

INNOVATIVE APPLICATIONS AND RESEARCH METHODS IN ENGINEERING SCIENCES



EDITORS

PROF. BIROL KILIÇ, PH.D.

ASSIST. PROF. SITKI ALPER OZDEMIR, PH.D.



DOI: 10.5281/zenodo.18094294

Innovative Applications and Research Methods in Engineering Sciences

Editors

Prof. Birol Kılıç, Ph.D.

Assist Prof. Sıtkı Alper Özdemir, Ph.D.

Publisher

Platanus Publishing®

Editors in Chief

Prof. Birol Kılıç, Ph.D. & Assist. Prof. Sıtkı Alper Özdemir, Ph.D.

Cover & Interior Design

Platanus Publishing®

The First Edition

December, 2025

ISBN

978-625-8513-29-5

©copyright

All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, or any information storage or retrieval system, without permission from the publisher.

Platanus Publishing®

Address: Natoyolu Cad. Fahri Korutürk Mah. 157/B, 06480, Mamak,
Ankara, Turkey.

Phone: +90 312 390 1 118

web: www.platanuspublishing.com

e-mail: platanuskita@gmail.com



Platanus Publishing®

CONTENTS

CHAPTER 1	7
<i>Impulse Response of Solid and Cut-Out Plates Supported by Corner Spring-Damper Systems</i>	
Can Gönenli & Okan Duymazlar	
CHAPTER 2	19
Advances and Mechanization in Capture Fisheries	
F. Ozan Düzbastılar	
CHAPTER 3	55
Microplastics in Cosmetics and Personal Care Products	
Sebnem Gun	
CHAPTER 4	69
Türkiye's Mobile Communication Technology Journey: An Anatomy of Digital Transformation from 3G to 5G	
Ali Özhan Akyüz & Onur İnan	
CHAPTER 5	87
A Simple Interactive Simulation and Learning Platform for Automotive Braking Systems	
Ali Özhan Akyüz & Kazım Kumaş & Mustafa Ayan& Durmuş Temiz	
CHAPTER 6	115
A Digital Platform for Examining the Biomedical Foundations and Clinical Accuracy of Blood Gas Analysis	
Ali Özhan Akyüz & Kazım Kumaş & Mustafa Ayan& Durmuş Temiz	
CHAPTER 7	135
Experimental PID Tuning for Magnetometer-Based Heading Control of an Autonomous Mobile Robot	
Osman Ünal	
CHAPTER 8	153
The Impact of Transformational Leadership on Sustainability and Business Performance Through ERP Systems and Supply Chain Integration	
İrem Candaş & Sertan Çal & Batin Latif Aylak & Mehmet Ali Taş	

CHAPTER 9	197
Scenario-Based Production and Capacity Planning Under Uncertainty: Pharmaceutical Industry	
Yağmur Erkan & Mehmet Ali Taş & Batin Latif Aylak	
CHAPTER 10	229
Medical Device Security and Anomaly Detection with AI: Risks, Models, and Future Perspectives in the IOMT	
Resul Kara	
CHAPTER 11	247
AI-Assisted Semantic Evaluation of EMC and Electrical Safety Standards: A Review and Framework Proposal	
Oğuz Gündüz & Gökhan Demirkıran	
CHAPTER 12	263
Assessment of Additive Impact on Fluidized Bed Combustion of Biofuels	
Zuhal Akyürek	
CHAPTER 13	273
Effects of Solar Panel Cooling Techniques on Energy Efficiency	
İbrahim Gunes	
CHAPTER 14	289
Evaluation of Environmental Recalcitrance of Biochar	
Zuhal Akyürek	
CHAPTER 15	299
Buldan Fabric Defects: Description, Analysis and Detection Using Image Processing Based Intelligent Models	
Erdem Dilmen & Emine Utkun Aktener & Selami Beyhan & Batuhan Bilgi	
CHAPTER 16	327
Environmental Isotope Hydrology and Recharge Dynamics of the Süleymanlı–İlica Spring System	
Yusuf Uras & Samet Ahmet Purdaş & Yağmur Uysal	
CHAPTER 17	357
Material Handling Concepts	
Yusuf Ziya Koç	

