & Huang, 1996). If B/G is bigger than 1.75, structure is elastic. If it is smaller than 1.75, it is brittle (Fox, 2010). If Bulk module is gained large, this means, structure is less compressible (Wooten, 1972). If Cauchy pressure (P) is negative, directional bond is dominant in the structure, if it has positive value metallic bond is dominant. P is generally positive for elastic materials and negative for brittle materials. Poisson’s ratio is a measure of compressibility. If ν is near to $\frac{1}{2}$, it means the material has tendency to incompressible property. ν can also be used for flexibility. If ν is bigger than $\frac{1}{3}$ material shows elastic property. If it is smaller than $\frac{1}{3}$ it means material is brittle. 0.25 and 0.5 are up and down limits for ν for solids. If ξ is zero this means atom remains in center of tetragonal structure. If $\xi=1$ then only bond twist is observed the first case, if an atom settles in the center of the tetrahedral structure, the tetrahedral crystallite structure deforms due to structural defects. In the second case, bond buckling defects are more dominant (Fox, 2010).

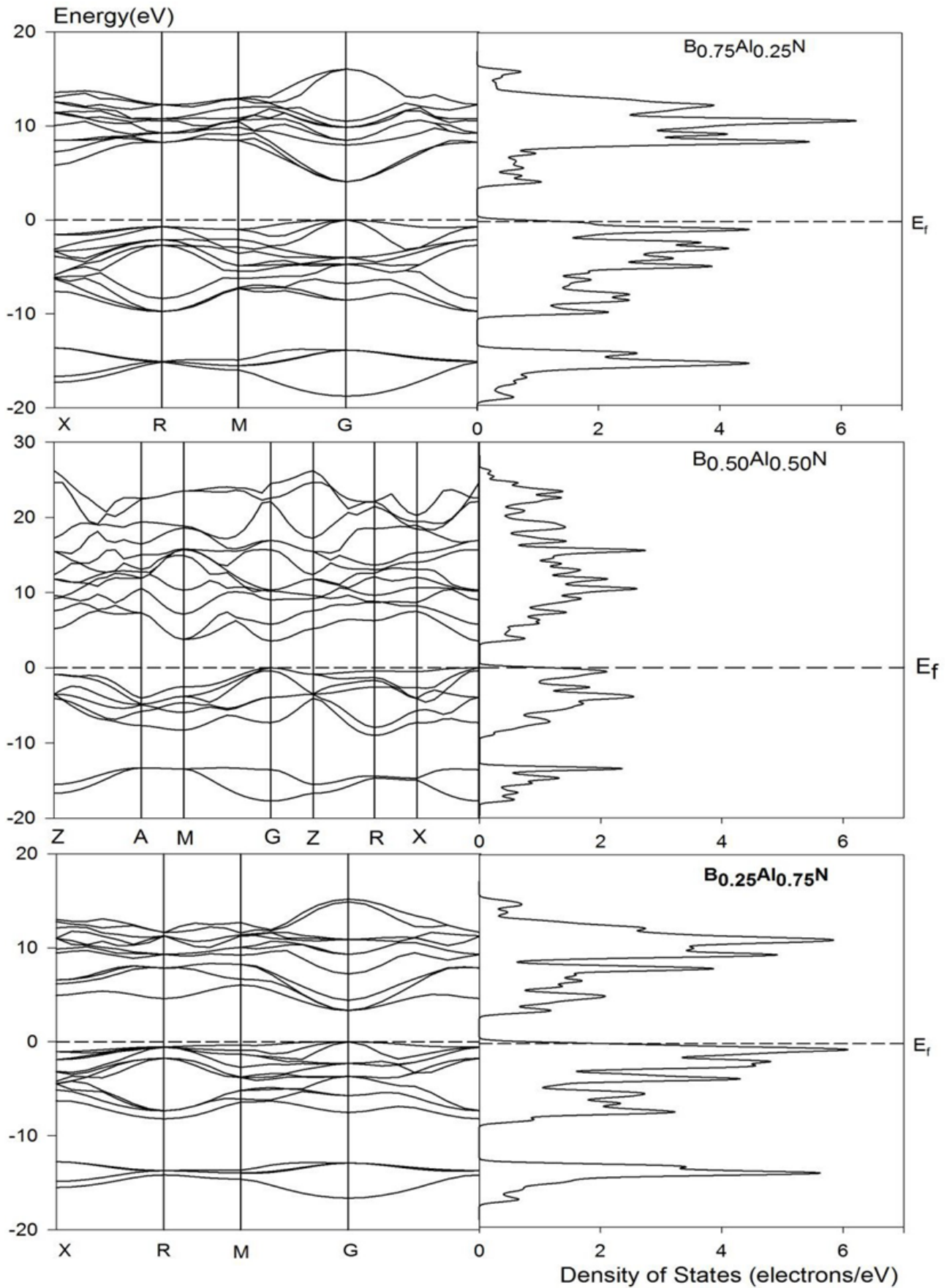


Figure 2. Band structure, band energy, forbidden band gap and state density function corresponding to high symmetry points at the Fermi energy level for $B_{1-x}Al_xN$ alloys.

Table 3. The calculated elastic constants (GPa) for $B_{1-x}Al_xN$.

$B_{1-x}Al_xN$	C_{11}	C_{12}	C_{13}	C_{33}	C_{44}	C_{66}
$B_{0.75}Al_{0.25}N$ (cubic)	588.1	195.8			308.6	
	600	39.2			236.7	
$B_{0.5}Al_{0.5}N$ (Tetragonal)	555 ^g (cubic)	176.5 ^g (cubic)	178.4	525.3	329 ^g (cubic)	116.6
$B_{0.25}Al_{0.75}N$ (cubic)	373.1	172.1			223.9	

[g; Frantsevitch]

Table 4. The estimated ν , B , G , B/G , Compressibility, ξ and P values for $B_{1-x}Al_xN$.

$B_{1-x}Al_xN$	ν	Bulk B (GPa)	Sherar G (GPa)	B/G	Compressibility (1/GPa)	Kleinman par. ξ	P (GPa)
$B_{0.75}Al_{0.25}N$ (cubic)	0.24	326.5	251.05	1.3	0.0030	0.48	-112.8
$B_{0.5}Al_{0.5}N$ (tetragonal)	0.28	278.9	186.9	1.5	0.0035	-	-197.5
$B_{0.25}Al_{0.75}N$ (cubic)	0.31	239.1	150.2	1.6	0.0042	0.59	-51.8

When the material is exposed to photon, photon interacts with electron of material. If energy of photon is equal to one of energy gap of the material electron is excited to outer energy levels. If energy of the photon is less than forbidden energy gap photon is not absorbed in the material and this material is transparent for this photon. Absorption and emission of photons cause transitions between energy levels.

$\varepsilon_2(w)$ is known as imaginary part of Dielectric function and is calculated by selection rule of wave functions and matrice components of momentum. $\varepsilon_1(w)$ is the real part of Dielectric function and is related to Kramers-Kronig function.

Other optical properties such as, refraction indice $n(w)$, absorption and extinction coefficients and function of energy loss $L(w)$ are derived from imaginary part of Dielectric function. Equations for Dielectric function, refraction indice, absorption and extinction coefficients and energy loss $L(w)$ are given below (Wooten, 1972).

$$\varepsilon_1(w) = 1 + \frac{2}{\pi} \int_0^{\infty} \frac{\varepsilon_2(w') w' dw'}{w'^2 - w^2} \quad (1)$$

$$\varepsilon_2 = \frac{ve^2}{2\pi\hbar m^2 w^2} \int d^3k \sum_{mn'} [\langle kn, p, kn' \rangle] 2f(kn) x [1 - f(kn')] \partial(E_{kn} - E_{kn'} - \hbar w) \quad (2)$$

Here $\hbar\omega$ is the energy of incident photon, p is momentum operator, $\frac{\hbar}{i} \frac{\partial}{\partial x}, kn >$ is an eigenvalue of the energy with wave function E_{kn} and $f(kn)$ is Fermi dispersion function.

$$n(\omega) = \frac{1}{\sqrt{2}} \left[\sqrt{\varepsilon_1^2(\omega) + \varepsilon_2^2(\omega)} + \varepsilon_1(\omega) \right]^{1/2} \quad (3)$$

$$L(\omega) = \text{Im} \left[\frac{-1}{\varepsilon(\omega)} \right] = \varepsilon_2(\omega) / \left[\varepsilon_1^2(\omega) + \varepsilon_2^2(\omega) \right] \quad (4)$$

$$\alpha(\omega) = \sqrt{2\omega} \left[\sqrt{\varepsilon_1^2(\omega) + \varepsilon_2^2(\omega)} - \varepsilon_1(\omega) \right]^{1/2} \quad (5)$$

Results of calculations gained from these equations are shown in Figure 3, 4, 5, 6. Peak centers of real part of Dielectric function are 8.34, 6.95, 5.8 eV for $x=0.25, 0.50$ and 0.75 , respectively. $\varepsilon_1(0)$ gives static Dielectric coefficient for Al content values $x=0.25, 0.50$ and 0.75 they are calculated as 4.17, 4.5 and 4.25 eV respectively. We mean dielectric constants corresponding to these energies.

According to doping values, absorption starts at 4.17, 4.50 and 4.25 eV energy values corresponding with imaginary part of dielectric coefficient. Values gained are in agreement with band gap energy and they imply transmission between valence and conduction band. This structure behaves like lucid material up to the value which dispersion value starts to rise. Peak centers for imaginary part of Dielectric coefficients are 11.29, 11.81 and 11.18 eV for $x=0.25, 0.50$ and 0.75 Al contents respectively. These values correspond with inter-band transitions (Wooten, 1972).

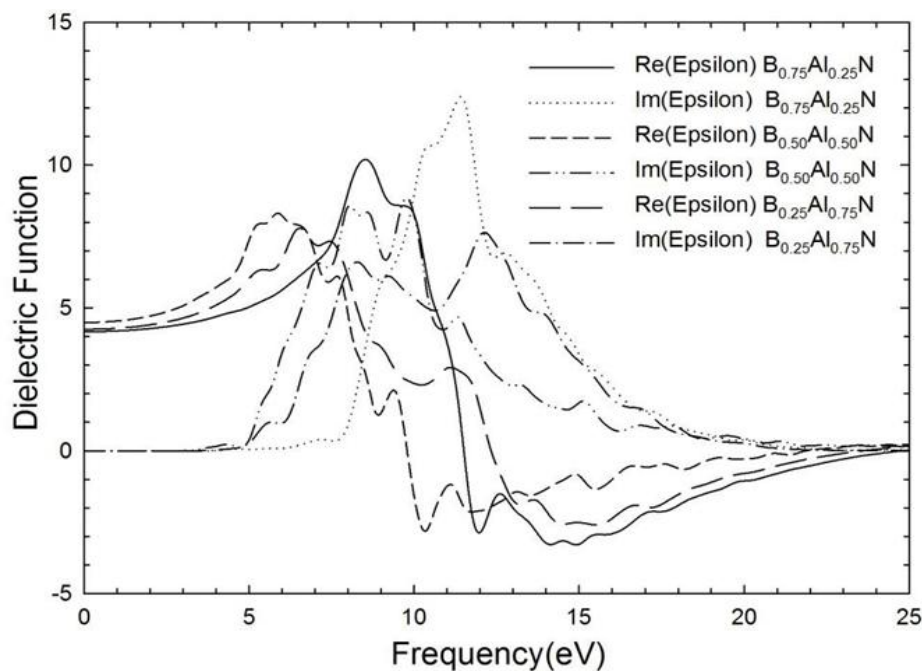


Figure 3. The real and imaginary functions of the obtained dielectric function for cubic ($B_{0.75}Al_{0.25}N$ and $B_{0.25}Al_{0.75}N$) and tetragonal ($B_{0.50}Al_{0.50}N$) of $B_{1-x}Al_xN$ alloys.

Refraction indices are gained as 2.05, 2.12 and 3.06 for $x=0.25$, 0.50 and 0.75 Al contents at $n(0)$ in accordance with curves of refraction indices. The results gained are in agreement with results of study done by Shunsuke Watanabe et al. (2003). Extinction coefficients are calculated as 4.08 for all three Al content situations. As shown in Figure 4 extinction coefficient values are in agreement with imaginary part of Dielectric coefficient.

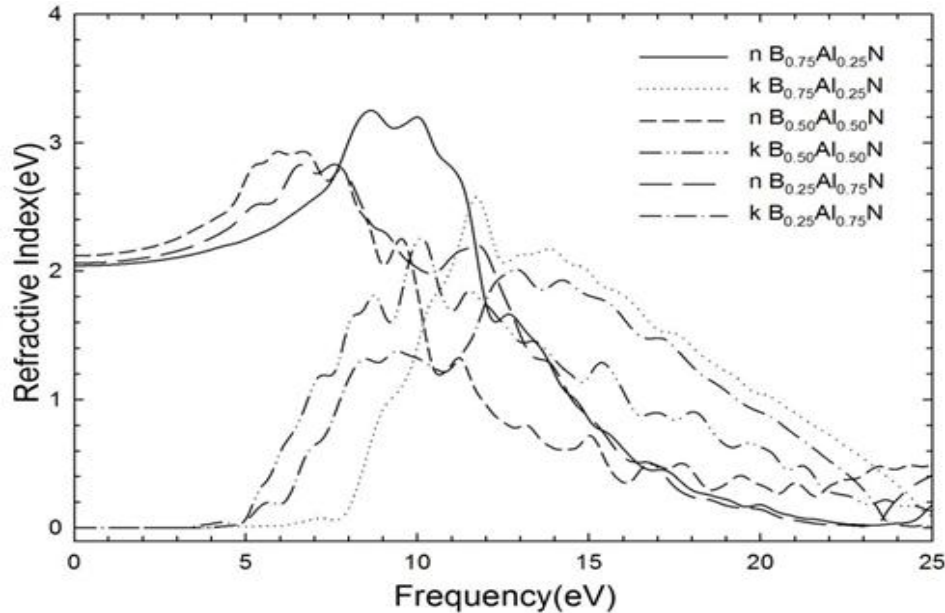


Figure 4. Refractive index n and extinction coefficient for cubic ($B_{0.75}Al_{0.25}N$ and $B_{0.25}Al_{0.75}N$) and tetragonal ($B_{0.50}Al_{0.50}N$) of $B_{1-x}Al_xN$ alloys.

For $B_{1-x}Al_xN$ alloys absorption coefficient values are calculated as 4.09, 3.58 and 3.38 for $x=0.25$, 0.50 and 0.75 Al contents, results are shown in Figure 5. At the point which absorption coefficient starts to increase, the imaginary part of dielectric function also starts to increase and are very close to each other. Energy gap at highest absorption occurs at approximately 32.21 eV value.

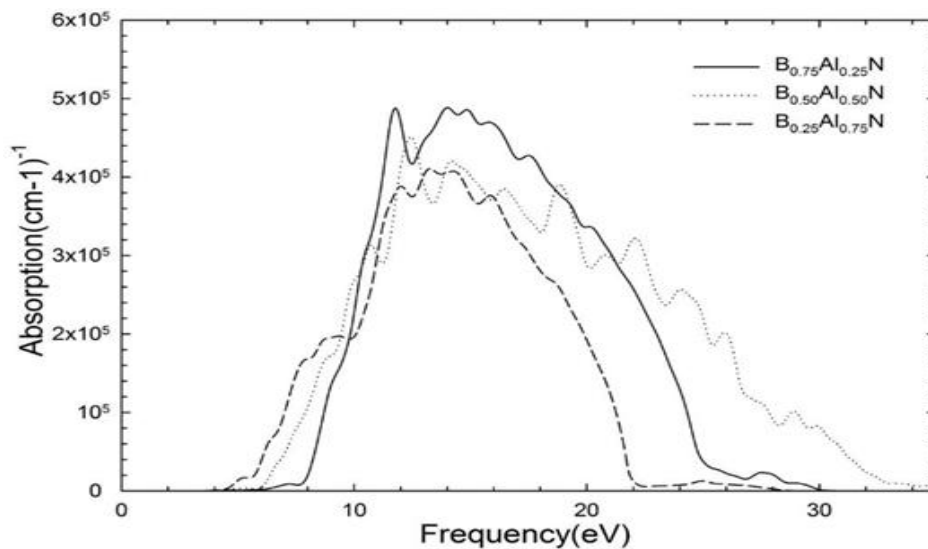


Figure 5. Absorption coefficient for cubic ($B_{0.75}Al_{0.25}N$ and $B_{0.25}Al_{0.75}N$) and tetragonal ($B_{0.50}Al_{0.50}N$) of $B_{1-x}Al_xN$ alloys.

Energy lose function of the electron is gained from real and complex part of Dielectric function and plot for it is shown in Figure 6. Lose function has variable peaks 21-32 eV energy values. Center of main peak in Losing function is named Plasmon frequency. Where the peak is maximum complex part of Dielectric constant is minimum. Depending on the values of x , Plasmon frequencies are gained as corresponding to 27.74, 26.60 and 21.76 eV energy values. If the value for Plasmon frequency is bigger than center of peak, samples behave like insulator, if it is smaller, it behaves like metal. Between 0-21 eV values absorption is very low for all three samples. The reason for this situation is that, imaginary part of Dielectric coefficient decays at this energy range (Frantsevich, 1982).

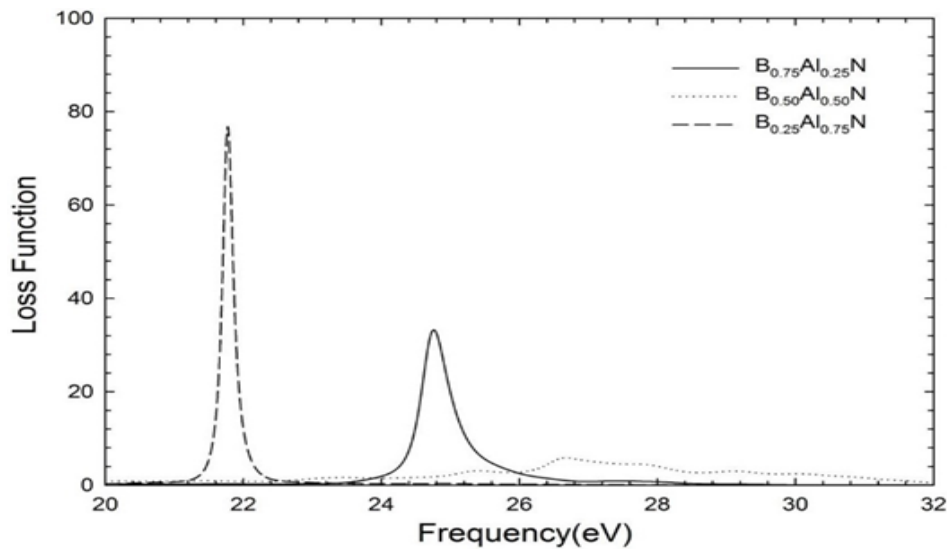


Figure 6. Energy Loss Function for cubic ($B_{0.75}Al_{0.25}N$ and $B_{0.25}Al_{0.75}N$) and Tetragonal ($B_{0.50}Al_{0.50}N$) of $B_{1-x}Al_xN$ alloys.

Because there are 4 atoms in primitive cell of $B_{0.75}Al_{0.25}N$ alloy, there are total 12 phonon branches, 3 of them are in acoustic mode and 9 of them are in optics mode. Latitudinal optics (TO) and longitudinal acoustic (LA) phonon modes overlap significantly, so there is no space in phonon dispersion curves and in DOS. For $B_{0.75}Al_{0.25}N$ alloy, central phonon modes in Brillon zone are found as 18.38, 21.06 and 30.14 THz. Because phonon frequency is not imaginary, this structure is dynamically stable. For this alloy, there are no documents in literature for dispersion curves. All three samples can be accepted as the same family member in terms of phonon frequency. Stable Bulk heat capacity (CV) dependent on temperature is determined by using dispersion and shown in Figure 7 and 8. Starting from low temperature through high temperature CV increase rapidly. For this sample, with increasing temperature CV approaches to conventional asymptotic Dulong-Petit limit. The value of this limit is 1445 Cal/cell. K (Born & Huang, 1996).

3. RESULTS AND DISCUSSION

In this study, elastic constants, optical and structural properties of $B_{1-x}Al_xN$ alloys are investigated by using CASTEP package program based on DFT. Elastic constants for zero pressure, Bulk module, Young module, Sherar module, compressibility, Poisson's ratio (ν) are calculated. Results gained are as follows: B/G ratio less than 1.75 implies samples has brittle property

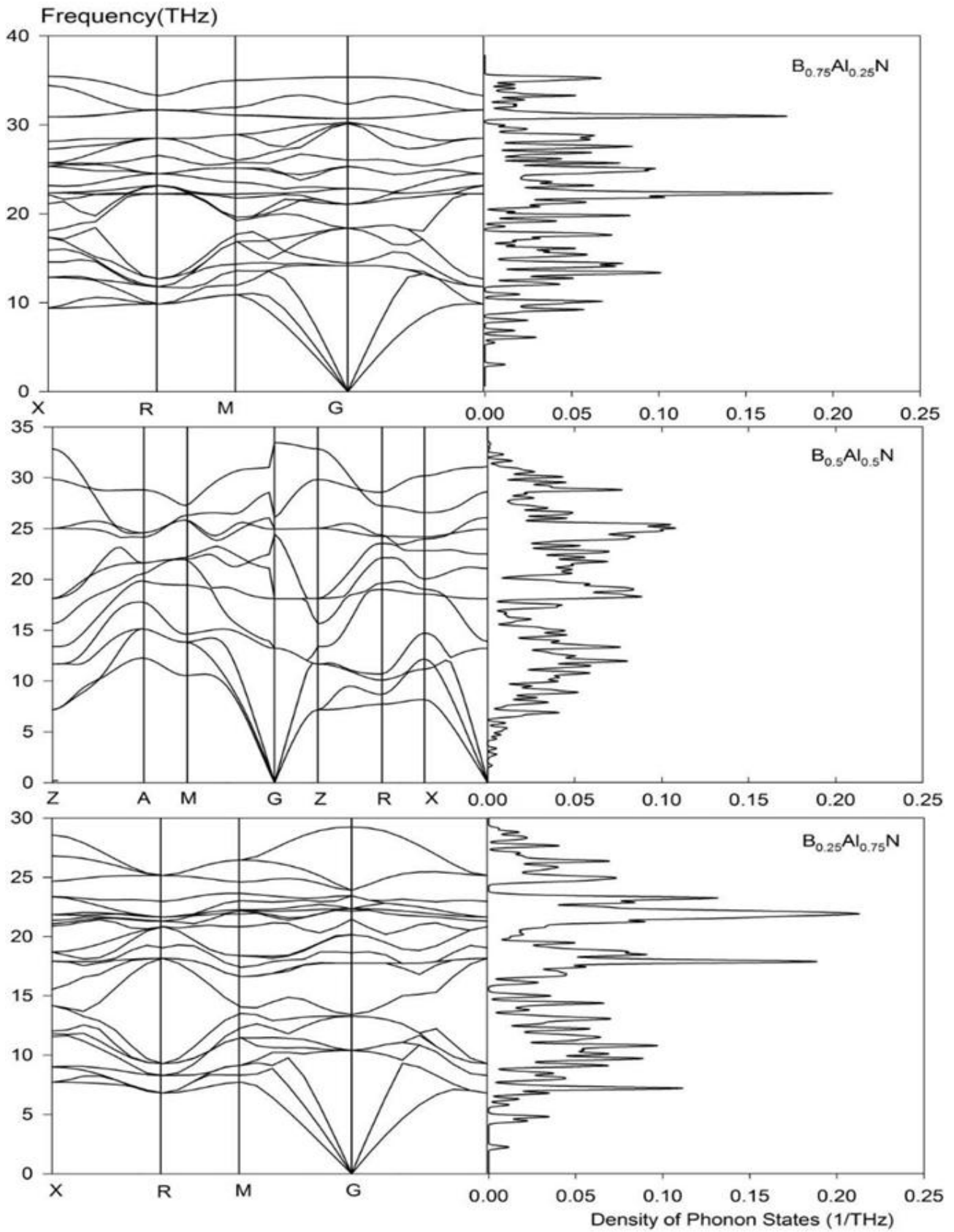


Figure 7. Phonon dispersion relation for cubic ($B_{0.75}Al_{0.25}N$ and $B_{0.25}Al_{0.75}N$) and tetragonal ($B_{0.50}Al_{0.50}N$) of $B_{1-x}Al_xN$ alloys.