CHAPTER 18..... 369
Renewable Energy Applications in University Campuses: Reducing Carbon Emissions Based On Solar Energy Systems (SES) at Mersin University
Hüdaverdi Arslan & Sena Nur Demirel & Anıl Erdağ Nomer

CHAPTER 19..... 399
Applicability of Multicriteria Decision-Making Approaches in Underground Dam Planning
Ahmet Emir Köse & Salim Serkan Nas & Yusuf Kaya

CHAPTER 20..... 427
Position Control of a DC Motor via PID Controller Optimized by Musk Ox Algorithm
Said Müftü & Erdem Dilmen

CHAPTER 21..... 449
AI in Combating Digital Fraud in Banking: Cases, Signals, and Methods
Ali Keleş &. Aytürk Keleş

CHAPTER 22..... 459
Hydrogen Combustion Technologies: From Engine Applications to Gas Turbines and Safety-Oriented Modeling
Şafak Melih Şenocak &. Yasin Varol

CHAPTER 23..... 483
Photovoltaic Solar Panel Technologies: A Comparison of Efficiency, Cost, and Lifetime Based on Material Types
Şafak Melih Şenocak &. Yasin Varol

CHAPTER 24..... 503
Hybrid Laser Aluminum Welding of Characteristics and Industrial Applications
Ferit Artkın

CHAPTER 25..... 519
Synthesis, End-Of-Life, and Economics of Bioplastics
Gamze Özçakır

CHAPTER 26 531

Natural Coagulants Used in Water and Wastewater Treatment and Current Studies

Hasan Kıvanç Yeşiltaş & Belgin Bayat

CHAPTER 27 543

Biomimetic Designs From A Textile Engineering Perspective

Züleyha Değirmenci & Ebru Çoruh

CHAPTER 28 561

**Calcium-Based Borohydride Systems for Hydrogen Storage:
Mechanochemical Synthesis and Catalytic Dehydrogenation**

Ahmet Fırat Karabulut

CHAPTER 29 575

Particleboard and Composite Materials from Agricultural Wastes

Ahmet Fırat Karabulut

CHAPTER 30 585

**From Neural ODEs to Liquid Time-Constant Networks and Closed-form
Continuous-time Policies: A Control-Oriented Review**

Asaf Sayıl

CHAPTER 31 595

From Discrete Neural Network Graphs To Liquid Time-Constant Networks

Asaf Sayıl

CHAPTER 1

Impulse Response of Solid and Cut-Out Plates Supported by Corner Spring-Damper Systems

Can Gönenli¹ & Okan Duymazlar²

¹ Assoc. Prof. Dr., Ege University, Ege Vocational School, Department of Mechanical and Metal Technologies, 0000-0001-9163-1569

² Lec., Ege University, Ege Vocational School, Department of Electronics and Automation, 0000-0002-1327-7493

INTRODUCTION

The dynamic behavior of plate structures under transient loading has been a subject of continuous interest in vibration analysis and structural mechanics. Rectangular plates are frequently encountered in engineering applications such as aerospace structures, mechanical frames, vehicle bodies, and civil engineering components. Their vibration characteristics are strongly influenced by boundary conditions and possible geometric discontinuities. Even small changes in stiffness, mass distribution, or damping can result in noticeable differences in their dynamic response. In particular, plates supported by discrete spring-damper elements at the corners exhibit responses that differ significantly from those of continuously supported or fully clamped plates. Corner supports introduce localized compliance and energy dissipation, making the vibration characteristics dependent not only on the global geometry of the plate but also on the stiffness and damping properties of the supports. Such conditions are representative of practical engineering situations where plates are mounted at discrete points rather than along their entire edges.

The vibration behavior of rectangular plates with discrete corner supports has long been an important subject in structural dynamics. Early investigations by Cox (1960) demonstrated that point-supported plates at the corners exhibit unique natural frequency characteristics compared to fully clamped or simply supported configurations. Subsequent studies, such as Reed (1965), compared different analytical methods for calculating the frequencies of corner-supported plates, highlighting the sensitivity of results to support modeling. Later, Kim (1987) extended the theoretical framework by analyzing the flexural vibration of rectangular plates with point supports, providing valuable analytical solutions that clarified the influence of support placement on modal behavior. These works established the foundation for understanding how localized supports alter the dynamic response of rectangular plates.