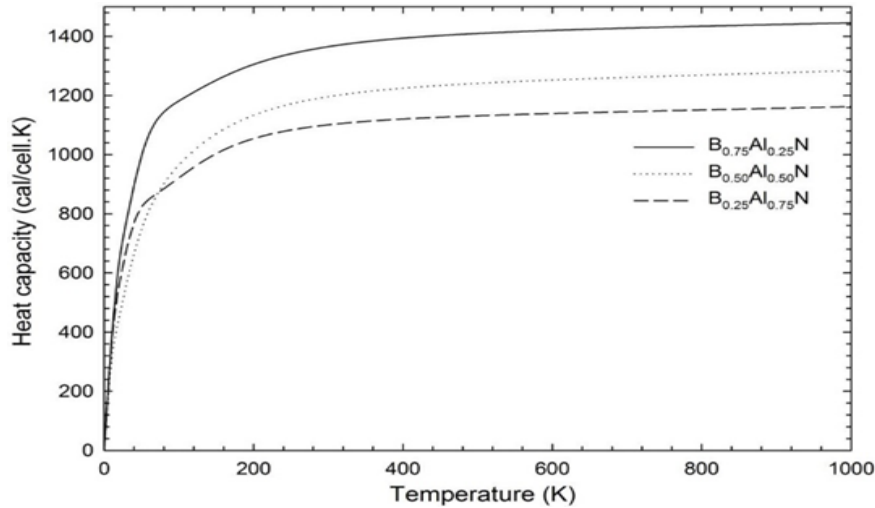


Figure 8. Heat capacities for cubic ($B_{0.75}Al_{0.25}N$ and $B_{0.25}Al_{0.75}N$) and tetragonal ($B_{0.50}Al_{0.50}N$) of $B_{1-x}Al_xN$ alloys.

4. CONCLUSION

Doping Al instead of B resulted with increasing Poisson's ratio. It is noticed that forces between atoms are mainly central forces and Poisson's ratio values for samples are in the range of 0.25 and 0.5. If Poisson's ratio value is bigger than 1/3 alloy is elastic if not it is brittle. In this study all samples present brittle property. The biggest Bulk module value is found as 326.5 GPa that implies the samples has minimum compressibility. Compressibility value is 0.0030 1/GPa. Its variation dependent on increasing X values is as $B_{0.75}Al_{0.25}N < B_{0.5}Al_{0.5}N < B_{0.25}Al_{0.75}N$. Kleinman parameters for $B_{0.75}Al_{0.25}N$ and $B_{0.5}Al_{0.5}N$ are as 0.48 and 0.59 respectively. It is not calculated for $B_{0.25}Al_{0.75}N$ because these samples have tetragonal unit cell. It is seen that in $B_{1-x}Al_xN$ alloys when Al is doped instead of B, forbidden band gap decreases and metallic property increase. Dielectric coefficients, refractive induce, absorption coefficient and energy loss function are gained by employing Kramer-Kronig equations. All samples present semiconductor behaviour which has transmission of direct band. It is seen that in the increasing values of imaginary part of Dielectric function, they are very close to forbidden band gap. Real part of Dielectric coefficient shows similar properties with refractive indice. Peak center of losing function reaches 27.74 THz value for $x=0.25$ and this is the highest Plasmon frequency.

AUTHOR CONTRIBUTIONS

Corresponding author Ö.B., drawing plots, A.K.B. and A.G., calculations, M.K.Ö., checking and editing, Ö.B.

ACKNOWLEDGEMENT

This work was supported by Presidency Strategy and Budget Directorate (Grant Number: 2016K121220).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Born, M., & Huang, K. (1996). *Dynamical theory of crystallite lattices*. Oxford university press.
- Bouarissa, N., & Saib, S. (2013). Elastic modulus, optical phonon modes and polaron properties in $\text{Al}_{1-x}\text{B}_x\text{N}$ alloys. *Current Applied Physics*, 13(3), 493-499. <https://doi.org/10.1016/j.cap.2012.09.021>
- Dulong, P. L., & Petit, A. T. (1819). *Recherches sur quelques points importants de la theorie de la chaleur*.
- Fox, M. (2010). *Optical properties of solids* (Vol. 3). Oxford university press.
- Frantsevich, I. N. (1982). Elastic constants and elastic moduli of metals and insulators. *Reference book*.
- Godlewski, M., Bergman, J. P., Holtz, P. O., Monemar, B., Bugajski, M., Regiński, K., & Kaniewska, M. (1995). Influence of Growth Conditions on Exciton Properties in Thin Quantum Wells of GaAs/AlGaAs. *Acta Physica Polonica A*, 88(4), 719-722. <https://doi.org/10.12693/aphyspola.88.719>
- Hosseini, S. M. (2008). Optical properties of cadmium telluride in zinc-blende and wurzite structure. *Physica B: Condensed Matter*, 403(10-11), 1907-1915. <https://doi.org/10.1016/j.physb.2007.10.370>
- Ilyasov, V. V., Zhdanova, T. P., & Nikiforov, I. Y. (2005). Electronic energy structure and x-ray spectra of wide-gap AlN and BN crystallites and $\text{B}_x\text{Al}_{1-x}\text{N}$ solid solutions. *Physics of the Solid State*, 47, 1618-1625. <https://doi.org/10.1134/1.2045343>
- Khenata, R., Bouhemadou, A., Sahnoun, M., Reshak, A. H., Baltache, H., & Rabah, M. (2006). Elastic, electronic and optical properties of ZnS, ZnSe and ZnTe under pressure. *Computational Materials Science*, 38(1), 29-38. <https://doi.org/10.1016/j.commatsci.2006.01.013>
- Kleinman, L. (1962). Deformation potentials in silicon. I. Uniaxial strain. *Physical Review*, 128(6), 2614. <https://doi.org/10.1103/PhysRev.128.2614>
- Larbi, M., Riane, R., Matar, S. F., Abdiche, A., Djermouni, M., Ameri, M., Merabet, N., & Oualdine, A. (2016). Abinitio studies of the structural, electronic, and optical properties of quaternary $\text{B}_x\text{Al}_y\text{Ga}_{1-x-y}\text{N}$ compounds. *Zeitschrift für Naturforschung B*, 71(2), 125-134. <https://doi.org/10.1515/znbn-2015-0149>
- Matori, K. A., Zaid, M. H. M., Sidek, H. A. A., Halimah, M. K., Wahab, Z. A., & Sabri, M. G. M. (2010). Influence of ZnO on the ultrasonic velocity and elastic moduli of soda lime silicate glasses. *International Journal of Physical Sciences*, 5(14), 2212-2216. <https://doi.org/10.5897/IJPS.9000313>
- Ponomareva, A. V., Isaev, E. I., Vekilov, Y. K., & Abrikosov, I. A. (2012). Site preference and effect of alloying on elastic properties of ternary B_2NiAl -based alloys. *Physical Review B*, 85(14), 144117. <https://doi.org/10.1103/PhysRevB.85.144117>
- Rai, D. P., Ghimire, M. P., & Thapa, R. K. (2014). A DFT study of BeX ($\text{X} = \text{S}, \text{Se}, \text{Te}$) semiconductor: modified Becke Johnson (mBJ) potential. *Semiconductors*, 48, 1411-1422. <https://doi.org/10.1134/S1063782614110244>
- Riane, R., Boussahla, Z., Matar, S. F., & Zaoui, A. (2008). Structural and electronic properties of zinc blende-type nitrides $\text{B}_x\text{Al}_{1-x}\text{N}$. *Zeitschrift für Naturforschung B*, 63(9), 1069-1076. <https://doi.org/10.1515/znbn-2008-0909>
- Shen, Y., & Zhou, Z. (2008). Structural, electronic, and optical properties of ferroelectric $\text{KTa}_{1/2}\text{Nb}_{1/2}\text{O}_3$ solid solutions. *Journal of Applied Physics*, 103(7). <https://doi.org/10.1063/1.2902433>

- Teles, L. K., Scolfaro, L. M. R., Leite, J. R., Furthmüller, J., & Bechstedt, F. (2002). Spinodal decomposition in $B_xGa_{1-x}N$ and $B_xAl_{1-x}N$ alloys. *Applied Physics Letters*, 80(7), 1177-1179. <https://doi.org/10.1063/1.1450261>
- Troullier, N., & Martins, J. L. (1991). Efficient pseudopotentials for plane-wave calculations. *Physical Review B*, 43(3), 1993. <https://doi.org/10.1103/PhysRevB.43.1993>
- Watanabe, S., Takano, T., Jinen, K., Yamamoto, J., & Kawanishi, H. (2003). Refractive indices of $B_xAl_{1-x}N$ ($x=0-0.012$) and $B_yGa_{1-y}N$ ($y=0-0.023$) epitaxial layers in ultraviolet region. *physica status solidi (c)*, (7), 2691-2694. <https://doi.org/10.1002/pssc.200303549>
- Wooten, F. (1972). *Optical properties of solids*. Academic Press.
- Wu, Z. J., Zhao, E. J., Xiang, H. P., Hao, X. F., Liu, X. J., & Meng, J. (2007). Crystallite structures and elastic properties of superhard Ir N 2 and Ir N 3 from first principles. *Physical Review B*, 76(5), 054115. <https://doi.org/10.1103/PhysRevB.76.054115>
- Zhai, H., Li, X., & Du, J. (2012). First-principles calculations on elasticity and anisotropy of tetragonal tungsten dinitride under pressure. *Materials Transactions*, 53(7), 1247-1251. <https://doi.org/10.2320/matertrans.M2011373>
- Zheng, J. C., Wang, H. Q., Huan, C. H. A., & Wee, A. T. S. (2001). The structural and electronic properties of $(AlN)_x(C_2)_{1-x}$ and $(AlN)_x(BN)_{1-x}$ alloys. *Journal of Physics: Condensed Matter*, 13(22), 5295. <https://doi.org/10.1088/0953-8984/13/22/322>



Forecasting Natural Gas Consumption by User Type Using Machine Learning: A Comparative Study

Süleyman TEKİN¹ Serhat PEKER² Nurhan DUDAKLI^{3*}

¹ İzmir Doğal Gaz Dağıtım A.Ş.

² Department of Management Information Systems, İzmir Bakircay University, İzmir, Türkiye

³ Department of International Trade and Business, İzmir Bakircay University, İzmir, Türkiye

Keywords	Abstract
Natural Gas Consumption Energy Forecasting Machine Learning User-Type Analysis Data Science Hyperparameter Optimization	This study aims to develop user-type-specific machine learning models to forecast natural gas consumption for residential and commercial user groups in İzmir, Turkey. Multiple Linear Regression, Random Forest, LightGBM, and XGBoost algorithms were implemented, and model performance was enhanced through hyperparameter optimization. The models were evaluated using MAE and RMSE metrics. Results indicate that LightGBM and Random Forest provided the most accurate forecasts overall, with LightGBM performing best in the residential group and Random Forest slightly outperforming others in the commercial group. In contrast, MLR underperformed due to the non-linear nature of the data. Residential consumption patterns were found to be more predictable, leading to lower error rates, whereas the commercial group exhibited higher variability and forecast challenges. The study highlights the importance of distinguishing user types and employing well-tuned machine learning algorithms for improved energy demand forecasting.
Cite	
Tekin, S., Peker, S., & Dudaklı, N. (2025). Forecasting Natural Gas Consumption by User Type Using Machine Learning: A Comparative Study. <i>GU J Sci, Part A</i> , 12(2), 632-651. doi:10.54287/guj.1689278	
Author ID (ORCID Number)	Article Process
0009-0009-4806-0671	Submission Date 02.05.2025
0000-0002-6876-3982	Revision Date 26.05.2025
0000-0002-5593-5335	Accepted Date 23.06.2025
	Published Date 30.06.2025

1. INTRODUCTION

In today's context, guided by the goals of sustainable development, energy management holds strategic importance for the efficient use of resources, reduction of environmental impacts, and ensuring energy security. Rapidly growing demand for energy due to increasing population, industrialization, and urbanization necessitates the development of more planned, efficient, and flexible energy systems (Smith et al., 2020). In particular, accurate and reliable forecasting of energy demand is critical to reducing dependency on fossil fuels and facilitating successful transitions to clean energy sources (Johnson et al., 2021).

Natural gas is a key energy source worldwide due to its economic and environmental advantages. In developing countries like Turkey, natural gas is widely used across residential, commercial, and industrial sectors. In residential settings, it is primarily consumed for heating, cooking, and hot water; whereas in commercial and industrial sectors, it is used in production processes, steam systems, and HVAC applications (Brown & Zhang, 2019). This diversity in user types leads to significantly different consumption patterns for each group.

*Corresponding Author, e-mail: nurhan.dudakli@bakircay.edu.tr

Natural gas demand is influenced by a variety of factors, including seasonal variations, ambient temperature, weekday/weekend effects, economic growth, energy prices, and population movement (Yu et al., 2014; Harold et al., 2015). Therefore, accurate demand forecasting plays a critical role not only from technical and economic perspectives but also for infrastructure planning, supply security, and cost optimization (Erdogdu, 2010; Jones, 2019). Particularly, short-term forecasts on the city or regional level provide indispensable insights for gas distribution companies in operational planning, inventory management, and emergency preparedness (Thornton et al., 2016).

Most studies on energy demand forecasting have focused on statistical models based on historical data (e.g., ARIMA, SARIMA, Box-Jenkins) and, more recently, on machine learning models (e.g., MLR, Random Forest, LightGBM, XGBoost, LSTM, CNN-LSTM) (Zha et al., 2022; Qin et al., 2023). Machine learning algorithms, in particular, offer high accuracy due to their ability to handle large, complex datasets and model non-linear relationships (Ke et al., 2017; Garcia et al., 2020).

The primary aim of this study is to model natural gas consumption in İzmir, Turkey's third largest city, by developing user-type-specific forecasting models for residential and commercial groups. To this end, Multiple Linear Regression (MLR), LightGBM, Random Forest, and XGBoost algorithms were applied, and their predictive performance was compared using MAE and RMSE metrics. Furthermore, model performance was enhanced through hyperparameter optimization, and the influence of different user group consumption patterns was analyzed.

This study offers both practical and theoretical contributions. Practically, it supports natural gas distribution companies in producing more accurate and timely forecasts by user group, thus enabling more effective decision-making in supply-demand planning. Theoretically, it compares the performance of different machine learning algorithms in modeling consumption patterns across user types and analyzes the effects of hyperparameter tuning on forecasting accuracy. As one of the few studies employing user-type-specific models with hyperparameter optimization, it provides a scalable methodology applicable to similar forecasting problems.

2. LITERATURE REVIEW

In recent years, there has been a significant increase in academic research on natural gas consumption forecasting, both in Turkey and globally. These studies, particularly those published between 2012 and 2024, mainly evaluate the impact of various modeling techniques on forecast accuracy.

2.1. Forecasting Studies in the Natural Gas Sector

In international literature, natural gas demand forecasting has become a vital area of research, especially in developing countries, for ensuring energy security, planning infrastructure investments, and forming energy policies. Early studies in this field predominantly employed econometric models, time series analyses, and

panel data approaches. However, in recent years, AI- and ML-based hybrid models have increasingly gained traction. For instance, in China, Ji et al. (2018) confirmed a long-term equilibrium relationship between income and price variables through econometric analysis. Shaikh and Ji (2016), using a logistic model, projected a significant increase in natural gas demand and import dependency by 2035. Yu et al. (2014), through a panel data analysis, found that income and price elasticities varied across regions. Similar trends were observed in countries like Bangladesh and Pakistan. Wadud et al. (2011) emphasized that income growth was a key determinant of demand in Bangladesh, whereas price changes had limited influence. Khan (2015) reported that in Pakistan, the electricity sector was the main driver of natural gas consumption, while the effect of prices was marginal.

Important findings also emerged from micro-level studies in Europe. In Ireland, Harold et al. (2015) analyzed the impact of building type, income level, and weather on natural gas consumption and found that temperature changes were the most significant factor. Thornton et al. (2016) statistically demonstrated a negative correlation between temperature and gas consumption in the UK, with demand increasing by 3–4% during cold winter days. A detailed comparison of these studies is provided in Table 1, summarizing methods, forecast periods, time resolutions, modeling techniques, and key findings. Overall, the international literature indicates that natural gas demand shows low price elasticity but high income sensitivity. Furthermore, exogenous environmental factors such as temperature and user type significantly influence consumption patterns and forecast accuracy.

Interest in machine learning and AI-based methods for natural gas forecasting has risen considerably in recent years. Alongside traditional models, advanced approaches such as Artificial Neural Networks (ANN), Support Vector Regression (SVR), CNN-LSTM, wavelet transformations, federated learning, whale optimization, and grey forecasting have been employed. For example, Beyca et al. (2019) reported that the SVR algorithm yielded the best results in daily gas demand forecasting in Istanbul. Su et al. (2019) significantly reduced forecast errors using a hybrid model that combined Bi-LSTM, genetic algorithms, and wavelet transformations on hourly data. Qin et al. (2023) demonstrated a 16.5% improvement in MAE using a federated contrastive learning approach over standard LSTM.

Similarly, Zha et al. (2022) showed that a CNN-LSTM-based model achieved a low MAPE of 7.7% in production-based forecasts, outperforming RNN (18%), Random Forest (23.2%), and traditional LSTM (46%). Qiao et al. (2020) developed a Volterra filtering-based approach that delivered high accuracy in short-term hourly forecasts. In some cases, traditional regression-based models also achieved comparable accuracy. Palinski (2019) found no significant difference between ANN, decision trees, and regression methods. In a Turkish study, Es (2021) reported that grey models performed better than SARIMA.

The variables most influential to forecast accuracy include temperature, historical consumption, prices, population, industrial output, and heating/cooling degree days (HDD/CDD). Moreover, temporal resolution

(hourly, daily, monthly) and differences among user groups also significantly impact model performance. These findings are summarized in Table 2, which compares ML-based methods in terms of input variables, forecast range, and performance advantages. In conclusion, these studies show that well-structured ML/AI-based models provide lower errors and higher forecasting accuracy than traditional methods. However, model performance sustainability depends on factors like data quality, input diversity, and hyperparameter tuning.

2.2. Natural Gas Demand Forecasting Studies in Turkey

Forecasting natural gas consumption is of critical importance for energy supply security, cost optimization, and infrastructure planning, particularly in the context of Turkey. Numerous academic studies conducted both in Turkey and globally have applied various modeling approaches, timeframes, data types, and performance metrics to achieve accurate forecasts.

One notable study by Erdogdu (2010) used the ARIMA model to analyze the price and income elasticities of natural gas demand in Turkey, concluding that price increases had a limited impact on demand. Similarly, Ozdemir et al. (2016) developed a hybrid regression model using Genetic Algorithm (GA) and Simulated Annealing (SA), achieving lower error rates than traditional methods. Early studies in literature predominantly relied on statistical methods. For instance, Dalkılıç and Topal (2014) performed monthly natural gas demand forecasts for Elazığ province using ARIMA and Box-Jenkins models. Tuna (2019) reported that the SARIMA model successfully captured seasonal variations in natural gas demand for Erzurum. Oruç and Çelik (2017) found that the Box-Jenkins method provided a high statistical fit in Isparta, while Akpınar and Yumuşak (2020) identified SARIMA as the best-performing model in Sakarya, with a MAPE of 11.32% and R^2 of 0.981.

With increasing data volume and complexity over time, machine learning (ML) methods have gained popularity. Yıldız (2015) compared ANN, Grey Forecasting, and Box-Jenkins models for Kayseri, finding that the ANN model produced the lowest error. Ayaz and Kılıç (2018) achieved $R^2 = 0.98$ using ANN in a daily prediction study conducted in Denmark. Ay (2018) used ANN for Yozgat and identified a consistent upward trend in monthly consumption. More recently, Balıkcı et al. (2024) demonstrated the applicability of ANN-based models for highly accurate daily and hourly forecasts in Istanbul, supporting infrastructure planning at city gate stations.

Broader applications of machine learning across Turkey have led to national-scale forecasts. Eren and Kaçtıoğlu (2017) predicted natural gas consumption in Turkey through 2030 using a Grey Forecasting model, estimating a 2.5-fold increase in per capita usage. Aydın et al. (2022) modeled national consumption using the MARS algorithm and found that increased use of renewable energy contributed to reduced demand. Kılınç and Çoban (2017) applied panel data analysis across OECD countries and identified a price elasticity of demand for natural gas at approximately 0.10. A comprehensive summary of these studies is provided in Table 3, comparing them in terms of methods, forecasting periods, time resolutions, input variables, and performance outcomes.

Table 1. Statistical and Econometric Approaches to Natural Gas Demand Forecasting

Source	Country/Region	Method	ML/Statistical	Forecast Horizon	Time Resolution	Input Variables	Model Performance Description
Erdogdu (2010)	Turkey	ARIMA	Statistical	Until 2030	Annual	Price, income	Comparative accuracy assessment
Ji et al. (2018)	China	Econometric model	Statistical	Until 2030	Annual	Sectoral income, fuel prices	Scenario-based projections
Shaikh and Ji (2016)	China	Logistic Model + LMA	Statistical	2020–2035	Annual	Population, energy policies, historical consumption	Forecast range between 330–590 bcm
Ozdemir et al. (2016)	Turkey	Hybrid GA-SA + Regression	Hybrid	2010–2030	Annual	GDP, population, growth rate	Lower error than traditional regression via GA-SA
Wadud et al. (2011)	Bangladesh	Dynamic econometric model	Statistical	Until 2025	Annual	GDP, prices	High income elasticity, low price responsiveness
Khan (2015)	Pakistan	Econometric model + Scenario	Statistical	2012–2020	Annual	GDP, prices, sectors	Sectoral simulation approach
Yu et al. (2014)	Cities in China	Panel data - FGLS	Statistical	2006–2009	Daily	Price, income, temperature, pipeline access	Pronounced price elasticity
Harold et al. (2015)	Ireland	Micro-econometric model	Statistical	2009–2011	Daily	Weather, building features, socio-economic variables	Demand responsiveness analyzed
Thornton et al. (2016)	United Kingdom	Statistical analysis	Statistical	1975–2013	Daily	Temperature, seasonality	Strong correlation: $r = -0.94$
Es (2021)	Turkey	Grey Model + SARIMA	Statistical	2010–2025	Monthly	Historical demand	New grey model outperformed SARIMA
Joshi (2021)	USA	Fixed-effects panel regression	Statistical	2011–2015	Monthly	Price, HDD, CDD	Low price elasticity; weather conditions were critical factors

Table 2. Machine Learning and AI-Based Forecasting Methods for Natural Gas Consumption

Source	Location	Method	ML/Statistical	Forecast Horizon	Time Resolution	Input Variables	Model Performance Description
Beyca et al. (2019)	Turkey – Istanbul	MLR, ANN, SVR	Machine Learning	Short-term	Daily	Historical consumption data	SVR yielded the best results
Su et al. (2019)	China	Wavelet + Bi-LSTM + GA	Machine Learning	Short-term	Hourly	Gas load data, temperature, etc.	Lower error rates compared to existing models
Qin et al. (2023)	China (11 cities)	Federated Contrastive Learning + LSTM	Machine Learning	Medium-term (3–12 mo)	Monthly	Historical consumption, clustered patterns	MAE improved by 16.52% compared to classic LSTM
Palinski (2019)	Poland	ANN, Decision Tree, Regression	Machine Learning	Medium-term	Monthly	EU countries' gas consumption, temperature, income, industrial data	ML was effective, but results similar to classical regression
Zha et al. (2022)	China	CNN-LSTM + Bagging + PURPS	Machine Learning	Medium-term	Monthly	Production history, forecast parameters	MAPE: CNN-LSTM (7.7%) < RNN (18%), RF (23.2%), LSTM (46%)
Qiao et al. (2020)	China	Volterra Filter + Whale Optimization	Machine Learning	Short-term	Hourly	Chaotic time series, temperature, etc.	Outperformed many other models

Table 3. Summary of Natural Gas Demand Forecasting Studies Conducted in Turkey

Source	Scope	Location	Method	ML/Statistical	Forecast Period	Time Resolution	Input Variables	Model Performance Description
Eren and Kaçtıoğlu (2017)	National	Turkey	Grey Forecasting Model (GM(1,1))	Statistical	2017–2030	Annual	Population, gas consumption	MAPE < 10% (high accuracy)
Ay (2018)	Local	Yozgat, Turkey	Artificial Neural Network (ANN), Mann-Kendall, Şen Test	Machine Learning	2006–2016	Monthly	Historical consumption data	High correlation, low error
Aydın et al. (2022)	National	Turkey	MARS (Multivariate Adaptive Regression Splines)	Machine Learning	2015–2021	Monthly	Climate, renewable energy share, macroeconomic indicators	Good fit, strong accuracy
Dalkılıç and Topal (2014)	Local	Elazığ, Turkey	ARIMA, Box-Jenkins	Statistical	2008–2013	Monthly	Historical consumption	Moderate error
Tuna (2019)	Local	Erzurum, Turkey	SARIMA, ARIMA, Logit	Statistical	2012–2018	Monthly	Consumption, seasonality	Low error with SARIMA
Oruç and Çelik (2017)	Local	Isparta, Turkey	Box-Jenkins, Regression	Statistical	2011–2015	Monthly	Historical consumption	High fit, $R^2 > 0.95$
Yıldız (2015)	Local	Kayseri, Turkey	Box-Jenkins, Grey Forecasting, ANN	Hybrid (ML + Stat.)	2006–2013	Monthly	Historical data, temperature	ANN yielded the lowest error
Ayaz and Kılıç (2018)	International	Denmark	ANN	Machine Learning	2010–2016	Daily	Temperature, calendar data	Low RMSE, high R^2
Kılınç and Çoban (2017)	International	OECD Countries	Panel Data Analysis	Statistical	1990–2014	Annual	Energy prices, GDP	High model fit coefficient
Aydın et al. (2022)	National	Turkey	MARS	Machine Learning	2015–2021	Monthly	Macroeconomic variables, temperature, renewables	Good fit, reliable accuracy
Akpınar and Yumuşak (2020)	Local	Sakarya, Turkey	ARIMA, SARIMA, Holt-Winters, TSD	Statistical	2014	Daily and Monthly	Univariate historical consumption	Best result with SARIMA (MAPE 11.32%, $R^2=0.981$)
Balıkçı et al. (2024)	Local	Istanbul, Turkey	Artificial Neural Networks (ANN)	Machine Learning	2008–2027	Daily and Hourly	Number of users, temperature, holidays, IGD	$R^2 = 0.98$
Present Study (2025)	Local	Izmir, Turkey	MLR, Random Forest, LightGBM, XGBoost	Machine Learning	2019–2023	Daily	Temperature, user type, day info, season	Lowest MAE/RMSE with LightGBM and RF

2.3. General Assessment and Contributions of This Study

The reviewed literature reveals that natural gas demand forecasting can be conducted at different geographic scales—national, regional, and city-level. In Turkey, studies conducted in cities like Istanbul (Balıkçı et al., 2024), Erzurum (Tuna, 2019), and Bayburt (Şengün, 2012) demonstrate that seasonal and daily consumption patterns can be effectively analyzed using time series-based models. High-accuracy predictions have been achieved using Box-Jenkins, SARIMA, and ARIMA models, and geographic scale has been shown to directly influence model selection and performance.

In recent years, machine learning and artificial intelligence approaches have gained prominence alongside traditional statistical methods. Algorithms like Random Forest, XGBoost, and LightGBM are frequently used in the literature for their ability to model non-linear structures and deliver high predictive accuracy. Hyperparameter optimization is a key technique that enhances the forecasting performance of these models. Garcia et al. (2020), for example, demonstrated that optimized Random Forest and LightGBM models significantly reduced error rates. Deep learning algorithms (e.g., LSTM, CNN, RNN) have also proven highly effective in short-term, high-frequency forecasting scenarios. However, their computational cost and preprocessing requirements are often noted as limitations (Evcim, 2019; Wang et al., 2021).

Because natural gas consumption patterns vary by user type, modeling residential and commercial consumers separately improves forecasting accuracy. Studies by Chen et al. (2018) and Aydemir and Kılıç (2020) showed that user-type-specific models yield more accurate results. However, literature systematically addressing such user group differentiation remains limited, with many studies focusing on only a single user type—thereby restricting overall model effectiveness.

In summary, existing literature contributes significantly to comparing different algorithms, integrating time series and machine learning approaches, and applying them across various scales. However, there is a lack of comprehensive studies that incorporate user-type segmentation, hyperparameter optimization, and model comparisons in a systematic manner. In dynamic consumption environments like Turkey, further research demonstrating the applicability of such models is urgently needed.

Unlike most prior studies that focus on national-level forecasting, aggregate consumption data, or single user types, our study introduces a novel, user-type-specific modeling framework applied to city-level data. It combines machine learning-based prediction with hyperparameter optimization for both residential and commercial users—an approach that is rarely integrated in earlier research. In this regard, the contributions of the present study are multifaceted. First, it evaluates residential and commercial user groups with separate model structures, allowing for a detailed analysis of their distinct consumption patterns. Second, it enhances model performance through hyperparameter optimization of modern machine learning algorithms, including MLR, Random Forest, LightGBM, and XGBoost. Third, it presents a regional energy demand forecasting framework specific to İzmir, generating insights useful for local decision-making. Finally, the scalable and

data-driven methodology proposed in this study offers a practical forecasting strategy to support energy policy and management processes.

3. METHODOLOGY

In this study, a comprehensive data analysis and machine learning-based methodological approach was adopted to forecast natural gas consumption in the province of İzmir, Turkey. The primary aim of the study is to develop user-type-specific prediction models for residential and commercial consumers, thereby supporting energy management decision-making processes.

In the first stage, datasets including historical consumption data, prices, weather conditions, and population—variables that enable the analysis of environmental, economic, and technological factors affecting natural gas usage—were collected. Next, the collected data were cleaned by identifying and removing missing values, outliers, and inconsistencies. Descriptive statistics were then calculated to evaluate the data structure, distribution, and inter-variable relationships. In the third stage, forecasting methods appropriate to the nature of the data and the study objectives were selected. In addition to the classical Multiple Linear Regression (MLR) approach based on statistical modeling, modern machine learning algorithms such as LightGBM, Random Forest, and XGBoost were employed. These algorithms were chosen for their ability to capture different data patterns and model non-linear relationships effectively. Finally, the developed prediction models were implemented on training and test datasets, and their performance was evaluated using two key error metrics: Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). A step-by-step overview of this methodological process is visually presented in Figure 1.

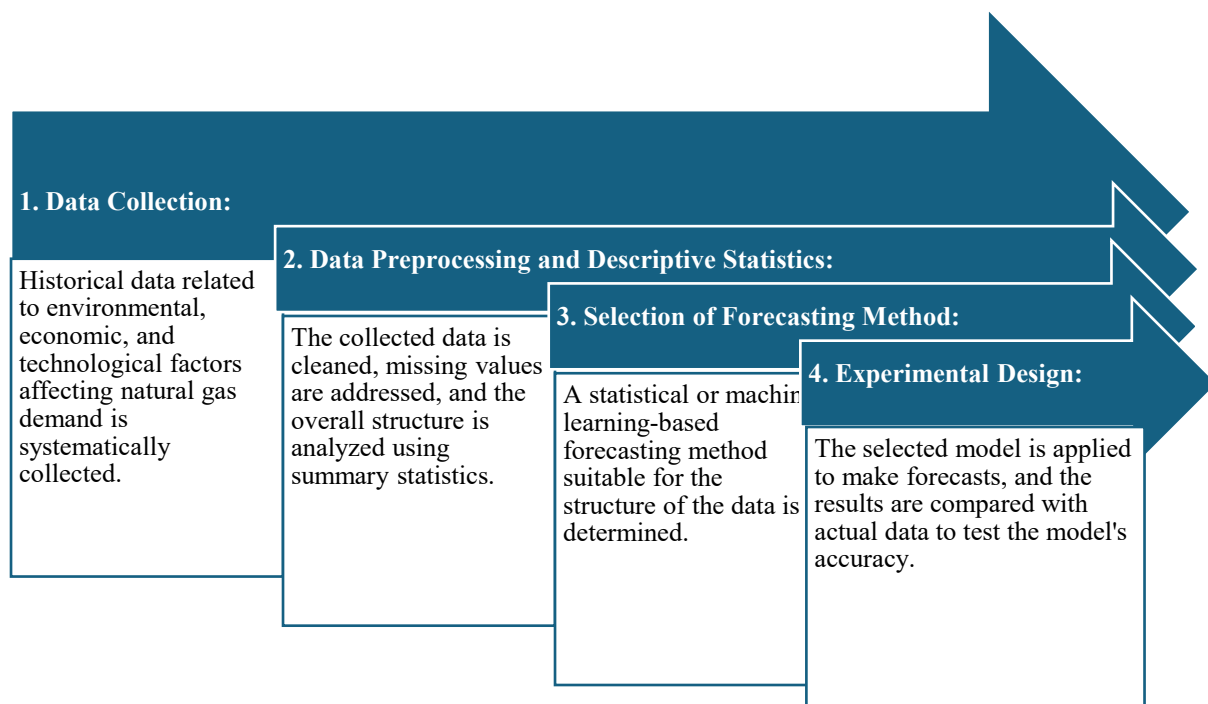


Figure 1. Natural Gas Demand Forecasting Process and Main Stages

The dataset used in this study includes natural gas consumption information representing residential and commercial user groups in İzmir. Separate models have been developed for each user group by segmenting the data based on numerical and categorical attributes. The selected dataset, data preprocessing steps, preferred forecasting methods, and experimental design process are discussed in detail in the following sections of the study.

3.1. Dataset

In this study, two separate datasets were used to forecast natural gas consumption across İzmir province, based on residential and commercial user types. The dataset includes historical natural gas consumption records of residential and commercial subscribers in İzmir. For the purpose of conducting user-type-specific analyses, the dataset was divided into residential and commercial groups.

The dataset includes the following features: *Season*, which represents the time of year during which consumption occurs; *User Type*, indicating whether the consumer is residential or commercial; *District Location*, categorizing regions based on temperature and geographic differences; *Number of Days*, denoting the duration (in days) over which the consumption was recorded; and *Consumption*, representing the volume of natural gas consumed in cubic meters (m³). The details and data types of these features are presented in Table 4.

Table 4. Dataset Features

Feature	Value	Type	Description
Season	{"1": "Spring", "2": "Summer", "3": "Autumn", "4": "Winter"}	Nominal	Indicates the season during which consumption occurred.
User Groups	{"1": "Residential", "2": "Commercial"}	Nominal	Specifies the type of natural gas user.
District Zone	{"1": 'Bergama', 'Kınık', 'Dikili', 'Aliağa', 'Foça', 'Menemen', 'Karşıyaka', 'Çiğli', 'Bayraklı', 'Bornova'}, {"2": 'Bayındır', 'Ödemiş', 'Kiraz', 'Beydağ', 'Tire', 'Kemalpaşa', 'Buca', 'Gazimur', 'Torbalı'}, {"3": 'Menderes', 'Seferihisar', 'Güzelbahçe', 'Urla', 'Narlıdere', 'Balçova', 'Karabağlar', 'Konak', 'Selçuk'}	Nominal	Represents the regional grouping of districts based on average temperature.
Number of Days	Real	Numeric	Indicates the number of days covered by the natural gas consumption record.
Consumption	Real	Numeric	Refers to the amount of natural gas consumed during a specific period (in cubic meters - m ³).

3.2. Data Preprocessing and Descriptive Statistics

A comprehensive data preprocessing procedure was implemented to make the dataset suitable for analysis. The dataset, which covers the natural gas consumption of residential and commercial users in İzmir, has a

multidimensional structure that required detailed examination. Several steps were applied to minimize the impact of missing and outlier values, optimize data types to reduce processing costs, and streamline the modeling process.

The datasets were refined according to the study's objectives. In the commercial dataset, 2,894 records containing missing or erroneous information were excluded from analysis. These excluded records were inconsistent with consumption continuity, misclassified in terms of user type, or displayed interregional inconsistencies. Similarly, 1,168,816 records were cleaned from the residential dataset, resulting in a clean and reliable dataset. Special attention was paid to records with high consumption variation or missing values.

The Tukey Fence method was used to remove missing values and detect anomalous consumption patterns. Values falling outside the standard deviation thresholds were excluded to obtain a more consistent dataset. Additionally, semantically insignificant subcategories within categorical features were merged to create a more coherent structure. As a result, the commercial dataset was reduced from 537,638 to 534,744 records, and the residential dataset was reduced from 43,707,706 to 43,590,890 records.

Nominal variables (*Season, User Type, District Location*) were converted to binary format using One Hot Encoding, making them suitable for modeling. In addition, all features other than the target variable were normalized to the [0, 1] range using the min-max scaling technique. Missing values were rechecked to ensure a clean dataset. Descriptive statistics were calculated for the numerical features. These include minimum, maximum, median, mean, and standard deviation values. It was found that the average daily consumption in the residential group was 57.29 m³, whereas in the commercial group, this value reached 493.24 m³, indicating a significantly higher variation. Detailed statistics for numerical variables are presented in Tables 5 and 6.

Table 5. *Descriptive Statistics for the Residential User Group*

Feature	Minimum	Maximum	Median	Mean	Standard Deviation
Number of Days	1	136	30	30.94	7.02
Consumption	0	847	19	57.29	77.08

Table 6: *Descriptive Statistics for the Commercial User Group*

Feature	Minimum	Maximum	Median	Mean	Standard Deviation
Number of Days	1	212	30	31.74	11.19
Consumption	0	22,368	84	493.24	1,446.88

The distribution of nominal variables was analyzed by calculating the frequency and percentage of each category. For instance, within the Season variable, "Autumn" exhibited the highest consumption rate, while the commercial user group demonstrated a wider variation in consumption patterns. Regarding the User Type variable, the "Residential" category accounted for the majority of total consumption.

This data preprocessing and analysis phase enabled the development of more accurate consumption forecasting models by considering differences between user types. Customized modeling approaches were adopted for both user groups, resulting in reliable outputs that can support energy management processes. Detailed statistical information for the categorical variables is provided in Tables 7 and 8.

Table 7. *Distribution of Categorical Variables for the Residential User Group*

Feature Name	Category	Count	Percentage (%)
Season	1	10,706,418	24.56
	2	11,227,187	25.76
	3	11,575,879	26.56
	4	10,081,406	23.13
District Zone	1	19,220,088	44.09
	2	13,885,792	31.85
	3	10,485,010	24.05

Table 8. *Distribution of Categorical Variables for the Commercial User Group*

Feature Name	Category	Count	Percentage (%)
Season	1	131,815	24.65
	2	136,244	25.48
	3	143,613	26.86
	4	123,072	23.02
District Zone	1	238,656	44.63
	2	155,686	29.11
	3	140,402	26.26

3.3. Machine Learning Methods

In this study, various machine learning algorithms were applied to forecast natural gas consumption. Four different algorithms were used to construct forecasting models: Multiple Linear Regression (MLR), LightGBM, Random Forest, and XGBoost. Each algorithm was carefully selected to capture different data patterns and provide the most accurate predictions based on user type.

MLR is one of the most fundamental statistical methods used to model the linear relationship between independent and dependent variables. This method assumes a linear function while analyzing how variables such as *Season* or *Number of Days* affect consumption. However, its performance may be limited when applied

to complex and nonlinear datasets. MLR is commonly used in energy forecasting and served as a baseline comparison model in this study (Kutner et al., 1998).

LightGBM is a fast, optimized gradient boosting algorithm. It is often preferred for complex data structures like energy forecasting due to its speed and lightweight nature in handling large datasets. Notably, it excels at efficiently processing categorical variables (Ke et al., 2017; Brown & Zhang, 2019).

Random Forest is an ensemble learning method that combines multiple decision trees to improve accuracy. It is particularly effective at modeling nonlinear relationships between independent and dependent variables. One of its key strengths is its ability to prevent overfitting by averaging the outputs of multiple trees (Breiman, 2001).

XGBoost is a widely used boosting algorithm known for its strong performance on large datasets. It is frequently applied in energy forecasting due to its ability to combine multiple optimization techniques to enhance accuracy while reducing computational time (Chen & Guestrin, 2016; Garcia et al., 2020).

Each of these methods offers unique advantages in modeling the multidimensional and complex relationships present in the dataset. The performance of each model was evaluated separately to assess prediction accuracy across user types.

3.4. Experimental Design

All experiments in this study were conducted using Python and relevant libraries including Dask, Scikit-learn, LightGBM, and XGBoost. The forecasting models were applied to datasets representing residential and commercial user groups in İzmir. These datasets featured a multidimensional structure, comprising independent variables such as *Season*, *Number of Days*, *District Location*, and *User Type*, with *Consumption* as the dependent variable.

For model training and evaluation, the data was split into training and testing sets using an 80%-20% ratio. This division allowed the machine learning algorithms to learn relationships from the training set while testing their generalization capabilities on unseen data. A randomized train-test split was preferred in this study due to the large volume and variability of the dataset, which ensured representative distribution in both sets. To improve the performance of each model, RandomizedSearchCV was chosen for hyperparameter tuning due to its efficiency in exploring a wide parameter space with limited computational cost. Compared to GridSearchCV, which exhaustively evaluates all parameter combinations, RandomizedSearchCV samples a fixed number of parameter settings from specified distributions, offering a practical trade-off between tuning quality and computational time. For algorithms such as LightGBM and XGBoost, parameters like "n_estimators", "max_depth", and "learning_rate" were fine-tuned using RandomizedSearchCV. The optimization process played a crucial role in managing training time efficiently while increasing accuracy

rates. Similarly, the impact of decision variable selection and configuration on optimization performance has been extensively analyzed in engineering problems such as infrastructure systems (Çetin et al., 2025).

The main objective of the experiments was to assess the impact of user groups on model performance and to analyze how these differences affect forecasting accuracy. Residential and commercial user groups were evaluated separately, and each model's performance within each group was compared. Results showed that LightGBM and Random Forest models consistently performed well across both user types, whereas MLR was less effective in capturing variation among user groups.

3.5. Performance Evaluation Metrics

Two primary metrics were used to assess the performance of the machine learning models: Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). These metrics quantify the discrepancy between the model's predicted consumption values and the actual values.

MAE quantifies the average magnitude of prediction errors by measuring the absolute differences between the predicted and actual values. It treats all deviations uniformly, regardless of direction, making it a simple and intuitive indicator of model accuracy. A smaller MAE value suggests more precise forecasts. The MAE is calculated as shown in the following equation.

$$\text{MAE} = \frac{1}{n} \sum_{t=1}^n (|E_t|) \quad (1)$$

The Root Mean Squared Error (RMSE) offers a more sensitive evaluation of prediction accuracy by giving greater weight to larger errors. It is computed as the square root of the average squared differences between predicted and actual values. A lower RMSE indicates better predictive performance. The formula for RMSE is provided in the following equation.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{t=1}^n (E_t)^2} \quad (2)$$

Both metrics were employed in this study to evaluate the effectiveness of the prediction models applied to both residential and commercial user groups. The comparison of MAE and RMSE values helped to identify which model performed better for each user type, and the results were interpreted accordingly.

4. RESULTS

This section presents the empirical results obtained from the application of machine learning algorithms to forecast natural gas consumption for residential and commercial user groups in İzmir. Each user group exhibits distinct consumption characteristics and variability patterns, necessitating a disaggregated analysis approach.

The evaluation focuses on comparing the predictive accuracy of Multiple Linear Regression (MLR), Random Forest, LightGBM, and XGBoost algorithms using two key performance metrics: Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). These metrics are used to assess the effectiveness of each model in capturing the underlying consumption patterns across different user types. The findings aim to identify which algorithms perform best under varying consumption behaviors and to determine the impact of hyperparameter tuning on overall model accuracy. Results are presented and discussed in detail under separate sub-sections for the residential and commercial user groups.

4.1. Residential User Group

The performance results of the machine learning models for the residential user group are presented in Table 9. LightGBM and Random Forest models delivered the best results in terms of MAE and RMSE, standing out in prediction accuracy for the residential user group. The LightGBM model achieved the lowest error rates with an MAE of 36.95 and an RMSE of 59.98, demonstrating superior performance through the optimization of hyperparameters such as learning rate and maximum depth. Similarly, the Random Forest model provided effective results with an MAE of 36.90 and an RMSE of 60.04. Increasing the number of trees (`n_estimators`) and optimizing the minimum number of samples required for a split (`min_samples_split`) enhanced the model's generalization capability.

Table 9. Residential User Group Forecasting Performance Comparison

Model	MAE	RMSE	Selected Parameters
MLR	52.76	72.07	-
LightGBM	36.95	59.98	{'n_estimators': 100, 'max_depth': 15, 'learning_rate': 0.05}
Random Forest	36.90	60.04	{'n_estimators': 200, 'min_samples_split': 10, 'max_depth': 10}
XGBoost	37.67	60.82	{'n_estimators': 50, 'max_depth': 10, 'learning_rate': 0.05}

XGBoost demonstrated a performance close to that of LightGBM and Random Forest, with an MAE of 37.67 and an RMSE of 60.82. However, due to its slightly higher error rates, it ranked behind these two models. In contrast, the MLR model, based on linear assumptions, yielded an MAE of 52.76 and an RMSE of 72.07, making it the least effective among the tested models. These findings highlight the superior performance of non-parametric models like LightGBM and Random Forest in forecasting residential natural gas consumption.

4.2. Commercial User Group

The performance results of machine learning models developed for the commercial user group are presented in Table 10. These models were evaluated using various metrics to assess their effectiveness in forecasting natural gas consumption patterns specific to commercial users.

Table 10. Commercial User Group Forecasting Performance Comparison

Model	MAE	RMSE	Selected Parameters
MLR	627.90	1423.07	-
LightGBM	614.50	1408.08	{'n_estimators': 100, 'max_depth': 15, 'learning_rate': 0.05}
Random Forest	614.27	1408.50	{'n_estimators': 200, 'min_samples_split': 10, 'max_depth': 10}
XGBoost	617.67	1408.61	{'n_estimators': 200, 'max_depth': 10, 'learning_rate': 0.05}

In the commercial user group, the LightGBM and Random Forest models stood out in terms of MAE and RMSE metrics. The LightGBM model achieved strong performance with MAE of 614.50 and RMSE of 1408.08. Hyperparameter optimization, particularly tuning the learning rate and maximum depth, played a significant role in achieving this performance. The Random Forest model showed comparable success with MAE of 614.27 and RMSE of 1408.34. Optimizing the minimum sample split enabled the model to effectively capture wide consumption variations.

The XGBoost model yielded competitive results with MAE of 614.66 and RMSE of 1408.61 but lagged slightly behind LightGBM and Random Forest due to slightly higher error rates. The MLR model, on the other hand, struggled to capture the complex consumption patterns of the commercial group, resulting in relatively poor performance with MAE of 627.89 and RMSE of 1423.07. These findings indicate that non-parametric models are more effective for forecasting consumption in the commercial user group. The higher error rates observed in the commercial group compared to the residential group are attributed to wider consumption variations. Despite these challenges, LightGBM and Random Forest still delivered effective forecasting results.

5. CONCLUSION

This study aimed to develop user-type-specific machine learning models to forecast natural gas consumption for residential and commercial users in İzmir. Multiple Linear Regression (MLR), Random Forest, LightGBM, and XGBoost algorithms were implemented. Each model was improved through hyperparameter optimization and evaluated using MAE and RMSE metrics. The key findings are:

- LightGBM and Random Forest algorithms delivered the lowest error rates across both user types, demonstrating the highest forecasting accuracy.
- XGBoost produced competitive results but ranked slightly behind the top two models in terms of errors.
- MLR underperformed, especially in the commercial user group, due to the non-linear nature of the data.

User-level analyses showed that residential consumption patterns were more predictable, contributing to better model performance. In contrast, the commercial group exhibited higher variability, leading to greater forecasting challenges. This demonstrates that non-parametric models are more capable of effectively modeling user-type differences.

Moreover, hyperparameter optimization significantly improved model accuracy—careful tuning of learning rate, maximum depth, and number of trees was shown to enhance performance. The study's contributions can be summarized as follows:

- Separate modeling of residential and commercial user groups allowed for a detailed analysis of consumption differences.
- Modern machine learning algorithms were integrated with hyperparameter tuning to develop user-specific models.
- The İzmir case study provided actionable insights at the regional level for decision-making in energy planning.