In addition to support conditions, geometric modifications such as cut-outs have been widely studied due to their practical relevance in lightweight structures. Kalita and Halder (2016) examined rectangular plates with central cut-outs and demonstrated how openings significantly modify natural frequencies and mode shapes, showing a reduction in stiffness and mass redistribution effects. Li et al. (2020) employed the discrete singular convolution method to analyze the free vibration of thin rectangular plates with cut-outs, offering a refined numerical approach that captured the influence of cut-out size and position on modal parameters. More recently, Majhi (2024) investigated laminated composite plates with square cut-outs using ANSYS-based simulations, further confirming the strong effect of cut-outs on vibration characteristics. Together, these studies emphasize that both boundary conditions and geometric discontinuities play a

critical role in shaping the vibration characteristics of plate structures, and they provide a strong basis for further numerical investigations using finite element modeling techniques (Dai et al., 2020). Beyond geometry and boundary constraints, the material characteristics of the plate also play a decisive role in its dynamic response. Variations in elastic modulus, density, and damping properties can lead to considerable changes in natural frequencies and mode shapes (Aydin, 2022; Ayvaz, 2023).

In this study, the impulse response of rectangular plates supported at the four corners by identical spring-damper systems is investigated. Two different geometrical models are considered: a solid plate and a plate with a central cut-out. The presence of the cut-out alters the mass and stiffness distribution of the system, thereby influencing the transient response and the decay characteristics of oscillations. The comparison between these two configurations is essential for understanding how geometric modifications affect vibration behavior. The analysis is carried out using the finite element method within the ANSYS APDL environment. This approach enables an accurate representation of the plate geometries, the cut-out, and the corner spring-damper connections. Impulse loading is applied to one corner of the plate, which excites a broad range of vibration modes simultaneously. This makes it possible to observe mode participation, resonance effects, and the overall damping performance in the time domain. The transient responses are extracted for all four corners, allowing a direct comparison of the displacement histories across different support points. The obtained results are interpreted in two complementary ways: first, in terms of the dynamic behavior associated with each geometric configuration, and second, through a comparative evaluation of the solid and cut-out plates. This dual perspective provides a clear understanding of how the introduction of a cut-out changes the vibration response of corner-supported plates. The chapter thereby offers a systematic numerical study that emphasizes the combined effects of geometry, support conditions, and impulsive excitation on the dynamic performance of plate structures.

MATERIAL AND METHOD

In this study, the dynamic response of rectangular plates supported by spring-damper elements at their corners was investigated under impulsive loading using the finite element method in ANSYS APDL. Two configurations were considered: a solid rectangular plate and a rectangular plate with a central cut-out.

Both models were constructed with identical outer dimensions of $l_1=600$ mm (length) and $l_2=400$ mm (width), and a uniform thickness of $t=10$ mm. The material was assumed to be isotropic aluminum with an elastic modulus of $E=70$ GPa, a density of $\rho=2700$ kg/m³, and a Poisson's ratio of $\nu=0.33$. The solid plate

was defined as a single closed area, while the cut-out plate was generated by subtracting an inner rectangle from the main geometry. The inner rectangle was offset by $x_x=100$ mm from each side, resulting in a centrally located opening that reduces the effective stiffness and mass of the plate.

The structural domain was discretized using SHELL63 elements (ANSYS, 2015), which account for both membrane and bending effects. The corner supports were modeled using COMBIN14 spring-damper elements (ANSYS, 2015), each assigned a stiffness of $k=100$ N/m and a damping coefficient of $c=2$ N·s/m. Lower reference nodes were fully constrained in all degrees of freedom, while the upper corner nodes were restricted in U_x and U_y , leaving U_z free to allow vertical motion. This configuration ensured that the springs and dampers acted only in the vertical direction.

Transient dynamic analysis was conducted for both plate configurations. An impulsive vertical load of $F_z=1$ N was applied to Node 1 for a duration of 0.05 s. After the load was removed, the response was tracked up to 10 s to capture the free vibration and damping behavior. Vertical displacements of all four corners were extracted and compared. The results were then interpreted both on a geometry-specific basis (solid vs. cut-out) and in a comparative manner to highlight the effect of the central opening on dynamic response. The parameters employed in the study are summarized in Table 1.