This study, while offering valuable insights into user-type-specific natural gas consumption forecasting, has several limitations that suggest directions for future research. First, although İzmir provides a diverse user base and regional representation, the findings may not be directly generalizable to other regions with different climatic, socio-economic, or infrastructural characteristics. Future studies should apply and validate the proposed models in different cities or countries to assess their geographic scalability and adaptability. Second, the models lack real-time data integration, hindering responsiveness to dynamic changes. In the future, incorporating live data streams into models could improve forecast relevance and timeliness in rapidly changing conditions. Finally, the focus of this research was on model accuracy and predictive comparison. Therefore, future studies could explore feature importance analysis to identify the most influential variables and inform decision-makers.

Overall, this study demonstrates that incorporating user-type distinctions and correctly configuring modern machine learning algorithms significantly improves model performance in natural gas consumption forecasting. Customized, data-driven forecasting models offer practical and effective decision-support tools in energy management, and this study contributes meaningfully to the limited body of research focusing on user-specific modeling approaches.

AUTHOR CONTRIBUTIONS

Conceptualization, S.T. and S.P.; methodology, S.T. and S.P.; fieldwork, S.T.; software, S.T.; title, S.P. and N.D.; validation, S.P. and N.D.; formal analysis, S.T. and S.P.; research, S.T., S.P. and N.D.; sources, S.T., S.P. and N.D.; manuscript-original draft, S.T., S.P. and N.D.; manuscript-review and editing, S.P. and N.D.; visualization, S.P. and N.D.; supervision, S.P.; project management, S.P. All authors have read and legally accepted the final version of the article published in the journal.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Akpınar, H., & Yumuşak, S. (2020). Sakarya ilinde doğal gaz talep tahmini: SARIMA modeli uygulaması. *Enerji ve Ekonomi Araştırmaları Dergisi*, 5(1), 23–35.
- Ay, M. (2018). Yozgat ilinde doğal gaz tüketiminde eğilim analizi: ANN uygulaması. *İstatistik ve Uygulamalı Bilimler Dergisi*, 10(1), 45–56.
- Ayaz, H., & Kılıç, A. (2018). Danimarka’da doğal gaz tüketimi tahmini: Yapay sinir ağları ile bir uygulama. *Uluslararası Enerji Ekonomisi ve Politika Dergisi*, 8(4), 100–108.
- Aydemir, S., & Kılıç, M. (2020). Türkiye’de doğal gaz tüketim tahmini için makine öğrenimi yöntemlerinin uygulanması. *Enerji ve Çevre Dergisi*, 29(3), 150–165.
- Aydın, R., Yüksel, S., Silahtaroglu, G., & Dinçer, H. (2022). Doğal gaz tüketiminin MARS algoritmasıyla tahmini ve yenilenebilir enerji ile ilişkisi. *Yapay Zeka ve Enerji Sistemleri Dergisi*, 8(2), 111–125.
- Balıkçı, A., Şahin, E., & Karaca, T. (2024). İstanbul’da şehir giriş istasyonları için doğal gaz tüketimi tahmini: ANN tabanlı bir model. *Enerji Sistemleri Mühendisliği Dergisi*, 12(1), 88–104.
- Beyca, A., Yılmaz, B., & Yıldız, S. (2019). Doğal gaz tüketim tahmini için SVR, MLR ve ANN karşılaştırması. *Mühendislik Bilimleri Dergisi*, 25(2), 157–168.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Brown, T., & Zhang, Y. (2019). The Role of Natural Gas in Emerging Energy Markets. *International Energy Journal*, 17(2), 56–70.
- Çetin, T., Turan, M. E., & Şenol, M. E. (2025). Effects of Decision Variables Selection on Sewer Optimization Problem. *Applied Sciences*, 15(9), 4836. <https://doi.org/10.3390/app15094836>
- Chen, T., & Guestrin, C. (2016). *XGBoost: A scalable tree boosting system*. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 785–794.
- Chen, W., Lin, Y., & Liu, W. (2018). A sectoral comparison of natural gas demand prediction. *Energy Economics*, 71, 365–378.
- Dalkılıç, M., & Topal, M. (2014). Elazığ ilinde doğal gaz talebinin ARIMA modeli ile tahmini. *Enerji ve Çevre Dergisi*, 6(3), 55–63.
- Erdogdu, E. (2010). Natural gas demand in Turkey. *Applied Energy*, 87(1), 211–219.
- Eren, T., & Kaçtıoğlu, S. (2017). Türkiye’deki doğal gaz tüketimi ve gri tahmin methodology tahmin edilmesi. *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 16(31), 23–41.
- Es, F. (2021). SARIMA ve gri model karşılaştırması: Türkiye doğal gaz talep tahmini. *Veri Bilimi ve Enerji Analizi Dergisi*, 3(2), 90–102.
- Evcim, M. (2019). Derin öğrenme ile kısa vadeli enerji tüketimi tahmini. *Elektrik Elektronik ve Bilgisayar Bilimleri Dergisi*, 11(1), 37–47.

- Garcia, L., Lee, M., & Smith, J. (2020). Hyperparameter tuning in energy forecasting models: Empirical evidence from LightGBM and Random Forest. *Energy Informatics*, 3(1), 1–15.
- Harold, J., Lyne, M., & Lyons, S. (2015). Microeconomic determinants of residential gas demand in Ireland. *Energy Policy*, 85, 260–269.
- Ji, Q., Zhou, D., & Geng, J. (2018). Natural gas consumption and economic growth: Empirical evidence from China. *Energy Policy*, 122, 17–28.
- Johnson, M., Lee, S., & Kim, H. (2021). Sustainable Energy Planning and Forecasting. *Journal of Energy Systems*, 14(3), 234–245.
- Jones, R. (2019). Energy Demand Forecasting for Smart Cities. *International Journal of Energy Policy Studies*, 12(2), 133–149.
- Joshi, M. (2021). Panel regression analysis of natural gas demand in the United States. *Energy Reports*, 7, 889–897.
- Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. Y. (2017). LightGBM: A highly efficient gradient boosting decision tree. *Advances in Neural Information Processing Systems*, 30, 3146–3154.
- Khan, M. A. (2015). Demand forecasting for natural gas in Pakistan. *Pakistan Journal of Energy and Environment*, 4(1), 33–44.
- Kılınç, N. Ş., & Çoban, O. (2017). Sanayi sektöründe enerji talep esnekliklerinin tahmini: OECD ülkeleri örneği, *Mehmet Akif Ersoy Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 9(20), 479-496. <https://doi.org/10.20875/makusobed.342575>
- Kutner, M. H., Nachtsheim, C. J., & Neter, J. (1998). *Applied regression analysis* (4th ed.). Wiley.
- Oruç, K., & Çelik, Ş. (2017). Isparta ili için doğal gaz talep tahmini. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 22(1), 31-42.
- Ozdemir, G., Aydemir, E., Olgun, M. O., & Mulbay, Z. (2016). Forecasting of Turkey natural gas demand using a hybrid algorithm. *Energy Sources, Part B: Economics, Planning, and Policy*, 11(4), 295-302.
- Palinski, A. (2019). Artificial neural networks vs. classical regression in energy prediction. *Energy Management Review*, 11(3), 78–91.
- Qiao, X., Jin, L., & Chen, M. (2020). Whale optimization algorithm-based Volterra filter for gas demand prediction. *Expert Systems with Applications*, 150, 113253.
- Qin, X., Zhang, X., & Yu, B. (2023). A federated contrastive learning approach for city-level gas consumption forecasting. *Applied Energy*, 321, 119312.
- Şengün, G. (2012). *Doğal Gaz Talep Tahmini Bayburt İli Üzerine Bir Uygulama*, Yayınlanmamış Yüksek Lisans Tezi, A.Ü. Sosyal Bilimler Enstitüsü, Erzurum.
- Shaikh, F., & Ji, Q. (2016). Forecasting China's natural gas consumption by 2035: Logistic model approach. *Resources Policy*, 49, 292–300.
- Smith, A., Brown, M., & Davis, R. (2020). Strategic energy management for urban resilience. *Energy Research and Social Science*, 68, 101573.

- Su, Y., Li, K., & Zhao, M. (2019). Hybrid short-term gas load forecasting using Bi-LSTM and GA-based wavelet transform. *Energy Procedia*, 158, 4517–4522.
- Thornton, H., Martin, C., & Evans, R. (2016). Modelling temperature effects on residential gas demand in the UK. *Journal of Environmental Economics and Policy*, 5(1), 48–66.
- Tuna, Ç. (2019). Erzurum ilinde doğal gaz talep tahmini: SARIMA modeli ile analiz. MSc Thesis, Atatürk University, Erzurum.
- Wadud, Z., Noman, A. M., & Ahmed, R. (2011). Forecasting gas demand in Bangladesh: An econometric approach. *Energy*, 36(10), 6120–6126.
- Wang, J., Chen, B., & Xu, T. (2021). Deep learning applications in short-term load forecasting. *Energy AI*, 2, 100036.
- Yıldız, C. (2015). Kayseri ilinde doğal gaz tüketimi tahmini: ANN, gri tahmin ve Box-Jenkins karşılaştırması. *Veri Analizi ve Uygulama Araştırmaları Dergisi*, 7(2), 66–78.
- Yu, Y., Yang, S., & Zhang, H. (2014). Regional variation in natural gas demand: Panel data analysis of Chinese cities. *Resources and Energy Economics*, 36(2), 291–310.
- Zha, D., Li, J., & Tang, Y. (2022). CNN-LSTM based hybrid model for natural gas production prediction. *Applied Energy*, 307, 118251.



Utilization of Electromyographic Signal Classification for Predicting Hand Gestures and Grasping Forces Through Artificial Neural Networks

Derya KARABULUT^{1*} Suzan Cansel DOĞRU¹

¹ Istanbul University-Cerrahpaşa, Faculty of Engineering, Mechanical Engineering Department, 34320, Avcılar/Istanbul,, Türkiye

Keywords	Abstract
Electromyography Artificial Neural Network Root Mean Square Classification	Electromyographic (EMG) signals have gained significant recognition for their potential applications in various fields, particularly in the classification of hand gestures and grasping force prediction. These signals provide critical information that can be utilized to develop systems capable of interpreting human intentions, making them indispensable in areas such as assistive technology and human-computer interaction. The primary objective of this study was to assess the effectiveness of different time-domain features extracted from EMG signals in predicting hand gestures and grasping force simultaneously. The study specifically tested features such as Root Mean Square (RMS), Variance of EMG (VAR), Waveform Length (WL), Integrated EMG (IEMG), Difference Absolute Standard Deviation Value (DASDV), and Difference Absolute Mean Value (DAMV), employing an artificial neural network (ANN) to evaluate their ability to predict hand movements and grasping force. EMG data were collected from the flexor carpi radialis muscle, which plays a significant role in hand movement control. After extracting the relevant time-domain features from the EMG signals, these were input into an ANN for further analysis. The results demonstrated that the RMS feature provided the highest accuracy for both prediction of hand gesture (success rate 90.0% for resting position, 93.3% for wrist flexion, and 86.7% for hand pronation) and grasping force (0.09 RMSD and 0.89 PCC values for resting position, 0.15 RMSD and 0.85 PCC for wrist position). These findings underscore the potential application of these EMG-derived features in practical systems, particularly in the development of myoelectric-controlled prostheses. Such advancements hold promise for enhancing the functionality and intuitiveness of prosthetic devices, offering users more efficient and effective control, and thus improving their overall quality of life in assistive technology contexts.

Cite

Karabulut, D., & Doğru, S. C. (2025). Utilization of Electromyographic Signal Classification for Predicting Hand Gestures and Grasping Forces Through Artificial Neural Networks. *GU J Sci, Part A, 12(2)*, 652-664. doi:10.54287/guj.1700850

Author ID (ORCID Number)	Article Process
0000-0002-1903-9525	Derya KARABULUT
0000-0002-6198-0861	Suzan Cansel DOĞRU
	Submission Date 16.05.2025
	Revision Date 21.05.2025
	Accepted Date 23.06.2025
	Published Date 30.06.2025

1. INTRODUCTION

The human hand is a highly functional part of the body, integral to nearly all daily activities. Upper limb amputation significantly impacts the quality of life of amputees (Sims et al., 2020). The loss of an upper limb severely affects work ability and overall well-being, prompting numerous studies on the control of hand gestures and grasping movements for prosthetic hands (Zecca et al., 2002; Jahani Fariman et al., 2016; Kundu et al., 2018; Noce et al., 2019; Parajuli et al., 2019). Consequently, electromyographic (EMG) signals have been commonly employed in the control of upper limb prosthetics (Oskoei & Hu, 2007; Micera et al., 2010; Atzori et al., 2014; Al-Timemy et al., 2015; Onay & Mert, 2020).

*Corresponding Author, e-mail: derya.karabulut@iuc.edu.tr

EMG, which reflects the electrical activation of contracting muscles (Basmajian & De Luca, 1985), has been utilized in diagnosing muscle diseases (Haig et al., 1996), in rehabilitation engineering (Oskoei & Hu, 2007), and for operating various human-machine interfaces (Zecca et al., 2002; Wu et al., 2022). Surface EMG (sEMG) is a well-established method due to its non-invasive nature and ease of application, allowing for the recording of muscle activity from superficial muscles and providing insights into neuromuscular activities (Micera et al., 2010; Ni et al., 2024). sEMG holds promise for reliable hand movement recognition and the provision of detailed movement information (Zhang et al., 2019). However, the classification of EMG signals for prosthetic control remains a challenging task due to their nonlinear characteristics and the dependencies between individuals (Gu et al., 2018). The success of EMG classification relies heavily on two key factors: an appropriate feature extraction method and an effective classifier.

Time-domain feature extraction is commonly used to reduce the dimensionality of EMG data without significantly altering the signal's amplitude-time characteristics (Tkach et al., 2010; Phinyomark et al., 2018). Compared to frequency-domain approaches, time-domain methods offer efficient computation and ease of implementation, making them suitable for myoelectric-controlled prosthetics (Hudgins et al., 1993; Geethanjali, 2016). Extracted EMG features are classified using various machine learning models to capture the nonlinear relationship between EMG signals and hand gestures or grasping forces (Phinyomark et al., 2018; Joshi et al., 2024). To enhance the performance of myoelectric-based upper-limb prosthetics, particularly in achieving accurate execution of force and position, several researchers have explored different feature patterns alongside classification techniques (Choi et al., 2009; Phinyomark et al., 2012; Váscónez et al., 2023). Among these techniques, artificial neural networks (ANNs) are the most widely employed and versatile classifiers for EMG signal feature classification (Veer & Sharma, 2016; Waris et al., 2018). Researchers (Farina et al., 2014) compared the accuracy of ANN, non-negative matrix factorization, and linear regression in mapping kinematics from EMG in offline tests. Their findings indicated that ANN exhibited a higher correlation between measured and estimated kinematics than the other methods.

Most existing studies have focused on either hand (or wrist) position/gesture recognition or grasping force prediction using EMG signals (Prakash et al., 2019; Zhang et al., 2019; Saikia et al., 2022). However, the main limitation of these studies is that force and position are estimated separately, which is not directly relevant to the nature of human movement. It is essential to predict both motion (kinematics) and force (kinetics) simultaneously to develop a functional and effective hand prosthetic capable of realistically mimicking human hand behavior in daily activities. Thus, the aim of our research was to predict both hand grasping force and gestures simultaneously using EMG signals. For this purpose, EMG signals were recorded during various hand gestures, while the grasping force was concurrently measured using a hand dynamometer. Time-domain features from the EMG signals were extracted and classified using an ANN to simultaneously predict grasping force and recognize hand gestures. To the best of our knowledge, this is the first study to simultaneously predict both grasping force and hand gesture patterns for the upper limb.

2. MATERIAL AND METHOD

2.1. Experimental Protocol

Ten healthy participants (3 females, 7 males; age: 20.9 ± 1.21 years, height: 174.1 ± 8.19 cm, weight: 79.5 ± 19.45 kg) with no musculoskeletal or neurological disorders participated in this study. All were right-handed and instructed to avoid upper extremity exertion prior to the experiment. Prior to the signal recording process, participants were provided with adequate information about the experiments and their informed consent was obtained. The workflow of the study is presented in Figure 1.

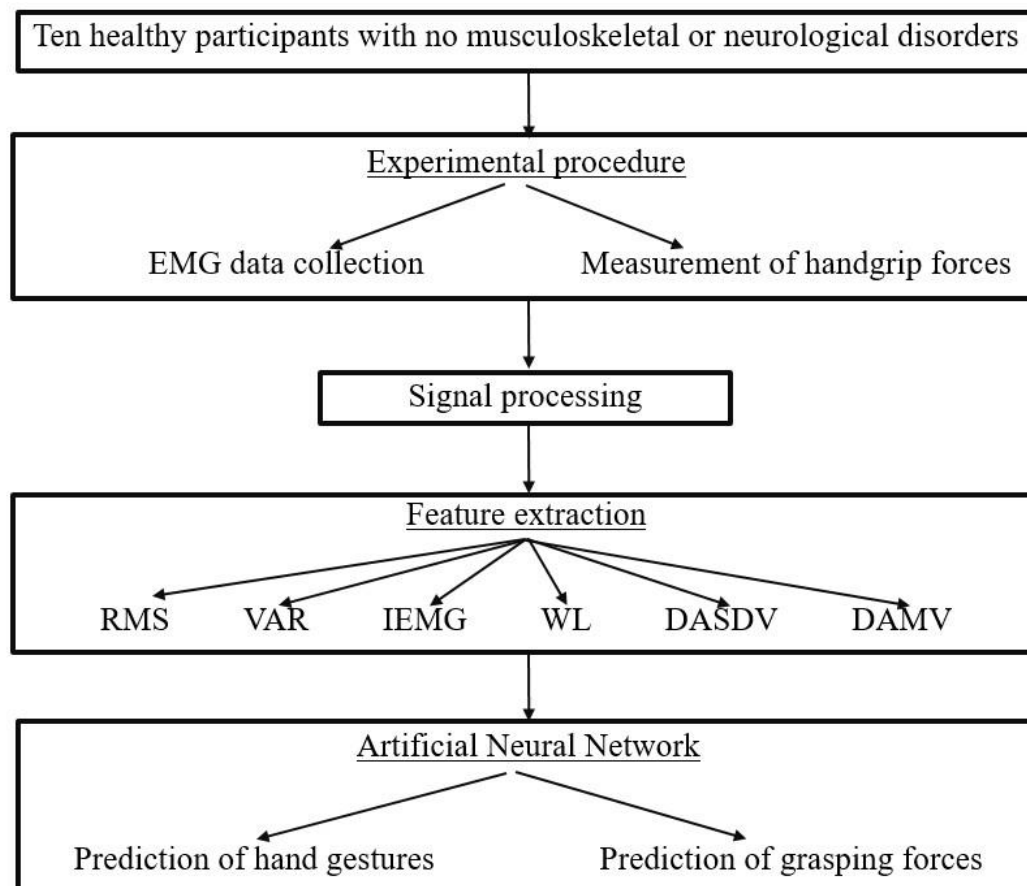


Figure 1. Flowchart of the study. RMS: Root mean square, VAR: Variance of EMG, IEMG: Integrated EMG, WL: Waveform length, DASDV: Difference absolute standard deviation value, DAMV: Difference absolute mean value.

sEMG signals were recorded from the flexor carpi radialis (FCR) muscle. The EMG signals were recorded using a Vernier EKG Sensor (EKG-BTA) and LabQuest 2 (Vernier Software and Technology, USA) with a 100 Hz sampling rate (Figure 2a). Electrode placement and skin preparation followed the SENIAM Project guidelines (Hermens et al., 1999). Pre-gelled, self-adhesive disposable surface electrodes were used. The electrodes contained an electrolytic gel interface. The skin at electrode sites was shaved and cleaned with alcohol. Muscle belly was identified and marked using a pen for accurate electrode placement. Electrodes were then attached to the marked site and the EMG receivers were connected to the electrodes.

Handgrip forces were measured using a strain-gauge-based isometric hand dynamometer (accuracy: ± 0.6 N, resolution: 0.2141 N, sampling rate: 100 Hz) (Vernier, Beaverton, USA). During the EMG and handgrip force recording, the participants sat upright in a neutral posture (Figure 2b). Each trial during the measurements had a duration of 3 seconds.

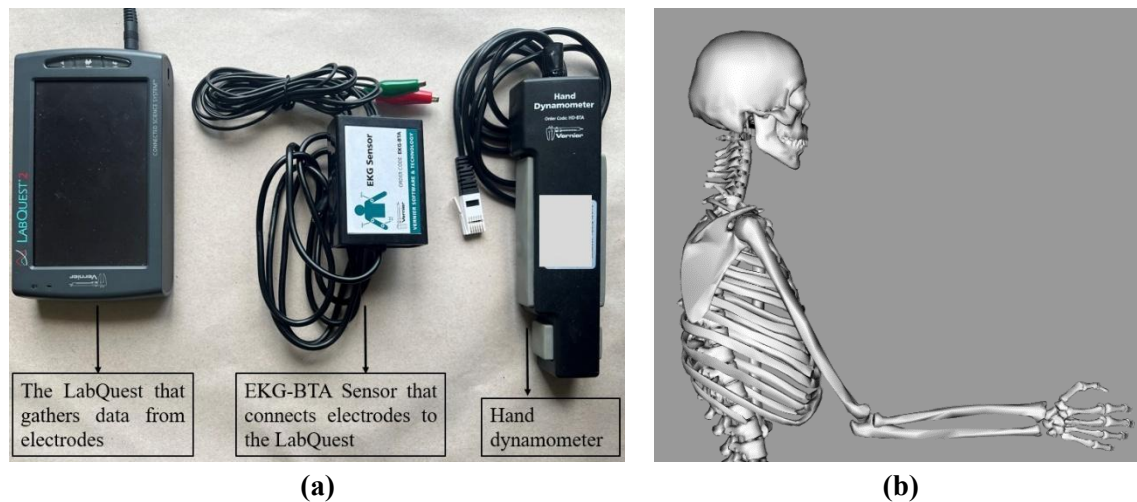


Figure 2. *a) Details of the experimental rig. b) The subjects' positioning during the EMG and handgrip force assessments involved the shoulder being placed in 30° abduction and the elbow in 90° flexion.*

The participants were instructed to firmly grasp the dynamometer and apply handgrip forces during three distinct wrist motion patterns: i) The wrist was positioned neutrally and remained motionless (resting position), ii) the wrist was in flexion with the fingers oriented upwards (wrist flexion), and iii) the hand was in pronation. They performed three trials for each wrist position, with a minimum rest period of ten seconds in neutral position with the hand open and flat on the table, free of muscle contraction between each repetition. After completing each set, the participants rested in a neutral position for at least two minutes before starting the next motion pattern.

For EMG signal normalization, participants were asked to perform a maximum voluntary contraction three times to maximize their grasping force, and thereby the amplitudes of the corresponding EMG signals. The participants maintained the same posture throughout all protocols. A three-minute rest period was provided between trials to prevent muscle fatigue (Kamavuako et al., 2009).

2.2. Signal Processing and Feature Extraction

The raw EMG signals underwent processing through a 6th-order band-pass filter, with cut-off frequencies of 15 Hz on the lower end and 450 Hz on the upper end, followed by full-wave rectification (Tkach et al., 2010). The rectified signals were divided into segments using sliding windows, with each window having a time span of 500 ms. The time gap between consecutive windows was set to 25 ms (Chen et al., 2013). All EMG signals were normalized based on those recorded during maximal voluntary contraction.

Various signal features can be applied to quantitatively describe EMG signals and reduce the size of the data (Zecca et al., 2002; Shin et al., 2024). In our research, since our goal was to correlate EMG signals with different hand gestures and handgrip forces of varying amplitudes, we employed the following time-domain signal features, which are commonly used to establish the nonlinear relationship between EMG signals and hand gestures (Phinyomark et al., 2012; Shin et al., 2024).

2.2.1. Root Mean Square (RMS)

Root mean square (RMS) is a well-accepted feature in the time domain analysis of EMG signal (Boostani & Moradi, 2003; Phinyomark et al., 2012). The mathematical expression of RMS can be defined as:

$$RMS = \sqrt{\frac{1}{N} \sum_{k=1}^N x_k^2} \quad (1)$$

where x_k represents the k^{th} sample and N is the total number of samples in each segment. These abbreviations are also applicable in the subsequent formulations.

2.2.2. Variance of EMG (VAR)

Variance (VAR) of EMG is expressed as a measure of EMG signal power (Zecca et al., 2002; Phinyomark et al., 2012). The mathematical expression of VAR can be defined as:

$$VAR = \frac{1}{N-1} \sum_{k=1}^N x_k^2 \quad (2)$$

2.2.3. Integrated EMG (IEMG)

Integrated EMG (IEMG) is a cumulative absolute value of EMG signal (Phinyomark et al., 2012). The calculation of IEMG feature is defined as:

$$IEMG = \sum_{k=1}^N |x_k| \quad (3)$$

2.2.4. Waveform Length (WL)

Waveform length (WL) is an estimation of the cumulative length of EMG waveform in a segment (Kamavuako et al., 2013). The waveform length of EMG is defined as:

$$WL = \sum_{k=1}^{N-1} |x_{k+1} - x_k| \quad (4)$$

2.2.5. Difference Absolute Standard Deviation Value (DASDV)

The difference absolute standard deviation value (DASDV) is the absolute deviation of the standard deviation between adjacent samples. It is calculated by taking the first-order differential value of the RMS feature (Phinyomark et al., 2014) and defined as follows:

$$DASDV = \sqrt{\frac{1}{N-1} \sum_{k=1}^{N-1} (x_{k+1} - x_k)^2} \quad (5)$$

2.2.6. Difference Absolute Mean Value (DAMV)

The difference absolute mean value (DAMV) is calculated for each data window using the following equation (Phinyomark et al., 2012).

$$DAMV = \frac{1}{N-1} \sum_{k=1}^{N-1} |x_{k+1} - x_k| \quad (6)$$

2.3. Artificial Neural Networks (ANN)

ANN is commonly utilized for classifying EMG signals due to its ability to manage the nonlinear nature of muscle activations (Waris et al., 2018; Saikia et al., 2022). In our research, we used ANN to predict hand gestures and handgrip force simultaneously. The neural network architecture consisted of one input layer, two hidden layers, and one output layer. The model used for training features had layers with 100, 60, 30, and 501 neuron. A log-sigmoid transfer function was used as the activation function. The number of epochs for the training phase was set to 1000. Additionally, the force values in the target set were normalized to a range between 0 and 1 to enhance the classification performance of the network. Furthermore, the training and testing datasets were carefully adjusted to ensure the optimal pairing of train-test data for the specific case under investigation. To achieve this, data from all trials were incorporated into the analysis. Leave-one-out cross-validation (LOOCV), a special type of k-fold cross-validation, was employed to train the network and assess the accuracy of its predictions (Arlot & Celisse, 2010). In this method, the network was trained using all data except for one trial, and the prediction (or test) was made for that specific trial. The LOOCV process was repeated until all trials had been used for both training and testing. As a result, the average error was calculated by taking the arithmetic mean of the errors from each test trial. The extracted EMG signal features, which served as inputs to the neural network, were also normalized with respect to the corresponding EMG signals recorded during the maximal voluntary contraction test (Soylu & Arpinar-Avsar, 2010).

2.4. Data Analysis

Hand gesture recognition was quantified by calculating the ratio of correct predictions to the total number of predictions (i.e., true and false predictions). To quantitatively assess the relationship between the predicted and

experimental force-time histories, root mean square differences (RMSD) and Pearson cross-correlation coefficients (PCC) were computed. An RMSD value of 0.1 indicates that the mean error in magnitude between the predicted and actual forces is 10%. PCC serves as a measure of similarity between two curves; a PCC value of 1 between two force-time histories suggests a 100% agreement between the predicted and experimental handgrip force profiles.

Statistical significance was analyzed using SPSS software (Version 21.0; SPSS; Chicago, IL, USA), with a significance threshold set at 0.05 ($p < 0.05$). The Shapiro-Wilk test was conducted to assess the normality of the average prediction values for hand gestures, as well as the average RMSD and PCC values. All data were statistically evaluated using Friedman's ANOVA, and the Mann-Whitney U test was applied to determine significant differences between groups. Bonferroni correction was used to adjust the p -value for multiple comparisons ($p < 0.016$).

3. RESULTS AND DISCUSSION

The results of hand gesture recognition and handgrip force prediction for all features were presented in Table 1. As shown in Table 1, the highest success rates for hand gesture prediction were achieved using RMS (90%) and WL (88%) ($p=0.013$). For the resting position, the best recognition rate was achieved using WL, while RMS provided the best performance for wrist flexion. Both RMS and WL produced the same detection rate for hand pronation. No statistically significant difference was observed between RMS and WL. As it can be seen from Table 2, the minimum average RMSD (0.09) and maximum PCC (0.89) values between the actual and predicted force-time histories were obtained from RMS ($p=0.014$).

Table 1. Hand gestures Recognition Rates for all Signal Features

	Resting position	Wrist flexion	Hand pronation	Total
RMS	27/30 (90.0%)	28/30 (93.3%)	26/30 (86.7%)	81/90 (90.0%)
VAR	27/30 (90.0%)	27/30 (90.0%)	22/30 (73.3%)	76/90 (84.4%)
WL	28/30 (93.3%)	26/30 (86.7%)	26/30 (86.7%)	80/90 (88.0%)
IEMG	27/30 (90.0%)	26/30 (86.7%)	21/30 (70.0%)	74/90 (82.2%)
DASDV	21/30 (70.0%)	17/30 (56.7%)	20/30 (66.7%)	58/90 (64.4%)
DAMV	25/30 (83.3%)	24/30 (80.0%)	17/30 (56.7%)	66/90 (73.3%)

RMS: Root mean square, VAR: Variance of EMG, WL: Waveform length, IEMG: Integrated EMG, DASDV: Difference absolute standard deviation value, DAMV: Difference absolute mean value.

Table 2. The Mean RMSD and PCC Results Obtained from all Subjects for all Features

	Resting position		Wrist flexion		Hand pronation	
	RMSD	PCC	RMSD	PCC	RMSD	PCC
RMS	0.09	0.89	0.15	0.85	0.20	0.84
VAR	0.12	0.81	0.16	0.83	0.26	0.75
WL	0.23	0.84	0.24	0.82	0.26	0.74
IEMG	0.14	0.88	0.17	0.81	0.21	0.85
DASDV	0.26	0.77	0.24	0.71	0.28	0.72
DAMV	0.24	0.79	0.22	0.75	0.31	0.73

RMS: Root mean square, VAR: Variance of EMG, WL: Waveform length, IEMG: Integrated EMG, DASDV: Difference absolute standard deviation value, DAMV: Difference absolute mean value, RMSD: Root mean square differences, PCC: Pearson cross-correlation coefficients.

In this study, the objective was to map the EMG signals from the forearm muscles to hand gestures and handgrip forces using an artificial neural network (ANN) combined with time-domain feature extraction techniques. To better simulate real-world conditions, we designed the current protocol, in which both hand gesture and applied force were predicted simultaneously. To our knowledge, this is the first study that attempts to predict both force and position concurrently. The analysis of hand gesture prediction results revealed that RMS and WL outperformed other features (Table 1) ($p = 0.013$). Contrary to the findings of the related study (Phinyomark et al., 2014), features derived from the first difference of the EMG time series, such as DASDV and DAMV, did not yield more accurate prediction results compared to their original feature counterparts. The inferior performance of DASDV and DAMV can be attributed to the classifier employed in our study, as the effectiveness of ANN classification is greatly influenced by network structure parameters (Phinyomark et al., 2012). When evaluating each hand gesture individually, it was noted that the resting position was predicted with the highest accuracy, while the prediction rate for hand pronation was significantly lower compared to other positions ($p = 0.011$) (Table 1). This outcome may be explained by the role of the associated muscles, as the FCR functions as radial abductors in addition to their primary roles. As a result, the EMG data from this muscle might not be sufficient to accurately represent the hand pronation pattern.

Based on the RMSD, PCC, and gesture prediction outcomes, RMS and WL consistently provided superior prediction performance for most positions compared to other features ($p = 0.014$).

Although the classification of certain hand motion patterns and the estimation of handgrip force using EMG signals from a single muscle yielded reasonable predictions in some instances, using EMG data from multiple muscles could result in significantly improved accuracy. This aligns with the findings of the researchers (Zhang et al., 2016), who observed a reduction in root mean square error when the FCR and ECRL muscles were included together in the prediction model. Similarly, researchers emphasized the need to account for co-activation when estimating the forces of forearm muscles (Brown et al., 2010).

Given that prediction accuracy is heavily dependent on the ANN structure, identifying an optimal ANN configuration remains a challenge (Demir et al., 2016). While the ANN provided promising results, varying the network structure parameters would yield different outcomes. The inherent instability and inconsistency of the ANN are significant drawbacks of this machine learning approach. Consequently, future research will focus on evaluating classifiers whose performance is less dependent on the optimization of structure and parameters, such as linear and quadratic discriminant analysis and k-nearest neighbors.

Several limitations of this study should be acknowledged. Firstly, only EMG data from one muscle were used. Although incorporating additional forearm muscles could improve prediction performance, this would likely introduce higher computational costs and greater nonlinearity. Secondly, our study focused on three hand gestures and isometric handgrip forces. To increase the generalizability of the methods used, future studies should include a broader range of hand/wrist motion patterns and dynamic muscle contractions.

4. CONCLUSION

This study comparatively assessed several time-domain EMG feature extraction techniques, including RMS, VAR, WL, IEMG, DASDV, and DAMV, for the concurrent estimation of hand gestures and handgrip force using ANN. The results demonstrated that RMS and WL provided effective outcomes for hand gesture recognition, while RMS and IEMG were the most accurate for predicting handgrip forces. The insights gained from this study are expected to contribute to the development of control systems for EMG-based applications, such as myoelectric-controlled prostheses. These findings highlight the importance of feature extraction methods in the classification performance of EMG-based control models for prosthetics. Future studies can be conducted to investigate factors such as expanding the dataset, considering different movement patterns, recording EMG signals from more muscles, and using different machine learning methods to improve the success rate and real-world applicability.

AUTHOR CONTRIBUTIONS

Conceptualization, D.K. and S.C.D.; methodology, D.K. and S.C.D.; software, D.K.; title, D.K. and S.C.D.; validation, D.K. and S.C.D.; laboratory work, D.K. and S.C.D.; research, D.K. and S.C.D.; sources, D.K. and S.C.D.; manuscript-original draft, D.K. and S.C.D.; manuscript-review and editing, D.K. and S.C.D.; visualization, D.K.; supervision, S.; funding, the authors have no received any financial support for the

research, authorship or publication of this study. All authors have read and legally accepted the final version of the article published in the journal.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Al-Timemy, A. H., Khushaba, R. N., Bugmann, G., & Escudero, J. (2015). Improving the performance against force variation of EMG controlled multifunctional upper-limb prostheses for transradial amputees. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 24(6), 650-661. <https://doi.org/10.1109/TNSRE.2015.2445634>
- Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics Survey*, 4, 40-79. <https://doi.org/10.1214/09-SS054>
- Atzori, M., Gijsberts, A., Castellini, C., Caputo, B., Hager, A. G. M., Elsig, S., Giatsidis, G., Bassetto, F., & Müller, H. (2014). Electromyography data for non-invasive naturally-controlled robotic hand prostheses. *Scientific Data*, 1(1), 1-13. <https://doi.org/10.1038/sdata.2014.53>
- Basmajian, J. V. & DeLuca, C. J. (1985). *Muscle Alive. Their Functions Revealed by Electromyography* (5th ed.). Baltimore, MD: Williams and Wilkins.
- Boostani, R., & Moradi, M. H. (2003). Evaluation of the forearm EMG signal features for the control of a prosthetic hand. *Physiological Measurement*, 24(2), 309. <https://doi.org/10.1088/0967-3334/24/2/307>
- Brown, S. H., Brookham, R. L., & Dickerson, C. R. (2010). High-pass filtering surface EMG in an attempt to better represent the signals detected at the intramuscular level. *Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine*, 41(2), 234-239. <https://doi.org/10.1002/mus.21470>
- Chen, X., Zhang, D., & Zhu, X. (2013). Application of a self-enhancing classification method to electromyography pattern recognition for multifunctional prosthesis control. *Journal of Neuroengineering and Rehabilitation*, 10, 1-13. <https://doi.org/10.1186/1743-0003-10-44>
- Choi, C., Micera, S., Carpaneto, J., & Kim, J. (2009). Development and quantitative performance evaluation of a noninvasive EMG computer interface. *IEEE Transactions on Biomedical Engineering*, 56(1), 188-191. <https://doi.org/10.1109/TBME.2008.2005950>
- Demir, U., Kocaoğlu, S., & Akdoğan, E. (2016). Human impedance parameter estimation using artificial neural network for modelling physiotherapist motion. *Biocybernetics and Biomedical Engineering*, 36(2), 318-326. <https://doi.org/10.1016/j.bbe.2016.01.002>
- Farina, D., Jiang, N., Rehbaum, H., Holobar, A., Graimann, B., Dietl, H., & Aszmann, O. C. (2014). The extraction of neural information from the surface EMG for the control of upper-limb prostheses: emerging avenues and challenges. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(4), 797-809. <https://doi.org/10.1109/TNSRE.2014.2305111>

- Geethanjali, P. (2016). Myoelectric control of prosthetic hands: state-of-the-art review. *Medical Devices: Evidence and Research*, 247-255. <https://doi.org/10.2147/MDER.S91102>
- Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, G., Cai, J., & Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern Recognition*, 77, 354-377. <https://doi.org/10.1016/j.patcog.2017.10.013>
- Haig, A. J., Gelblum, J. B., Rechten, J. J., & Gitter, A. J. (1996). AAEM practice topics: technology assessment: the use of surface EMG in the diagnosis and treatment of nerve and muscle disorders. *Muscle & Nerve: Official Journal of the American Association of Electrodiagnostic Medicine*, 19(3), 392-395. [https://doi.org/10.1002/\(sici\)1097-4598\(199603\)19:3%3C392::aid-mus21%3E3.0.co;2-t](https://doi.org/10.1002/(sici)1097-4598(199603)19:3%3C392::aid-mus21%3E3.0.co;2-t)
- Hermens, H. J., Freriks, B., Merletti, R., Stegeman, D., Blok, J., Rau, G., Disselhorst-Klug, C., & Hägg, G. (1999). European recommendations for surface electromyography. *Roessingh Research and Development*, 8(2), 13-54. [https://doi.org/10.1016/S1050-6411\(00\)00027-4](https://doi.org/10.1016/S1050-6411(00)00027-4)
- Hudgins, B., Parker, P., & Scott, R. N. (1993). A new strategy for multifunction myoelectric control. *IEEE Transactions on Biomedical Engineering*, 40(1), 82-94. <https://doi.org/10.1109/10.204774>
- Jahani Fariman, H., Ahmad, S. A., Hamiruce Marhaban, M., Alijan Ghasab, M., & Chappell, P. H. (2016). Hand movements classification for myoelectric control system using adaptive resonance theory. *Australasian Physical & Engineering Sciences in Medicine*, 39, 85-102. <https://doi.org/10.1007/s13246-015-0399-5>
- Joshi, D. C., Kumar, P., Joshi, R. C., & Mitra, S. (2024). AI-Enhanced Analysis to Investigate the Feasibility of EMG Signals for Prosthetic Hand Force Control Incorporating Anthropometric Measures. *Prosthesis*, 6(6). <https://doi.org/10.3390/prosthesis6060106>
- Kamavuako, E. N., Farina, D., Yoshida, K., & Jensen, W. (2009). Relationship between grasping force and features of single-channel intramuscular EMG signals. *Journal of Neuroscience Methods*, 185(1), 143-150. <https://doi.org/10.1016/j.jneumeth.2009.09.006>
- Kamavuako, E. N., Rosenvang, J. C., Bøg, M. F., Smidstrup, A., Erkocevic, E., Niemeier, M. J., Jensen, W., & Farina, D. (2013). Influence of the feature space on the estimation of hand grasping force from intramuscular EMG. *Biomedical Signal Processing and Control*, 8(1), 1-5. <https://doi.org/10.1016/j.bspc.2012.05.002>
- Kundu, A. S., Mazumder, O., Lenka, P. K., & Bhaumik, S. (2018). Hand gesture recognition based omnidirectional wheelchair control using IMU and EMG sensors. *Journal of Intelligent & Robotic Systems*, 91, 529-541. <https://doi.org/10.1007/s10846-017-0725-0>
- Micera, S., Carpaneto, J., & Raspopovic, S. (2010). Control of hand prostheses using peripheral information. *IEEE Reviews in Biomedical Engineering*, 3, 48-68. <https://doi.org/10.1109/RBME.2010.2085429>
- Ni, S., Al-qaness, M. A., Hawbani, A., Al-Alimi, D., Abd Elaziz, M., & Ewees, A. A. (2024). A survey on hand gesture recognition based on surface electromyography: Fundamentals, methods, applications, challenges and future trends. *Applied Soft Computing*, 112235. <https://doi.org/10.1016/j.asoc.2024.112235>

- Noce, E., Bellingegni, A. D., Ciancio, A. L., Sacchetti, R., Davalli, A., Guglielmelli, E., & Zollo, L. (2019). EMG and ENG-envelope pattern recognition for prosthetic hand control. *Journal of Neuroscience Methods*, 311, 38-46. <https://doi.org/10.1016/j.jneumeth.2018.10.004>
- Oskoei, M. A., & Hu, H. (2007). Myoelectric control systems—A survey. *Biomedical Signal Processing and Control*, 2(4), 275-294. <https://doi.org/10.1016/j.bspc.2007.07.009>
- Onay, F., & Mert, A. (2020). Phasor represented EMG feature extraction against varying contraction level of prosthetic control. *Biomedical Signal Processing and Control*, 59, 101881. <https://doi.org/10.1016/j.bspc.2020.101881>
- Parajuli, N., Sreenivasan, N., Bifulco, P., Cesarelli, M., Savino, S., Niola, V., Esposito, D., Hamilton, T. J., Naik, G. N., Gunawardana, U., & Gargiulo, G. D. (2019). Real-time EMG based pattern recognition control for hand prostheses: A review on existing methods, challenges and future implementation. *Sensors*, 19(20), 4596. <https://doi.org/10.3390/s19204596>
- Phinyomark, A., Phukpattaranont, P., & Limsakul, C. (2012). Feature reduction and selection for EMG signal classification. *Expert Systems with Applications*, 39(8), 7420-7431. <https://doi.org/10.1016/j.eswa.2012.01.102>
- Phinyomark, A., Quaine, F., Charbonnier, S., Serviere, C., Tarpin-Bernard, F., & Laurillau, Y. (2014). Feature extraction of the first difference of EMG time series for EMG pattern recognition. *Computer Methods and Programs in Biomedicine*, 117(2), 247-256. <https://doi.org/10.1016/j.cmpb.2014.06.013>
- Phinyomark, A., N. Khushaba, R., & Scheme, E. (2018). Feature extraction and selection for myoelectric control based on wearable EMG sensors. *Sensors*, 18(5), 1615. <https://doi.org/10.3390/s18051615>
- Prakash, A., Sharma, S., & Sharma, N. (2019). A compact-sized surface EMG sensor for myoelectric hand prosthesis. *Biomedical Engineering Letters*, 9(4), 467-479. <https://doi.org/10.1007/s13534-019-00130-y>
- Saikia, A., Mazumdar, S., Sahai, N., Paul, S., & Bhatia, D. (2022). Performance analysis of artificial neural network for hand movement detection from EMG signals. *IETE Journal of Research*, 68(2), 1074-1083. <https://doi.org/10.1080/03772063.2019.1638316>
- Shin, J., Miah, A. S. M., Kabir, M. H., Rahim, M. A., & Al Shiam, A. (2024). A methodological and structural review of hand gesture recognition across diverse data modalities. *IEEE Access*. <https://doi.org/10.1109/ACCESS.2024.3456436>
- Sims, T., Donovan-Hall, M., & Metcalf, C. (2020). Children's and adolescents' views on upper limb prostheses in relation to their daily occupations. *British Journal of Occupational Therapy*, 83(4), 237-245. <https://doi.org/10.1177/0308022619865179>
- Soylu, A. R., & Arpinar-Avsar, P. (2010). Detection of surface electromyography recording time interval without muscle fatigue effect for biceps brachii muscle during maximum voluntary contraction. *Journal of Electromyography and Kinesiology*, 20(4), 773-776. <https://doi.org/10.1016/j.jelekin.2010.02.006>

- Tkach, D., Huang, H., & Kuiken, T. A. (2010). Study of stability of time-domain features for electromyographic pattern recognition. *Journal of Neuroengineering and Rehabilitation*, 7, 1-13. <https://doi.org/10.1186/1743-0003-7-21>
- Vásconez, J. P., López, L. I. B., Caraguay, Á. L. V., & Benalcázar, M. E. (2023). A comparison of EMG-based hand gesture recognition systems based on supervised and reinforcement learning. *Engineering Applications of Artificial Intelligence*, 123, 106327. <https://doi.org/10.1016/j.engappai.2023.106327>
- Veer, K., & Sharma, T. (2016). A novel feature extraction for robust EMG pattern recognition. *Journal of Medical Engineering & Technology*, 40(4), 149-154. <https://doi.org/10.3109/03091902.2016.1153739>
- Waris, A., Niazi, I. K., Jamil, M., Englehart, K., Jensen, W., & Kamavuako, E. N. (2018). Multiday evaluation of techniques for EMG-based classification of hand motions. *IEEE Journal of Biomedical and Health Informatics*, 23(4), 1526-1534. <https://doi.org/10.1109/JBHI.2018.2864335>
- Wu, Y., Liang, S., Yan, T., Ao, J., Zhou, Z., & Li, X. (2022). Classification and simulation of process of linear change for grip force at different grip speeds by using supervised learning based on sEMG. *Expert Systems with Applications*, 206, 117785. <https://doi.org/10.1016/j.eswa.2022.117785>
- Zecca, M., Micera, S., Carrozza, M. C., & Dario, P. (2002). Control of multifunctional prosthetic hands by processing the electromyographic signal. *Critical Reviews™ in Biomedical Engineering*, 30(4-6). <https://doi.org/10.1615/CritRevBiomedEng.v30.i456.80>
- Zhang, S., Guo, S., Gao, B., Huang, Q., Pang, M., Hirata, H., & Ishihara, H. (2016). Muscle strength assessment system using sEMG-based force prediction method for wrist joint. *Journal of Medical and Biological Engineering*, 36, 121-131. <https://doi.org/10.1007/s40846-016-0112-5>
- Zhang, L., Liu, G., Han, B., Wang, Z., & Zhang, T. (2019). sEMG based human motion intention recognition. *Journal of Robotics*, 2019(1), 3679174. <https://doi.org/10.1155/2019/3679174>



Gazi University

Journal of Science

PART A: ENGINEERING AND INNOVATION

<http://dergipark.org.tr/guj.1647168>

Effect of Yolo-Based Vehicle Detection and Manuel Vehicle Classification Methods on Simulation-Based Intersection Analysis

Bahadır Ersoy ULUSOY¹ Dogukan HAZAR² Yalcın ALBAYRAK² Sevil KOFTECI^{1*} Mehmet Arikan YALCIN¹ Alican DEMİRAL¹

¹ Akdeniz University, Faculty of Engineering, Department of Civil Engineering, 07070, Antalya, Türkiye

² Akdeniz University, Faculty of Engineering, Department Electrical and Electronics Engineering, 07070, Antalya, Türkiye

Keywords	Abstract
Signalized Intersection	In complicated intersections, for measuring intersection performance using simulation programs (PTV Vissim, Aimsun, Sidra, etc.), decomposing camera recording data procured for vehicle counting according to vehicle type and movement direction is laborious. In this study, a software application was developed using YOLO (You Only Look Once) object detection model to classify vehicle types and estimation the number of vehicles from images obtained with camera recordings. The system aims to provide high-accuracy results in significantly less time compared to manual visual counting methods. The proposed model achieved a detection accuracy of 91.8%. The software was tested on video data from the Sampi Intersection, and the results were validated by comparing them to manual counting. Based on the obtained data, performance analyses were conducted using the PTV Vissim software by evaluating vehicle delays, determining service levels, and making comparative assessments. As a result of the simulation analysis, the vehicle delays, which are 38.82 seconds in the manual count method and 37.36 seconds in the software counting method, are between 35.1-55 seconds and correspond to the service level value D. The results demonstrate that YOLO-based vehicle counting and classification systems can offer an efficient and reliable alternative for traffic monitoring and intersection analysis.
Simulation	
Vehicle Count	
Vissim	
Artificial Intelligence	

Cite

Ulusoy, B.E., Hazar, D., Albayrak, Y., Kofteci, S., Yalcin, M.A., & Demiral, A. (2025). Effect of Yolo-Based Vehicle Detection and Manuel Vehicle Classification Methods on Simulation-Based Intersection Analysis. *GU J Sci, Part A, 12(2)*, 665-690. doi:10.54287/guj.1647168

Author ID (ORCID Number)	Article Process
0000-0003-3367-8312	Bahadır Ersoy ULUSOY
0009-0009-2285-2515	Dogukan HAZAR
0000-0002-5632-0869	Yalcın ALBAYRAK
0000-0002-5096-2545	Sevil KOFTECI
0000-0002-8916-1411	Mehmet Arikan YALCIN
0009-0001-8241-8253	Alican DEMİRAL

1. INTRODUCTION

Real-time traffic management systems are necessary for the rearrangement of traffic, especially in intersections. In addition to the studies on signalized intersections performance in the literature, many studies have been made and ideas have been presented on traffic counting methods in line with the developing technological conditions. Analyzing the performance of intersections in transportation and traffic engineering and making appropriate interventions when deemed necessary is vital for traffic safety and flow. Vehicle counting data are necessary data for intersection performance analysis in simulation programs. (PTV Vissim, Aimsun, Sidra, etc.). It is very difficult to manually decompose the camera recording data obtained for vehicle counting, especially in complex intersections, according to the vehicle type and the direction of the vehicle's

*Corresponding Author, e-mail: skofteci@akdeniz.edu.tr

movement, and it is necessary to work with a large team. Determination number of vehicles is a key element for developing traffic management systems in urban areas, highway surveillance, as well as vehicle parking management systems.