Table 1. Parameters used in the solid and cut-out plate models

Parameter	Solid Plate	Cut-Out Plate
Software & Method	ANSYS APDL – FEM (Transient Analysis)	Same
Plate Element	SHELL63	Same
Support Element	COMBIN14 (spring-damper)	Same
Material	Aluminum, ($E=70$) GPa, ($\rho=2700$) kg/m ³ , ($\nu=0.33$)	Same
Outer Dimensions	(600 $\times$ 400) mm	Same
Thickness	10 mm	Same
Geometry	Solid rectangular plate	Central cut-out, inner rectangle offset 100 mm from each edge
Mesh Size	10 mm target element size	Same
Corner Support Parameters	($k=100$) N/m, ($c=2$) N·s/m (each corner)	Same
Boundary Conditions	Lower reference nodes fully fixed; upper corners U_x , U_y fixed; U_z free	Same

Parameter	Solid Plate	Cut-Out Plate
Loading	Impulse ($F_Z=1$) N at Node 1, duration 0.05 s	Same
Analysis Duration	10 s (with free vibration after 0.05 s)	Same
Output	Time histories of ($U_Z(t)$) at corner nodes	Same

RESULTS

The results obtained from the transient analyses are presented and discussed in this section. The solid and cut-out plate models used in the numerical study are shown in Figure 1 and Figure 2. Both geometries were analyzed under identical boundary and loading conditions to enable a direct comparison of their dynamic behaviors.

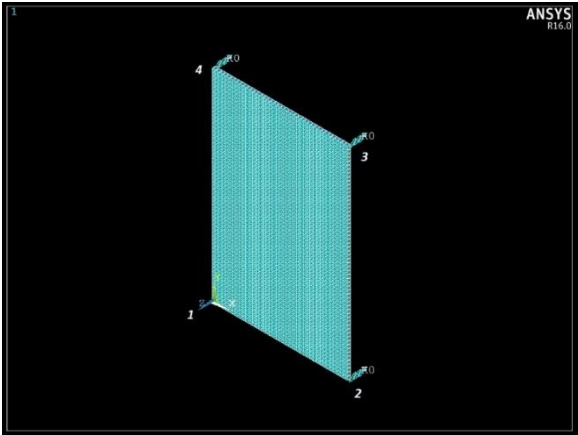


Figure 1. Solid plate model showing the orientation of corner points

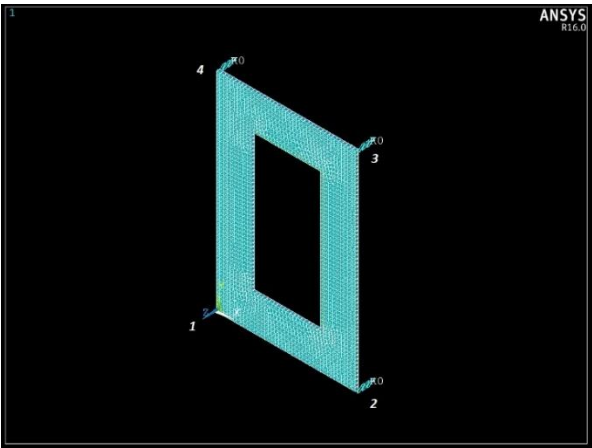


Figure 2. Cut-out plate model showing the orientation of corner points

The transient response of the solid plate under a 1 N impulse applied at corner point P1 is illustrated in Figure 3. The figure presents the time–amplitude curves for the four corner points (P1–P4), showing the vertical displacement variations during the free vibration phase.

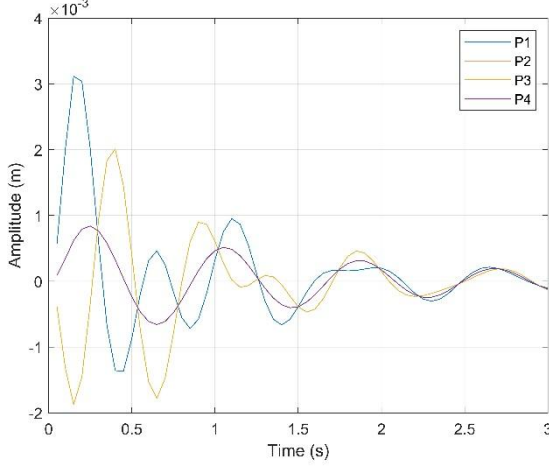


Figure 3. Transient response of corner points (P1–P4) for the solid plate

It can be observed that the point of impact (P1) exhibits the highest initial amplitude due to the direct application of the impulsive force. The remaining corners respond with smaller amplitudes and phase delays, indicating the propagation of vibration through the plate structure. As time progresses, the amplitudes of all points decay, demonstrating the damping effect introduced by the spring–damper supports. The oscillations gradually converge toward equilibrium, confirming the stability of the system and the effectiveness of the boundary damping characteristics.