Real-time vehicle number determination and monitoring, from security cameras are an essential part of many applications of intelligent transportation system based smart cities. Usually, this task causes some problems in the application, such as videos not being rendered or not being detected in real time and/or tracking errors. Real-time traffic management systems, which are an important part of vehicle counting, have recently become popular through low-cost surveillance cameras, powerful mobile video analytics devices and cloud infrastructure (Oltean et al., 2019). Indolia et al. (2018) examines the conceptual understanding of neural network (CNN) along with its three most common architectures, and learning algorithms. A model-based vehicle tracking and classification system that is robust in most circumstances has been presented by (Gupte et al., 2002). An intelligent transport system (ITS) has an important role in monitoring and managing traffic density by counting the number of vehicles. Information on vehicle counting can be used to reduce the number of traffic jams (e.g. Dynamic traffic redirection, automatic traffic light control) (Sooksatra et al., 2019). The number of vehicle and vehicle classification is one of the fundamental steps in traffic management (Cruz et al., 2019). It is also important data in transportation planning. Chung and Sohn (2018) devised a deep convolutional neural network in order to determine the number of vehicles on a road segment based on video images. Kulkarni et al. (2025) studied a model for the detection of objects in traffic. The model is based on leader-follower (LT) and vehicle tracking (AT) using the “domain” from the Wiedemann-99 model. Thanks to the model, they were able to accurately detect vehicle counts on lane-less roads compared to other models.

Vehicle detection can be accomplished in different ways, such as background subtraction or optical flow. The primary step of the traffic monitoring system is vehicle detection. Some of the traditional methods for vehicle detection are microwave detectors, inductive loops, passive magnetic sensors etc (El-Khoreby et al., 2019). Vehicle counting is an important part of the ITS. Compared with traditional vehicle detection methods, a convolutional CNN can achieve real-time speed with high accuracy for the vehicle detection. Yao (2019) provides an effective and robust tool count method by combining CNN and virtual coils. The research reported by Zheng and Mike (2012) investigated the empirical determination of errors in manual traffic counting. They said in their paper that manual counting has not been well studied and reported, despite its extensive use in national traffic counting and assessment. Pal'o et al. (2019) studied manual counting and automatic vehicle counters. In this work, two traffic-related studies were conducted and the results obtained from manual counting and automatic counting were compared with each other. In their study, Findley et al. (2011) aimed to compare the results of roadside data collected using typical manual methods with data collected by data collection vehicles travelling with traffic. Lin and Sun (2018) proposed for traffic flow counting system that utilizes the YOLO framework in their paper. For this aim, they used three blocks. These blocks include the detector that generates the bounding box of vehicles and the buffer that stores the coordinates of vehicles. The

counter they used was responsible for vehicle counting. The videos were used to verify and analyze the system. The efficiency of the developed system, and the results indicated that their system achieves high counting results under with sufficient ambient light. Bao et al. (2016) designed a low-cost as well as energy-efficient type of multi-function traffic magnetic sensor which is working wireless. A basic vehicle detection algorithm (DWVDA) based on the magnetic traffic sensor with less complex computing was introduced to determine the number of vehicles in low-traffic volume conditions in this study. The experimental results have shown that developed algorithms yield a satisfactory performance in both low-traffic volume conditions and heavy traffic volumes. The second algorithm was proven to achieve better results in heavy traffic volume. The results show that the proposed first algorithm can make efficient and accurate vehicle detection in jam flow conditions and can be employed in field traffic environments. Toth et al. (2013) developed a manual traffic counting application for Android tablets. They aimed to investigate the issue of data truth based on manual counts. The results of this study showed the benefits of the verification process. They suggested that using their developed application could minimize arising from human error in manual counting. Thus this application provides more accurate traffic count data for use in transport planning studies. Li et al. (2020) proposed a framework for counting vehicles and estimating traffic flow parameters such as flow, density in dense traffic conditions. The evaluation experiments show that the proposed framework can efficiently determine the number of vehicles. Additionally, by using the proposed framework, the estimation of traffic parameters raised high accuracy. Dai et al. (2019) develop a video-based vehicle counting framework. For this purpose, they used a three-component process of object. These are detection, object tracking and trajectory processing. In this study, the vehicles were counted according to the vehicle types and their routes to obtain accurate traffic flow information. Obtained results showed that the accuracy of their vehicle counting method reaches more than 90%. Kolluri and Das (2023) studied the YOLO-v5 and RetinaNet models to address the challenges associated with detecting pedestrians in traffic. They also used the KELM algorithm to classify pedestrians. In this way, they were able to detect pedestrians more accurately in traffic. In their study, Bhosale et al. (2023) used deep learning techniques in video surveillance in order to improve the problems experienced in tracking the position of vehicles in traffic with traditional methods. Thus, they developed a more accurate approach than traditional approaches. In their study, Ge et al. (2023) tested a model for complete and accurate identification of vehicles on a bridge. Their YOLO-v4 based model provides a two-step approach to vehicle appearance and tracking. As a result of their study, they found that the model is faster and more accurate in predicting small objects. Ma et al. (2023) studied how to reduce the problems (e.g. background image clarity, object detection, scale changes, etc.) that may occur during object detection in Unmanned Aerial Vehicles (UAV) traffic. By using YOLO-v5s based RTOD-YOLO model, they planned to avoid the problems mentioned in the above sentence. As a result, they were able to make more accurate and clearer detections thanks to an improved detection method model.

As can be seen; many studies have been conducted on the collection and separation of traffic data which are important in matters that intelligent traffic systems, intersection performance analysis, and current situation

analysis, etc. In our country, researches and studies have been conducted on subjects such as vehicle tracking, object detection, image analysis. The common purpose of these studies serve is to reduce the manpower, excess cost, labor, and loss of time and to eliminate the damage risks that may arise in asphalt surface counting methods from traditional counting methods. Altun et al. (2005) within the scope of traffic engineering applications, determination of vehicle crossing intervals at level intersections that are not signalized and not controlled with stop signs, AADT (annual average daily traffic) estimation, traffic accident analysis, traffic flow estimation and classification on highways, partially closed roads capacity and queuing calculation issues have been studied extensively. It is stated that transportation engineering problems are affected by factors that cannot be expressed mathematically such as road, human, vehicle, and environmental conditions. As a result of this study, considering the developments in computer technology, it is thought that faster and more reliable solutions will be produced for transportation engineering problems with artificial neural networks technique in the near future.

In the study conducted by Yıldırım et al. (2019) vehicle counts were performed using image processing techniques and it was proven that it could be applied as an alternative to existing systems. Sivaraman and Trivedi (2013) researched detailed advances in vehicle detection. They discussed monocular, stereo vision, and active sensor-vision fusion for on-road vehicle detection; and they provided a survey of recent works in the literature. The usability of the obtained data in solving the mentioned traffic problems has been shown. It was mentioned that besides the identification and counting of vehicles, classification of vehicles can be performed, and thus the usage rates of the regions or lanes designated on the highways can be determined according to vehicle types. Papageorgiou (2006) provides a wider spectrum of understanding on the principles of traffic flow theory and how a traffic flow simulation model is developed and utilized for effective transportation planning. Within the scope of the study conducted by Pazar et al. (2020); with the artificial intelligence-based camera with the vehicle detection system, vehicle detection was prioritized at the first stage.

Today, with the developing technology, it is an inevitable fact that new methods are derived and developed with different artificial intelligence processes. The flow diagram of the study is given in Figure 1. In this study, unlike other studies in the literature, the results of the visual or manual counts in the analysis and the results of the counts made with the software in the analysis were compared using the software. Obtained results were showed that with a rate of 91.8%, high accuracy of the count was made by the software This comparison was examined in terms of delay, queue length, travel time, service level parameters for intersection performance, and it was aimed to determine the effect of margins of error to occur on the analysis results. For this aim, microsimulation was performed with Vissim Package Software. As a result of the simulation analysis, the vehicle delays, which are 38.82 seconds in the manual count method and 37.36 seconds in the software counting method, are between 35.1-55 seconds and correspond to the service level value D.

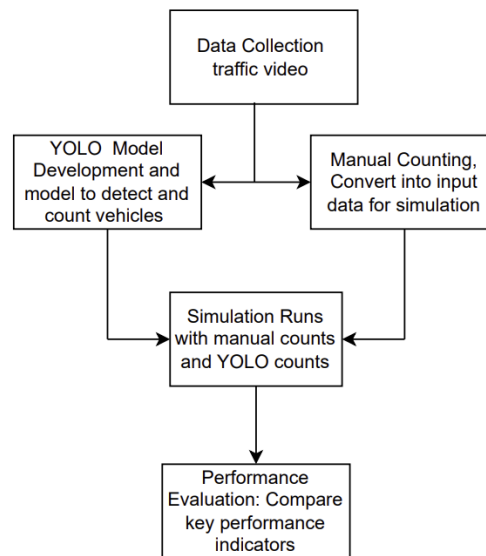


Figure 1. Steps in the study

2. MATERIALS

With the specially developed software, the counting of the vehicles was made according to their types. The section where counting is made is the section where the current arm numbered 3 is located, and counting and analysis have been made only on the straight movements of the vehicles. To control and evaluate these counts, data were also obtained by manually counting from video camera recordings. In the study, to model the intersection, to make performance analysis, and to compare simulation analysis with these data, the PTV Vissim program was used.

2.1. Working Area

Sampi Intersection in Muratpaşa district of Antalya province is located at the junction of Burhanettin Onat Street, Portakal Çiçeği Street, Metin Kasapoğlu Street, Tevfik Işık Street, and Paşa Kavakları Street. The satellite image and traffic flow information of the intersection is shown in Figure 2a-2b . The representation of the intersection in the program is as in Figure 3.

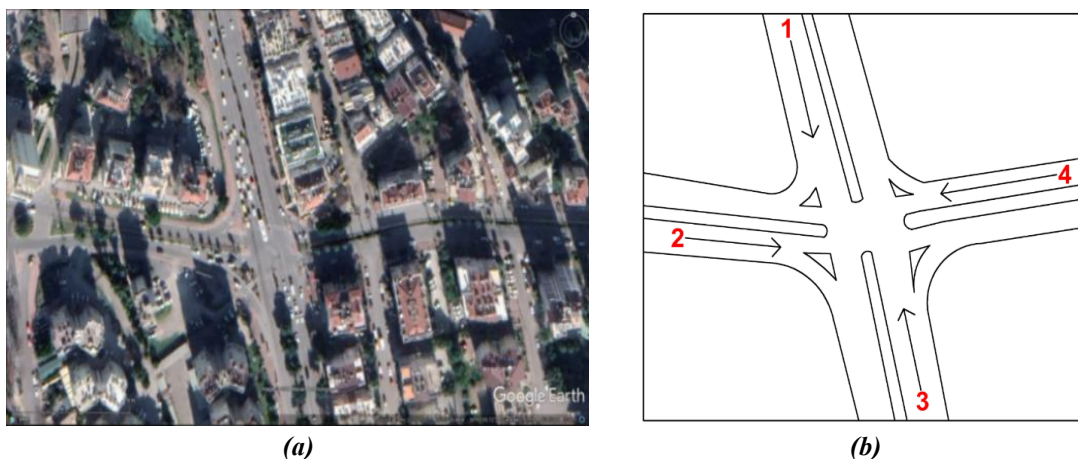


Figure 2. *a) Satellite view, b) Traffic flow plan of Sampi Intersection (Demiral & Köfteci, 2019)*

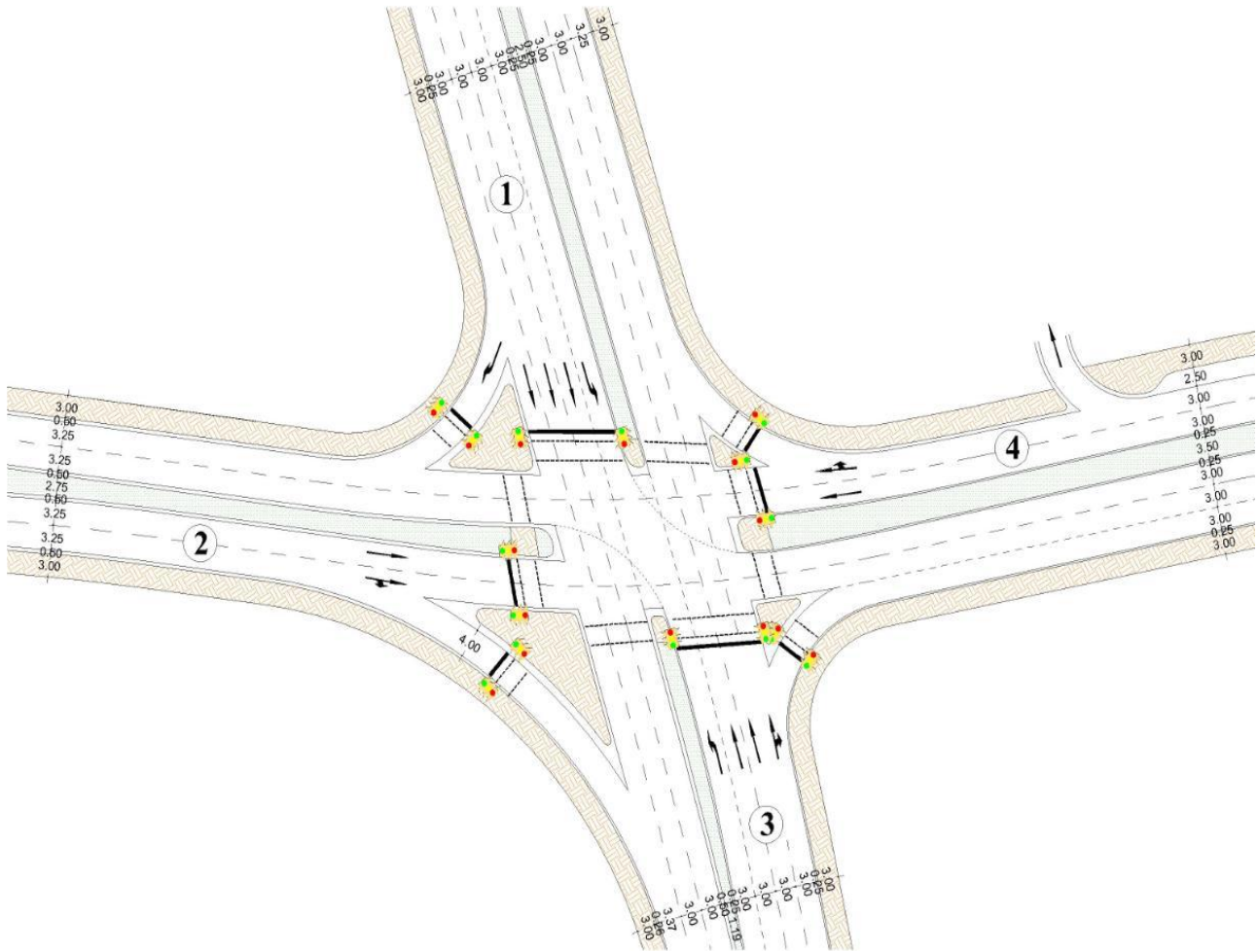


Figure 3. The survey plan of the Sampi Intersection (Demiral & Köfteci, 2019)

There is a small traffic island at right turns for vehicles coming from all 4 directions. Vehicles coming from direction number 1 have the opportunity to turn right, continue straight, and turn left. Vehicles coming from direction number 2 can continue their movements by turning right or going straight. Vehicles coming from direction number 3 have the opportunity to turn right, turn left, and continue their movements straight. Vehicles coming from direction number 4 have the ability to turn right or go straight.

2.2. Works at The Intersection (Camera Setup)

Camera recordings were made between 08.00-09.00 in the morning and 17.00-18.00 in the evening when the traffic density was the most on five days a week. To make these records, first of all, the necessary permissions were obtained by official correspondence with the Antalya Governorship and Antalya Provincial Police Directorate Traffic Inspection Branch Directorate. A mechanism has been established with action cameras that can shoot 360 degrees in the central traffic island at the crossroads. The cameras were raised to 6 meters from the ground using interlocking water pipes, and by fixing cameras peak points to the traffic sign in the central traffic island, shooting quality has been improved, and camera vibration which may occur during recording has been prevented. Camera installation studies can be seen Figure 4.

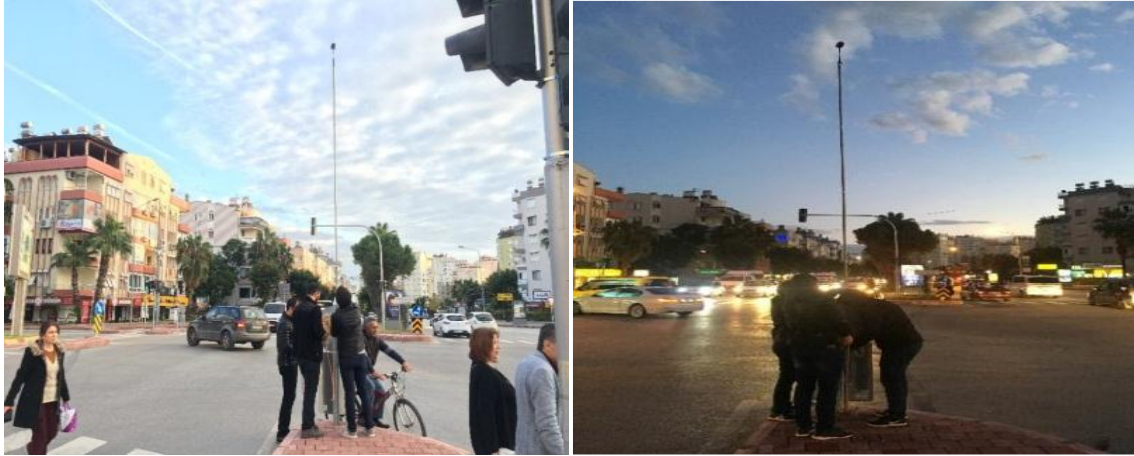


Figure 4. Camera installation studies (Demiral & Köfteci, 2019)

3. METHODS

3.1. Manual Count Method and Data Acquisition

After the camera recording studies, the images were transferred to the computer. Traffic data were obtained by watching these recordings. Figure 5 shows the monitoring of camera images and counting studies.



Figure 5. Monitoring of camera images and counting studies (Demiral & Köfteci, 2019)

This section is the section where the data created by counting and noting them individually by human means of 3 12-minute videos of Sampi intersection, for which all necessary permissions for camera recording have been obtained (Table 1). These counts have been arranged to be used in the analysis to be carried out in the Vissim program. In this program, vehicle types are available as Car, HGV (Heavily Loaded Vehicles), Bus, Tram (Tramway), Truck (Truck-Van), Pedestrian, Bike (Bicycle-Motorcycle). In this study, it was necessary to establish a correlation between the types of vehicles in counting with software, manually counting, and Vissim. Because the effect of vehicles on intersection performance varies. The effects of trucks and automobiles on traffic at any signalized intersection are not the same. To give an example, the effects of trucks and automobiles on traffic at any signalized intersection are not the same. While counting, it is ensured that the intersection simulated in the program is analyzed with values that reflect the reality by classifying according to Vissim vehicle types.

Table 1 Manually counting data

Vehicle types	Manually counting data (According to Vissim vehicle types)	
	Number of vehicles	Vehicle ratio
Car	646	0,7564
HGV	5	0,0059
Bus	176	0,2061
Bike	27	0,0316
Total	854	1

3.2. Software Method and Data Acquisition

Python; is an object-oriented, interpretive, unitary (modular), and interactive high-level programming language (Wikipedia, 2020) . OpenCV (Open Source Computer Vision) is an open-source coded image processing library (Pişkin, 2020). YOLO is an algorithm of the Darknet library that has recently become popular in the field of object detection (Mesci, 2019). The Microsoft Common Objects in Context (MSCOCO) dataset contains 91 common object categories with 82 of them having more than 5,000 labeled in-stances. In total the dataset has 2,500,000 labeled instances in 328,000 images (Lin et al., 2014). Artificial intelligence is the ability of a computer or a computer-controlled robot to perform various activities similar to an intelligent object (Demiral & Köfteci, 2019). Artificial neural networks are a subtitle formed under the artificial intelligence concept. Artificial neural networks collect information about objects and make generalizations. Then make decisions about those examples by using the information they have learned when compared to examples that they have never seen. Artificial neural networks collect information about objects and make generalizations. Then they can make decisions about those examples by using the information they have learned when faced with examples that they have never seen (Sengoz, 2017).

3.2.1. Algorithm and Method Used

If the algorithm of the study is summarized, the way it works is as follows: 1. Detect objects in input data according to the YOLO algorithm and COCO data set. 2. Filter out objects that are work and are above the desired ratio. 3. Since the object detection will be done again in each new video frame, check whether the found objects have been found before or if a new object entered the video frame. 4. According to the information obtained from the user, check these found objects to whether they are suitable for counting, if appropriate check these objects whether they have been counted before, if not, count that object. 5. Take a loop this until the video frames are finished and give us the counted objects and their types as a result. Video is processed frame by frame; a loop rotates in the program for each frame and the commands within that loop are executed. For this reason, each new video frame is freestanding from the previous one according to the program. Hence, the objects in each new video frame are also unique according to the program. Nonetheless,

an object in the video frame that has been processed in reality is its own version, perhaps one second later. To control this situation, the "tracker" method of the "Dlib" library was used in this study.

As seen in Figure 6, the center of gravity of the objects in the video frame where the operation is performed is calculated. The distance between the center of gravity of the objects in the previous frame taken action and the center of gravity of the objects in the new frame is calculated with the help of the hypotenuse. If the center of gravity of the objects in the previous frame is much closer to one of the new centers of gravity than the others, it is figured out that this object is most likely the same as the object in the previous frame and only moves. On the other hand, the object at the center of gravity which is away from all the centers of gravity is probably an object that just entered the video in that video frame. At the same time, the user is asked to determine a counting area and direction, and count control is performed accordingly. After checking whether the found objects are new or are the state that the old object has been replaced, it is checked whether that object is in the counting area. In the counting area, the object is counted, and the "has it been counted" part which is a property of its class is changed to "True" so that in the next frame, if the same object is still in the video frame, it is not checked whether it is counted in vain and no processing load is encountered. At the same time, this "has it been counted" feature is necessary so that a counted object is not counted again.

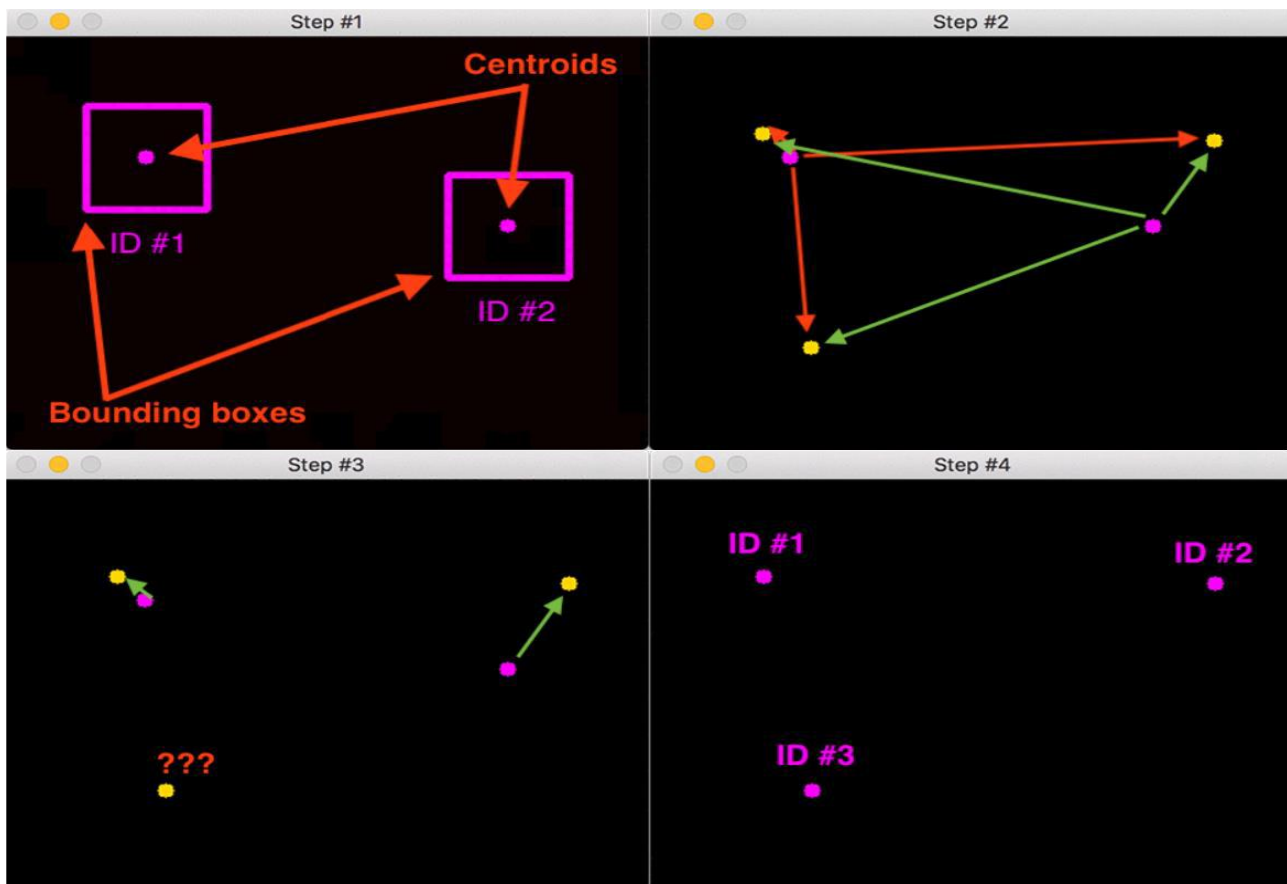


Figure 6. Object tracking between video frames using centroid distance

3.2.2. Software-Video Analysis

The Python version of the OpenCV library has been used for that the video can be read and be performed operations on the video by the software. When the video is read with OpenCV, it is converted into a 3-D matrix in Python. A 3-layer matrix array that each pixel having a value between 0-255 representing Blue, Green, and Yellow tones are formed, respectively. In this way, using algebraic methods and mathematics, the desired operation on this image data comes to a state that can be done. A pre-trained model utilizing the YOLOv3 algorithm and the COCO data set was used for object detection. The weight, configuration, and COCO data set label files of this model were read by using OpenCV's "readNetFromDarkNet ()" method. After a function was written, the picture frame and sensitivity variables were taken as input. The function passes the picture frame it receives through the trained model, and in consequence, according to the weight and configuration files, gives the objects it predicts in the picture as output to us. These estimates found were filtered with a high probability of the sensitivity variable received as input. The estimation percentage of the objects it predicts, the class name, the x, y coordinate of the frame in which the object is located, height, and width information were returned with a Python dictionary variable. In the application where the analysis will be made, a terminal screen like in Figure 7 awaits us. The terminal code arguments were explained in Table 2.

```
> python main.py -i input.mp4 -o outputname -v -l -s
0%|          | 0/919 [00:00<?, ?it/s]
INFO[Drawing]: Left Click: Draw | Right Click: Finish the drawing | Double Right Click: Clean the drawing | 'q': Exit button
INFO[Drawing]: This function return the mask numpy array, original frame and lines of region
INFO[Drawing]: Returned the mask numpy array, original frame and lines of region
INFO[Video]: 919 Total Frames in Video
INFO[Program]: Started Program...
5%|██        | 45/919 [04:01<01:40, 8.67it/s]
```

Figure 7 Terminal screen

After the selection of the counting area, the terminal screen at the start of the analysis is shown in Figure 6.

Table 2. Terminal code arguments

Argument	Explanation
-i, --input	It is the name and extension of the video which is wanted to be analyzed in the "input" folder at the project's main directory. E.g.: inputvideo.mp4
-o, --output	The name of the analysis output video to be saved in the "output" folder in the project's main directory. E.g., outputvideo.mp4
-v, --video360	If the analysis video is a 360-degree video, this is the argument that must be entered.
-r, --resizeframe	If the analysis video is in large resolution and if it is desired to be scaled down, this is the argument that must be entered.
-d, --database	The argument to be entered, if the analysis outputs are wanted to be saved to the database.
-l, --log	It is the argument that must be given if the analysis outputs are wanted to be saved in the project's main directory with a log.txt file.
-s, --savevideo	The argument to be entered, if the analysis is wanted to be saved as a video.

After the information is selected, a sample code block is run as follows (1).

`python main.py -i input.mp4 -o outputname -v -l -s (sample code block)` (1)

After the "Script" runs, a screen like in Figure 8a welcomes us. Direction selection process shown in Figure 8b.



Figure 8. a) Select direction screen, b) Screen after direction selection.

Then, with the help of the cursor, the desired region is selected in a polygon shape, and the process is completed with the "q" key after selection. The mask frames are shown in Figure 9a-9b. At the same time, instructions on how to use the "Select Direction Screen" are printed in the terminal command line window. This information is shown in Figure 10. While selecting the counting area, the mask frame created for the counting algorithm is also drawn on the screen.

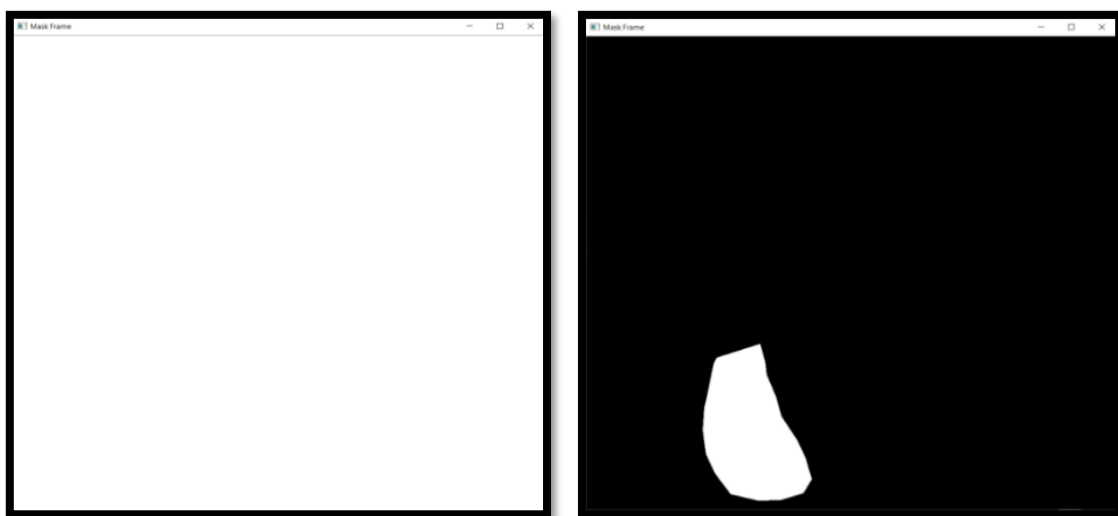


Figure 9. a) Mask frame blank state, b) Mask frame drawn state

```
INFO[Drawing]: Left Click: Draw | Right Click: Finish the drawing | Double Right Click: Clean the drawing | 'q': Exit button
INFO[Drawing]: This function return the mask numpy array, original frame and lines of region
```


Figure 10. Direction selection screen usage instructions

The plotted output of the information of the analyzed video is as in Figure 11.

**Figure 11.** Video output of vehicle tracking, classification, and counting visualization

In an example in Figure 11, "Down" has been selected as direction information, in other words, vehicles going down from the selected red counting area will be counted. Also, vehicles have their type and "ID's" on them. At the same time, there is a line that follows the path they followed in their center of gravity. It becomes that the labels of the vehicles crossing the line and counted are red and their information is "counted". On the upper left, there is information such as the status, the counted vehicle type, the total number of vehicles, in which video frame / total number of frames, at what second / total video duration, respectively. An example of the JSON data in the "log.txt" file, which is the analysis log output (Figure 12), is as follows:

```
{ "bicycle":0,
  "motorbike":1,
  "car":1,
  "bus":0,
  "truck":0,
  "total_cars":2,
  "video_sec":30.63,
  "video_min":0.51,
  "video_name":"input-mp4",
  "total_analysis_time_sec":39.6,
  "total_analysis_time_min":0.66 }
```

Figure 12. *Analysis of log output*

"bicycle", "motorbike", "car", "bus", "truck" give information about how many vehicles are counted in these types. The "total_cars" key gives information about how many vehicles have been counted in total so far. The "video_sec", "video_min", "video_name" keys give information such as the seconds, minute, and the name of the video, respectively. The "total_analysis_time_sec", "total_analysis_time_min" keys give the seconds and minute equivalents of the analysis time, respectively. If the “—database” argument is given while the script is running, the JSON information is also saved in the database as in “log.txt” above. After all these processes are completed, the counting results of software image analysis are given in Table 3.

Table 3. *Counting data made with the software*

Vissim vehicle types	Counting data made with the software	
	Number of vehicles	Vehicle ratio
Car	532	0,6786
HGV	4	0,0051
Bus	239	0,3048
Bike	9	0,0115
Total	784	1

3.3. Intersection Performance Analysis

For the intersection analysis, software counting & visual counting traffic data obtained from camera recordings were used. In this study, PTV Vissim micro-scale traffic simulation software was used for traffic analysis.

PTV Vissim is a program that can perform highway simulation, intersection simulation, and public transportation simulations together. According to different traffic situations, using simulation methods is an important feature in terms of detecting the problem in advance (Akbaş, 2001). The traffic flow model in VISSIM is a not only discrete but also stochastic, time-step-based microscopic model. In this model, driver-vehicle units are accepted as single entities (Fellendorf & Vortisch, 2001).

Simulators are needed, to assess the benefits of ITS in a planning mode, to generate scenarios, and to predict network behavior at the operational level. This type of usage by traffic engineers can give them an overall picture of the traffic as well as the ability to assess current problems and project possible solutions immediately (Boxill & Yu, 2000). Simulation studies are of great importance when planning transportation. Because it is possible with simulation programs to perform analyzes that are difficult or even impossible to do with analytical and numerical methods and to produce scenarios related to today and the future by revealing effective visual results (Güler, 2017). The characteristics of computer-aided traffic analysis programs were given in Table 4.

Table 4. Computer-aided traffic analysis programs (MassDot, 2012)

Program	Unsignalized analysis	Signalized analysis	Roundabout analysis	Signal optimization	Network analysis	Simulation
HCS 2010 Ver 6.1	Yes	Yes	No	Yes	No	No
Synchro Ver 8.0	Yes	Yes	No	Yes	Yes	No
SIDRA Intersection Ver 5.1	Yes	Yes	Yes	Yes	No	No
Transyt-7F Release 11.32	No	No	No	Yes	Yes	Yes
SimTraffic Ver 8.0	Yes	Yes	No	No	Yes	Yes
TSIS-CORSIM Ver 6.0	No	No	No	No	Yes	Yes
GDOT Roundabout Analysis Tool Ver 1.1	No	No	Yes	No	No	No
Vissim Ver 5.40	Yes	Yes	Yes	Yes	Yes	Yes

The service levels of intersections and roads are the most important criteria for the quality and performance of the road. This criterion may vary according to the topographic conditions of the region where the road is located. Performance parameters are evaluated by the General Directorate of Highways based on HCM (Highway Capacity Manual). Service level class is more flexible than sub-parameters such as delay, queue length. In the analysis results, the PTV Vissim program determines the level of service (LOS) class according to the VehDelay (All) / Vehicle delay parameter.

Table 5. Level of service criteria (An et al., 2015)

LOS by Volume-to-Capacity Ratio		Signalized Intersection	Roundabout
$v/c \leq 1.0$	$v/c > 1.0$	Control Delay (s/veh)	Control Delay (s/veh)
A	F	≤ 10	0-10
B	F	>10-20	>10-15
C	F	>20-35	>15-25
D	F	>35-55	>25-35
E	F	>55-80	>35-50
F	F	>80	>50

It is shown that the connection between service level classification and vehicle delay (sec) at signalized intersections in Table 5. If the vehicle delay parameter value is 10 seconds or less, the service level class is A. If it is from 10 seconds to 20 seconds, the level of service is B; if it is from 20 seconds to 35 seconds, it is C; if it is from 35 seconds to 55 seconds, it is D; if it is from 55 seconds to 80 seconds, the intersection is at E

level of service. If the vehicle delay parameter is more than 80 seconds, the intersection service level class is determined as F; and it is the lowest rank of the level of service. In this study, average tail length, maximum tail length, vehicle delay, stop delay, pauses, and service level is considered as analysis parameters.

Table 6. *Vehicle counting with software and manual way*

Vissim vehicle types	Counting data made with the software		Manually counting data	
	Number of vehicles	Vehicle ratio	Number of vehicles	Vehicle ratio
Car	532	0,6786	646	0,7564
HGV	4	0,0051	5	0,0059
Bus	239	0,3048	176	0,2061
Bike	9	0,0115	27	0,0316
Total	784	1	854	1

When looking at Table 6, it is seen that a total number of 784 vehicles are counted in the counting made with the software, and 854 vehicles are counted in the manual counts. This shows that with a rate of 91.8%, high accuracy of the count was made by the software. There are several reasons why we may not have the same accuracy in both methods at vehicle classification subject. The most important of these is that the analyzed video data has been taken from a 360-degree camera. Therefore, as the objects approach the counting area, they come out of the real image due to the 360-degree feature of the camera. Artificial intelligence, trained according to the data given to it, has difficulties in perceiving these objects. The second reason is that the artificial intelligence model used in this study is a pre-trained model that is not specifically trained for the region-specific tools analyzed. The third reason is that as a result of the camera being placed at a high position so that it can see the entire intersection, vehicles such as the "Bike" class remain very small in the image given to the artificial intelligence. Therefore, it is more difficult for artificial intelligence to distinguish such classes than normal. Solution suggestions for these reasons are presented in the "Recommendations" section.

The interface of the Vissim program is shown in Figure 13. Sampi Intersection has been modeled in terms of signal durations, lane numbers, lane connections, lane widths, and vehicle mobility capability. (Figure 14)

For the Sampi Intersection, as a result of the counting study intended the current coming from the selected arm, flow data for this arm were entered in the Vissim program. This section is the section where the vehicle components, vehicle types, desired speed preferences, and the ratio in the total number of vehicles of the vehicle types are entered. The entry to Vissim of the data of the counts made manually is shown in Figure 15a. In the manual count, a total of 854 vehicles were counted in 3 videos. These vehicles are separated as Car, HGV, Bus, and Bike by classifying them by the way appropriate to Vissim. The rate of these vehicle types is shown in Figure 16a. It is made that vehicle entry into Vissim according to the count made with software. (Figure 15b) The vehicle ratios, according to their types have been entered into Vissim by calculating which vehicle type how much of the number of incoming traffic belongs to. (Figure 16b)

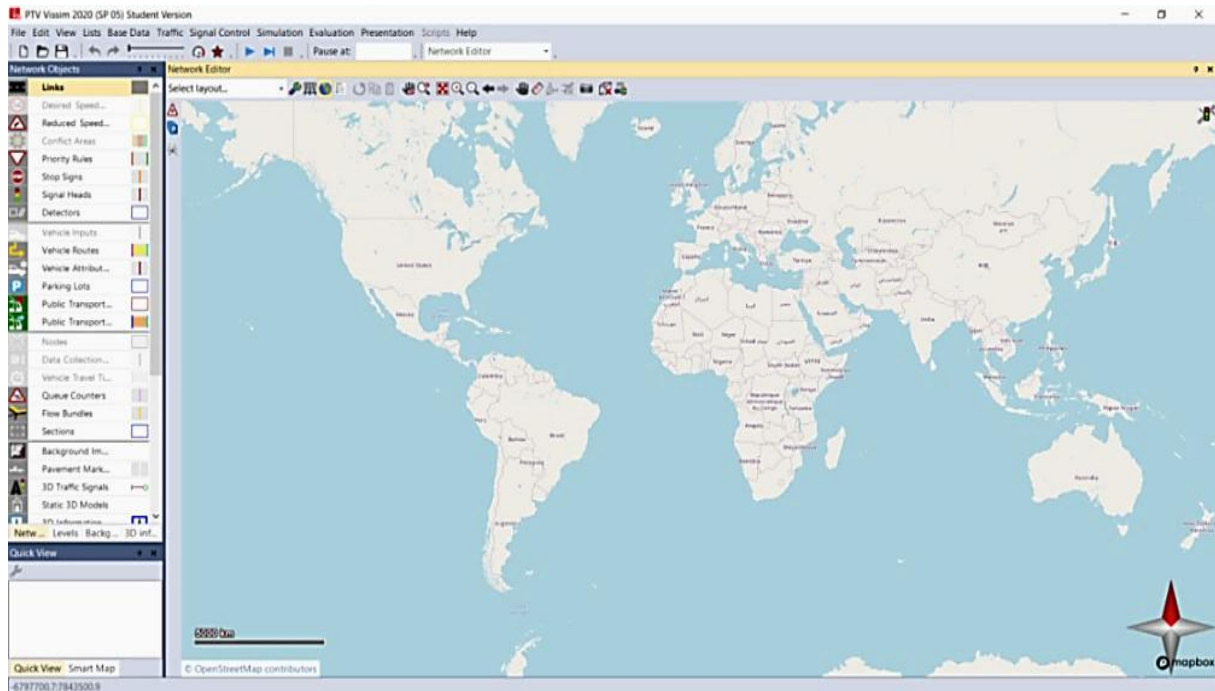


Figure 13. Vissim program interface



Figure 14. Sampi Intersection Vissim model

Since the 'node result' section will be checked for analysis, the boundaries of the area to be analyzed are determined from the 'nodes' button (Figure 17). Analysis of the current from number 3 arm of the intersection examined with 'Nodes results' has been made in terms of parameters such as QLen, QLenMax, LOS (All), VehDelay (All), StopDelay (All), and Stops (All). A three-dimensional screenshot of the traffic is given in Figure 18.

Cou	No	Name	Link	Volume(0)	VehComp(0)
1	1		4: KOL3	854,0	1: Default
2	2		5: KOL4	0,0	1: Default
3	3		3: KOL2	0,0	1: Default
4	4		1: KOL1	0,0	1: Default

(a)

Cou	No	Name	Link	Volume(0)	VehComp(0)
1	1		4: KOL3	784,0	1: Default
2	2		5: KOL4	0,0	1: Default
3	3		3: KOL2	0,0	1: Default
4	4		1: KOL1	0,0	1: Default

(b)

Figure 15. a) Manual count data vehicle entry, b) Vehicle entry with software counting data

Count: 4	VehType	DesSpeedDistr	RelFlow
1	100: Car	70: 70 km/h	0,7564
2	200: HGV	50: 50 km/h	0,0059
3	300: Bus	50: 50 km/h	0,2061
4	610: Bike Man	60: 60 km/h	0,0316

(a)

Count: 4	VehType	DesSpeedDistr	RelFlow
1	100: Car	70: 70 km/h	0,6786
2	200: HGV	50: 50 km/h	0,0051
3	300: Bus	50: 50 km/h	0,3048
4	610: Bike Man	60: 60 km/h	0,0115

(b)

Figure 16. a) Vehicle rates entry according to the counting data made manually, b) Vehicle rates entry according to the counting data made with the software



Figure 17. For the results, determining the frontiers with the 'Nodes' command

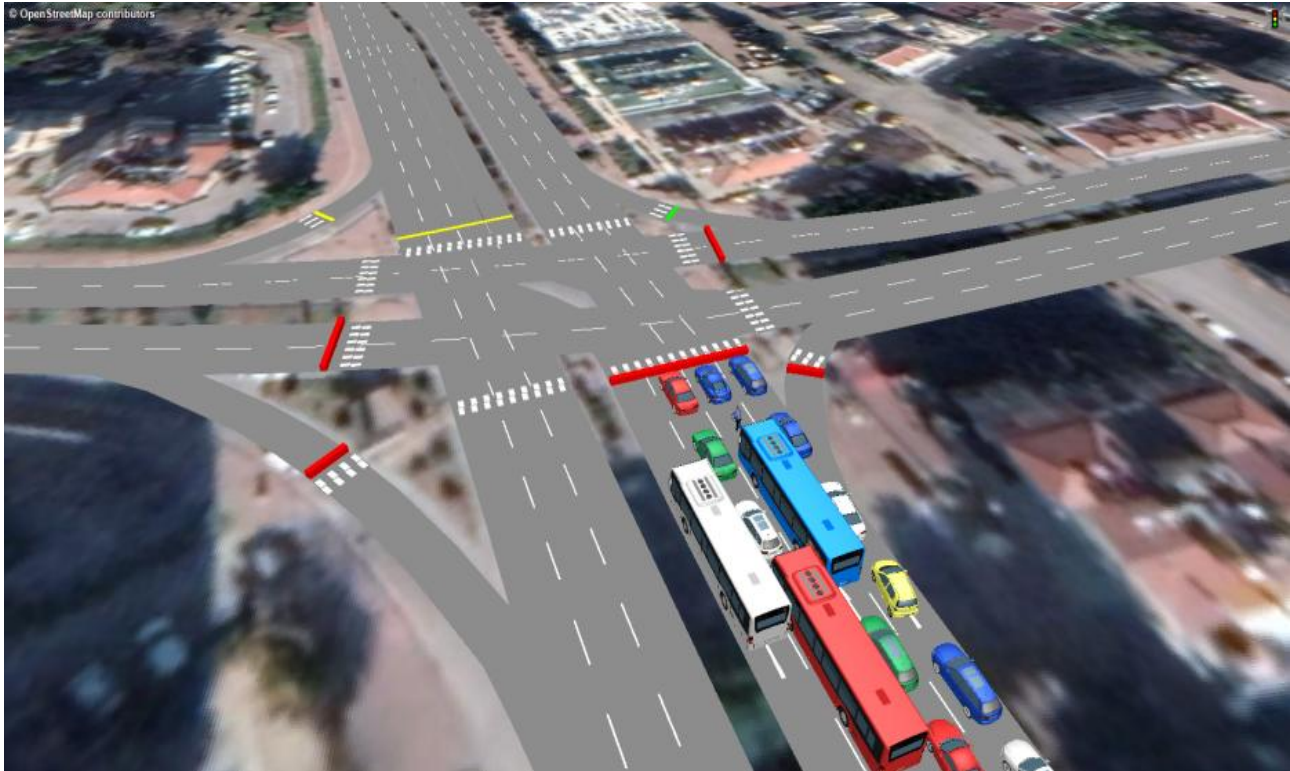


Figure 18. A three-dimensional screenshot of the traffic, taken during analysis

4. FINDINGS AND DISCUSSIONS

According to the manual count data, the average queue length (QLEN), maximum queue length (QLEN MAX), vehicle delay (VEH DELAY), stopping delay (STOP DELAY), pauses (STOPS) in the simulation analysis for the direct straight movement of the arm current number 3 are shown in Table 7.