The transient response of the cut-out plate under a 1 N impulse applied at corner point P1 is shown in Figure 4. The figure illustrates the time-amplitude variations of the four corner points (P1–P4), obtained from the transient dynamic analysis.

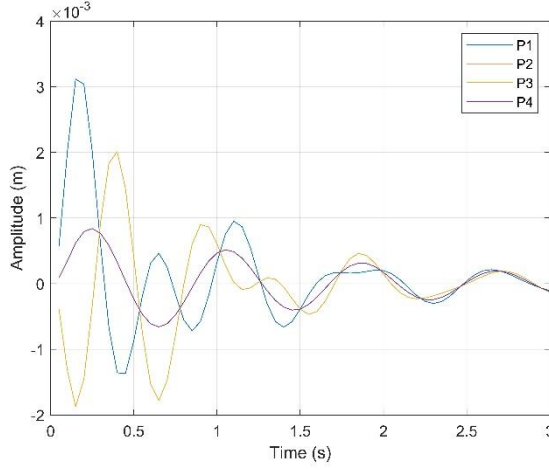


Figure 4. Transient response of corner points (P1–P4) for the cut-out plate

Compared with the solid plate, the cut-out model exhibits slightly altered vibration characteristics due to the reduced stiffness and redistributed mass around the central opening. The peak amplitude at P1 remains the highest because of the direct impulse application, but the subsequent oscillations decay faster than in the solid plate, indicating a change in the energy dissipation behavior. The presence of the central cut-out modifies the structural coupling between corners, resulting in minor phase differences and amplitude attenuation across the plate.

The comparison of the P1 corner responses for the solid and cut-out plates is presented in Figure 5. Both models were subjected to the same 1 N impulsive load applied at the P1 corner.

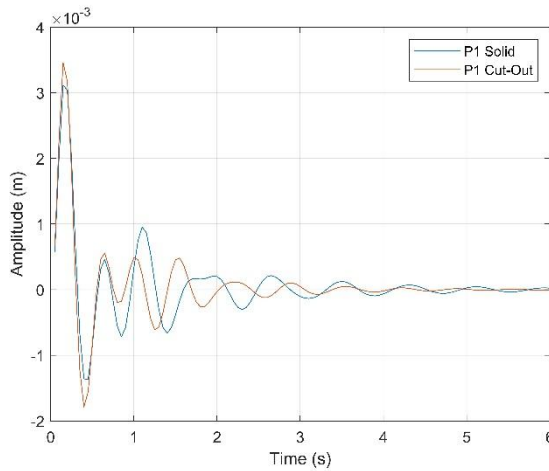


Figure 5. Comparison of transient responses at point P1

As seen in the figure, the initial peak amplitudes of the two models are nearly identical since the applied impulse and local stiffness at the loading point are comparable. However, subsequent oscillations differ in amplitude and decay rate. The cut-out model exhibits slightly faster attenuation of vibration amplitude, which can be attributed to its reduced stiffness and redistributed mass around the central opening. The solid plate, being structurally stiffer, sustains the oscillation for a longer period before reaching equilibrium. These results clearly indicate that the removal of material in the central region alters the global dynamic stiffness and damping characteristics of the plate system.

The comparison of the transient responses at P2 for the solid and cut-out plates is illustrated in Figure 6. Both models were subjected to a 1 N impulse applied at P1, and the resulting vibration propagation at the adjacent corner (P2) was recorded.