Table 7. Simulation analysis results according to manual count and software count data

Data types	TIMEINT	QLEN (m)	QLEN MAX (m)	LOS (ALL)	VEH DELAY (ALL) (sec)	STOP DELAY (ALL) (sec)	STOPS (ALL)
Manual count data results	0-3600	28,4	86,42	LOS_D	38,82	29,43	70,73
Software data analysis results	0-3600	29,96	86,42	LOS_D	37,36	27,89	68,38

As a result of these values in Table 7, as Level of Service (LOS) is calculated according to the vehicle delay in signalized intersections, the service level is determined as D as shown in Table 5. The Vissim program also indicated the service level according to vehicle delay as D as a result of the analysis.

According to the software count data, the average queue length (29.96 m), the maximum queue length (86.42 m), the vehicle delay (37.36 m), the stopping delay is (27.86 m), the pauses are (%68.38) are shown in Table 7. Since the vehicle delay parameter is 37.36 seconds, the Vissim program determined the service level as D class. The results of the analysis are shown in Figure 19a-19b as Vissim screenshots and Figure 20a-20b. Figure 21 shows the place of vehicle delays due to counting with software and manual counting among service level classes.

According to Table 7, between manual counting method and software counting method; with 1.56 m as 5.21% for queue length parameter, 0% for maximum queue length parameter, with 1.46 sec as 3.76% for vehicle delay parameter, with 1.54 sec as 5.23% for stop delay parameter, 2.35% error rates for the stops parameter are seen. Charef et al. (2024) used the YOLO algorithm to detect drivers wearing helmets. The counting error between the algorithm and manual measurement of values periodized by hours was at most 20%. Hashmi et al. (2024) counted by day and vehicle type and found that the margin of error of the model was at most 20% in the morning. They found this value for large vehicle types compared to small counts. In this study, the error values found were smaller in percentage terms. This is thought to be due to the use of a wide range of counting data and the fact that the study was carried out under favorable conditions.



Figure 19. a) Vissim analysis results according to manual count data, b) Vissim analysis results according to counting data with software

The number of vehicles and their classes was obtained by manual counting and by analyzing images with the software. Afterward, vehicle simulations were made to present the data obtained from these two methods and Vissim vehicle classes in a common language. Sampi Intersection was drawn and modeled in Vissim according

to a satellite image, traffic flow plan, and survey plan. After studies on the connection of vehicle types with each other, software count data and manual count data were arranged according to Vissim vehicle types, and the number and types of vehicles were entered into the Vissim program.

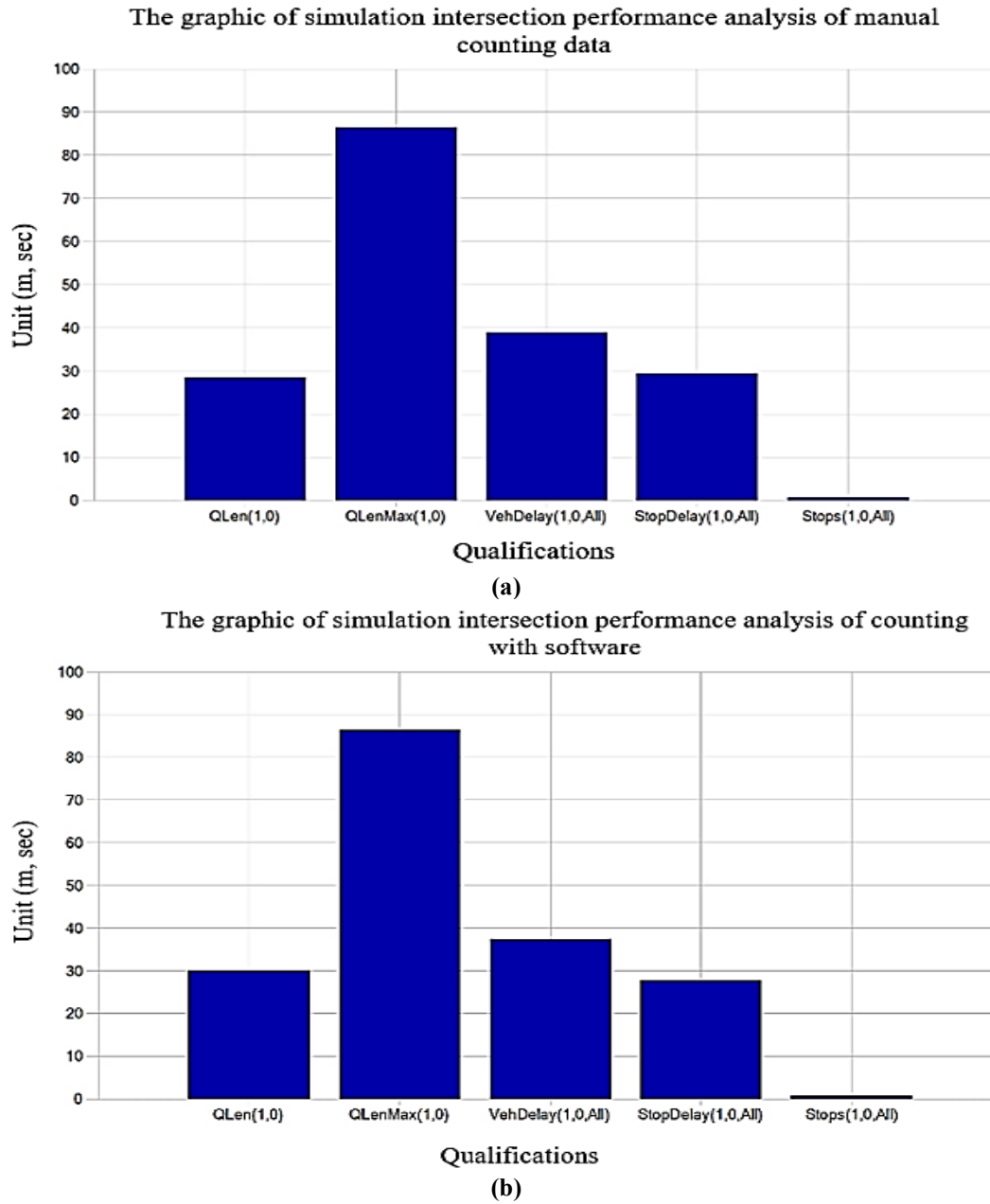


Figure 20. *a) The graphic of simulation intersection performance analysis of manual counting data, b) The graphic of simulation intersection performance analysis of software counting data*

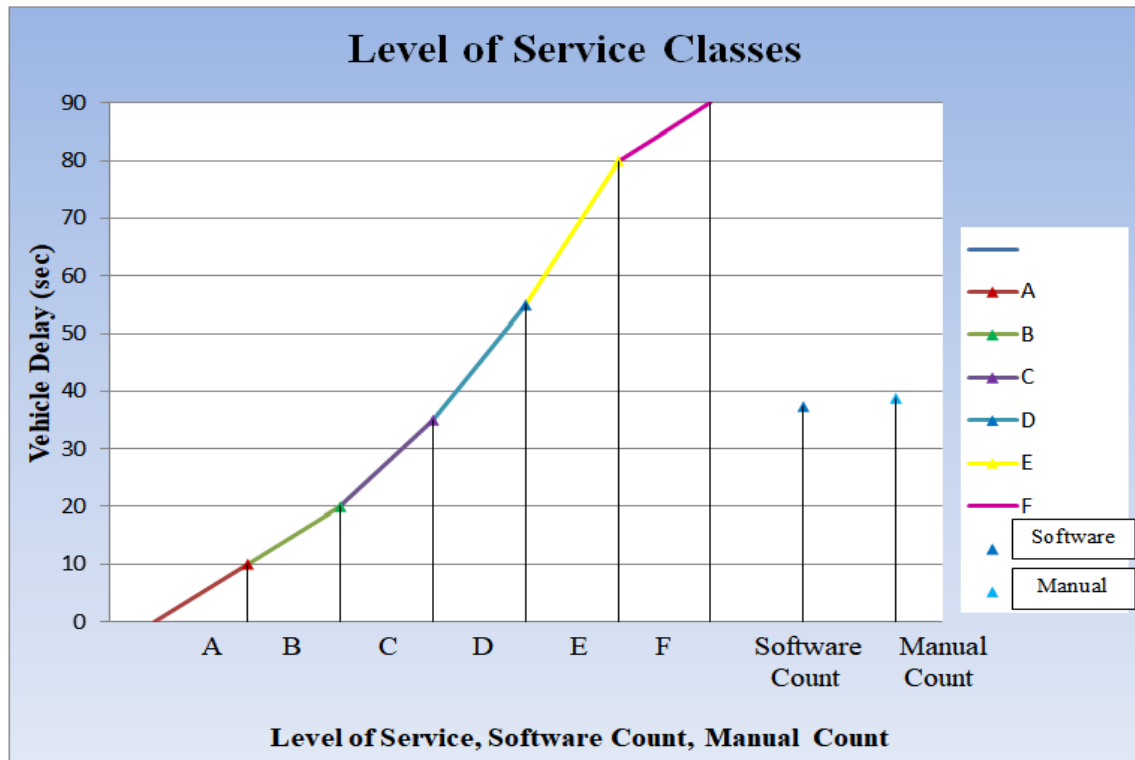


Figure 21. The place of vehicle delays due to counting with software and manual counting among service level classes

According to the results of this study, while 854 vehicles have been counted by the manual counting method, it has been counted 784 vehicles by the software. There are reasons for the differences that occur between vehicle numbers and types. These reasons;

- i. There are errors in the differentiation of car and bus vehicle types due to the change in the size of the video images according to the distance of the vehicles in motion to the area selected in the software,
- ii. The types of vehicles included in the vehicle classification of the software may differ according to the vehicle types of our country,
- iii. And some vehicles cannot be recognized or counted from the area specified for counting.

5. CONCLUSIONS

In this study, 360-degree camera images of Sampi Intersection in Antalya province were analyzed using image processing technologies and artificial neural network models on Python. In the image analysis, the counting and classification of the vehicles continuing to move straight from the current arm number 3 of the intersection were made. It was used manual counting as another method. The camera images were watched, and the counting and classification of the vehicles were done manually. The purpose of the study is the observation and the demonstration of how much the effect of the software data on the simulation program analysis results makes a difference according to the simulation program analysis results of the data obtained from a grueling but accurate method such as manual counting. As a result of the simulation analysis, the vehicle delays, which are 38.82 seconds in the manual count method and 37.36 seconds in the software counting method, are between

35.1-55 seconds and correspond to the service level value D. Since the software-induced error rate is at the level not to affect the result of the intersection performance analysis, it has been proven that the artificial intelligence and vehicle count method is an acceptable method in traffic engineering.

6. RECOMMENDATIONS

The reasons for not achieving the same rate in classification, despite having a high accuracy rate in counting as a result of the study, are specified under Table 6. To achieve a better accuracy rate in classification and therefore in counting, it is recommended to retrain the artificial neural network with a visual data set of significant size for the counting region tools and a computer with powerful hardware. Also, if the environment where the artificial intelligence can better distinguish the objects can be provided on the side of the camera where the video data is taken, this accuracy rate will increase even more. It is important for image analysis to obtain quality images for studies on these subjects and software databases to have a large library. It is important that the vehicle classification of software can be configured to comply with the conditions of Turkey for a correct evaluation of the data. It can be provided that image analysis is done with high performance with good technological hardware.

It can be benefited from simulation programs for analysis will be done with the purpose of determination of intersection performance. With the development of studies, it can be ensured that counts are successfully performed with artificial intelligence on vehicle count and classification. Thus, faster and more reliable data can be obtained by reducing the human factor and workload. Because of this way, by performing the intersection analyzes in a fast and healthy way, the prospective project results can be predicted in a short time, information can be obtained by making retrospective analyzes, and situation assessment can be made with a current situation analysis.

FUNDING

This research was supported by the “Akdeniz University Scientific Research Projects Department” with the Project code: FYL-2018-3590

AUTHOR CONTRIBUTIONS

Conceptualization, S.K. and Y.A.; methodology, S.K. and Y.A.; fieldwork, A.D.; software, B.E.U. and D.H.; title, S.K. and Y.A.; validation, M.A.Y., B.E.U. and D.H.; formal analysis, S.K., Y.A. and M.A.Y.; research, B.E.U. and D.H.; sources, M.A.Y., B.E.U., Y.A. and D.H.; data curation, S.K., Y.A., B.E.U., D.H.; manuscript-original draft, S.K., Y.A., B.E.U., D.H.; manuscript-review and editing, S.K., Y.A., M.A.Y.; visualization, B.E.U. and D.H.; supervision, S.K., Y.A.; project management, S.K., Y.A.; funding, S.K., Y.A. All authors have read and legally accepted the final version of the article published in the journal.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

- Akbaş, A. (2001). *Kent İçi Ulaşımında Ana Arterlerdeki Ulaşım Performansının Simülasyon Tabanlı Olarak Değerlendirilmesi*. Ulaştırma Kongresi, 75–86.
- Altun, İ., DüNDAR, S., & Yöntem, K. (2005). *Yapay sinir ağları ile trafik akım kontrolü*. Deprem Sempozyumu, 23–25.
- An, H. K., Yue, W. L., & Kim, D. S. (2015). A Proposal of Two Signals Roundabout Analysis Method Using SIDRA6. *KSCE Journal of Civil and Environmental Engineering Research*, 35(5), 1111–1121. <https://doi.org/10.12652/KSCE.2015.35.5.1111>
- Bao, X., Li, H., Xu, D., Jia, L., Ran, B., & Rong, J. (2016). Traffic Vehicle Counting in Jam Flow Conditions Using Low-Cost and Energy-Efficient Wireless Magnetic Sensors. *Sensors*, 16(11), 1868. <https://doi.org/10.3390/S16111868>
- Bhosale, M. K., Patil, S. B., & Patil, B. B. (2023). Automatic Video Traffic Surveillance System with Number Plate Character Recognition Using Hybrid Optimization-Based YOLOv3 and Improved CNN. *International Journal of Image and Graphics*, 25(4), 2550041. <https://doi.org/10.1142/S021946782550041X>
- Boxill, S. A., & Yu, L. (2000). An evaluation of traffic simulation models for supporting its. *Houston, TX: Development Centre for Transportation Training and Research, Texas Southern University*.
- Charef, A., Jarir, Z., & Quafafou, M. (2024, May 2-4). *Enhancing Road Safety: Automated Traffic Violation Detection and Counting System Using YOLO Algorithm*. Proceedings of the 2024 Mediterranean Smart Cities Conference (MSCC), Morocco. <https://doi.org/10.1109/MSCC62288.2024.10697076>
- Chung, J., & Sohn, K. (2018). Image-Based Learning to Measure Traffic Density Using a Deep Convolutional Neural Network. *IEEE Transactions on Intelligent Transportation Systems*, 19(5), 1670–1675. <https://doi.org/10.1109/TITS.2017.2732029>
- Cruz, F. R. G., Santos, C. J. R., & Veal, L. A. (2019, November 29 - December 1). *Classified Counting and Tracking of Local Vehicles in Manila Using Computer Vision*. Proceedings of the 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), Philippines. <https://doi.org/10.1109/HNICEM48295.2019.9072808>
- Dai, Z., Song, H., Wang, X., Fang, Y., Yun, X., Zhang, Z., & Li, H. (2019). Video-based vehicle counting framework. *IEEE Access*, 7, 64460–64470. <https://doi.org/10.1109/ACCESS.2019.2914254>
- Demiral, A. C., & Köfteci, S. (2019). Analysis of intersection performance with package program: Antalya Muratpaşa Sampi intersection example. *International Journal of Environmental Science and Technology*, 16(9), 5319–5324. <https://doi.org/10.1007/s13762-019-02415-2>
- El-Khoreby, M. A., Abu-Bakar, S. A. R., Mokji, M. M., Omar, S. N., & Malik, N. U. R. (2019, September 17-19). *Localized Background Subtraction Feature-Based Approach for Vehicle Counting*. Proceedings of the 2019 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), Malaysia, (pp. 324–328). <https://doi.org/10.1109/ICSIPA45851.2019.8977795>

- Fellendorf, M., & Vortisch, P. (2001). Validation of the microscopic traffic flow model VISSIM in different real-world situations. *Transportation Research Board 80th Annual Meeting*, 11.
- Findley, D. J., Cunningham, C. M., & Hummer, J. E. (2011). Comparison of mobile and manual data collection for roadway components. *Transportation Research Part C: Emerging Technologies*, 19(3), 521–540. <https://doi.org/10.1016/J.TRC.2010.08.002>
- Ge, L., Dan, D., Koo, K. Y., & Chen, Y. (2023). An improved system for long-term monitoring of full-bridge traffic load distribution on long-span bridges. *Structures*, 54, 1076–1089. <https://doi.org/10.1016/J.ISTRUC.2023.05.103>
- Güler, H. (2017). A new approach for road traffic accidents: Crash analysis segments model. *Pamukkale University Journal of Engineering Sciences*, 23(6), 707–717. <https://doi.org/10.5505/pajes.2016.81542>
- Gupte, S., Masoud, O., Martin, R. F. K., & Papanikolopoulos, N. P. (2002). Detection and Classification of Vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 3(1), 37–47. <https://doi.org/10.1109/6979.994794>
- Hashmi, H. T., Ud-Din, S., Khan, M. A., Khan, J. A., Arshad, M., & Hassan, M. U. (2024). Traffic Flow Optimization at Toll Plaza Using Proactive Deep Learning Strategies. *Infrastructures*, 9(5), 87. <https://doi.org/10.3390/INFRASTRUCTURES9050087>
- Indolia, S., Goswami, A. K., Mishra, S. P., & Asopa, P. (2018). Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach. *Procedia Computer Science*, 132, 679–688. <https://doi.org/10.1016/J.PROCS.2018.05.069>
- Kolluri, J., & Das, R. (2023). Intelligent multimodal pedestrian detection using hybrid metaheuristic optimization with deep learning model. *Image and Vision Computing*, 131, 104628. <https://doi.org/10.1016/J.IMAVIS.2023.104628>
- Kulkarni, M. M., Chaudhari, A. A., Srinivasan, K. K., Chilukuri, B. R., Treiber, M., & Okhrin, O. (2025). Leader–follower identification with vehicle-following calibration for non-lane-based traffic. *Transportation Research Part C: Emerging Technologies*, 171, 104940. <https://doi.org/10.1016/J.TRC.2024.104940>
- Li, S., Chang, F., Liu, C., & Li, N. (2020). Vehicle counting and traffic flow parameter estimation for dense traffic scenes. *IET Intelligent Transport Systems*, 14(12), 1517–1523. <https://doi.org/10.1049/IET-ITS.2019.0521>
- Lin, J. P., & Sun, M. Te. (2018, November 30 – December 2). *A YOLO-Based Traffic Counting System*. Proceedings of the 2018 Conference on Technologies and Applications of Artificial Intelligence (TAAI), Taiwan. (pp. 82–85). <https://doi.org/10.1109/TAAI.2018.00027>
- Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., & Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 8693 LNCS(PART 5), 740–755. https://doi.org/10.1007/978-3-319-10602-1_48

- Ma, X., Wei, W., Dong, J., Zheng, B., & Ma, J. (2023, June 18-23). RTOD-YOLO: *Traffic Object Detection in UAV Images Based on Visual Attention and Re-parameterization*. Proceedings of the 2023 International Joint Conference on Neural Networks (IJCNN), Australia. <https://doi.org/10.1109/IJCNN54540.2023.10191514>
- MassDot (2012, January 5). *A Guide on Traffic Analysis Tools Revised*. https://nanopdf.com/download/a-guide-on-traffic-analysis-tools-revised-october-5-2012_pdf
- Mesci, Y. (2019, April 29). *YOLO Algoritmasını Anlamak. Son yıllarda nesne tespiti alanında...* | by Yiğit Mesci | *Deep Learning Türkiye* | Medium. (Accessed: 11/04/2025) <https://medium.com/deep-learning-turkiye/yolo-algoritmas%C4%B1n%C4%B1-anlamak-290f2152808f>
- Pazar, Ş., Bulut, M., Uysal, C., (2020). Yapay Zeka Tabanlı Araç Algılama Sistemi Geliştirilmesi, *Bilim, Teknoloji ve Mühendislik Araştırmaları Dergisi*, 1(1), 31-37.
- Oltean, G., Florea, C., Orghidan, R., & Oltean, V. (2019, October 23-26). *Towards Real Time Vehicle Counting using YOLO-Tiny and Fast Motion Estimation*. Proceedings of the 2019 IEEE 25th International Symposium for Design and Technology in Electronic Packaging (SIITME), Romania, (pp. 240–243). <https://doi.org/10.1109/SIITME47687.2019.8990708>
- Paľo, J., Caban, J., Kiktová, M., & Černický. (2019). The comparison of automatic traffic counting and manual traffic counting. *IOP Conference Series: Materials Science and Engineering*, 710(1), 012041. <https://doi.org/10.1088/1757-899X/710/1/012041>
- Papageorgiou, G. N. (2006). Towards a Microscopic Simulation Model for Traffic Management: A Computer-Based Approach. *IFAC Proceedings Volumes*, 39(12), 403–411. <https://doi.org/10.3182/20060829-3-NL-2908.00070>
- Pişkin, M. (2020, May 27). *OpenCV Nedir? Bileşenleri ve Alternatifleri* | Mesut Pişkin | Blog. <https://mesutpiskin.com/blog/opencv-nedir.html>
- Sengoz, N. (2017, Mart 4). *Yapay Sinir Ağları - Derin Öğrenme* | Deep Learning. <https://www.derinogrenme.com/2017/03/04/yapay-sinir-aglari/>
- Sivaraman, S., & Trivedi, M. M. (2013). Looking at vehicles on the road: A survey of vision-based vehicle detection, tracking, and behavior analysis. *IEEE Transactions on Intelligent Transportation Systems*, 14(4), 1773–1795. <https://doi.org/10.1109/TITS.2013.2266661>
- Sooksatra, S., Yoshitaka, A., Kondo, T., & Bunnun, P. (2019, November 26-29). *The density-aware estimation network for vehicle counting in traffic surveillance system*. Proceedings of the- 2019 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), Italy, (pp. 231–238). <https://doi.org/10.1109/SITIS.2019.00047>
- Toth, C., Suh, W., Elango, V., Sadana, R., Guin, A., Hunter, M., & Guensler, R. (2013). Tablet-Based Traffic Counting Application Designed to Minimize Human Error. *Transportation Research Record*, 2339, 39–46. <https://doi.org/10.3141/2339-05>
- Wikipedia (2020). Python. <https://tr.wikipedia.org/wiki/Python>

- Yao, L. (2019, December 9-11). *An Effective Vehicle Counting Approach Based on CNN*. Proceedings of the 2019 IEEE 2nd International Conference on Electronics and Communication Engineering (ICECE), China, (pp. 15–19). <https://doi.org/10.1109/ICECE48499.2019.9058582>
- Yıldırım, Z. B., Özdemir, B. E., & Eren, E. (2019). *Trafikteki Araç Sayımlarının Farklı Görüntü İşleme Teknikleri Kullanılarak Karşılaştırılması*. Proceedings of the 2nd International Congress on Engineering and Agriculture, (pp. 242–248).
- Zheng, P., & Mike, M. (2012). An Investigation on the Manual Traffic Count Accuracy. *Procedia - Social and Behavioral Sciences*, 43, 226–231. <https://doi.org/10.1016/J.SBSPRO.2012.04.095>

JOURNAL OF SCIENCE

PART A: ENGINEERING AND INNOVATION



| Correspondence Address |

Gazi University
Graduate School of Natural and Applied Sciences
Emniyet Neighborhood, Bandırma Avenue
No:6/20B, 06560, Yenimahalle - ANKARA
B Block, Auxiliary Building

| e-mail |

gujsa06@gmail.com

| web page |

<https://dergipark.org.tr/tr/pub/gujsa>

e-ISSN 2147-9542