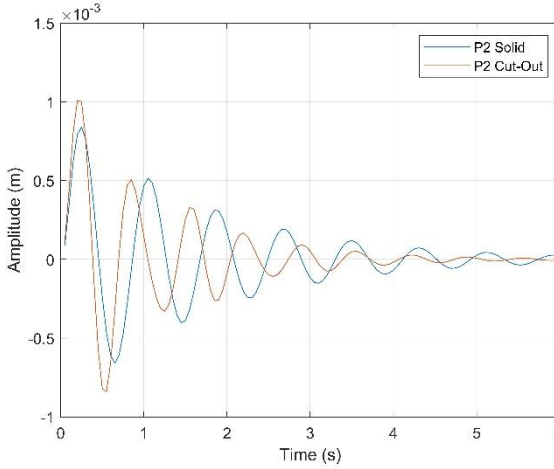


Figure 6. Comparison of transient responses at point P2

The results show that the P2 response amplitudes are significantly lower than those at the impact point, as expected due to the energy dissipation through the plate structure. The cut-out model demonstrates a slightly faster decay in amplitude, reflecting the reduction in stiffness and energy transfer efficiency caused by the central opening. The solid plate, having a continuous geometry, maintains oscillations for a longer period, indicating stronger coupling between the corners. The phase shift between the two models also suggests that the removal of material alters the global modal characteristics of the system.

The transient responses at P3 for both the solid and cut-out plate configurations are presented in Figure 7. These results represent the vibration propagation toward the opposite corner of the impact point (P1), where the transmitted energy and damping effects become more evident.

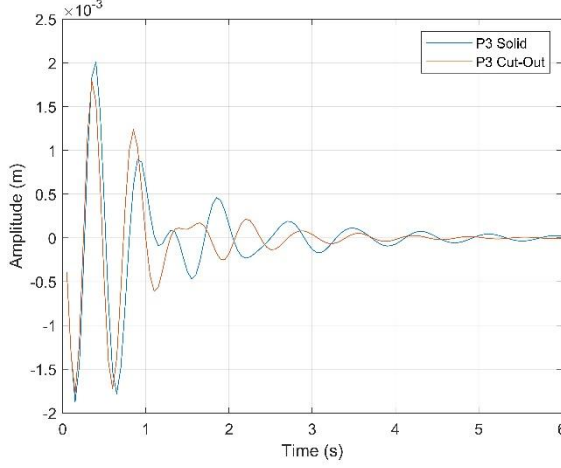


Figure 7. Comparison of transient responses at point P3

It can be seen that the initial amplitudes at P3 are lower than those observed at P1 and P2, indicating that only a portion of the input energy reaches the far corner through structural coupling. The solid plate exhibits slightly higher displacement amplitudes and longer oscillation duration compared to the cut-out model, which dissipates energy more rapidly due to its reduced stiffness and mass concentration around the central opening. The difference in phase alignment also reflects the influence of geometric discontinuity on the modal characteristics of the system.

The transient responses at P4, representing the diagonal corner opposite to the impact point (P1), are shown in Figure 8 for both the solid and cut-out plate configurations. This point provides valuable information on how the vibrational energy propagates through the plate and how geometric discontinuities influence the global response.

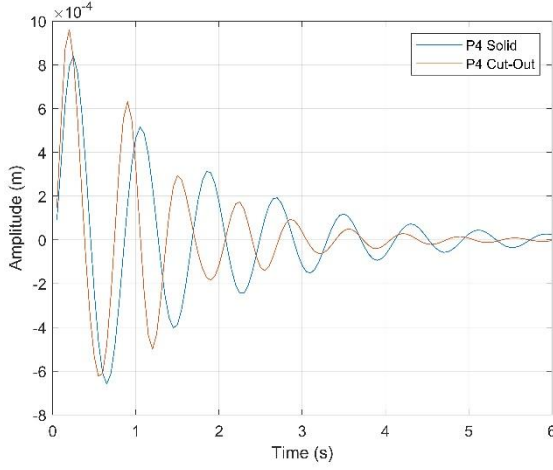


Figure 8. Comparison of transient responses at point P4

The results reveal that the displacement amplitudes at P4 are the lowest among all corner points, as most of the input energy is dissipated before reaching this location. The cut-out model displays slightly faster damping and smaller steady-state oscillations, confirming that the central opening alters the stiffness distribution and increases local flexibility, which affects energy transmission efficiency. In contrast, the solid plate retains a slightly stronger oscillatory response due to its continuous geometry and higher global stiffness. The close alignment of the two curves after approximately 3 seconds indicates that both systems eventually reach similar equilibrium behavior.

CONCLUSIONS

In this study, the dynamic responses of solid and cut-out rectangular plates supported by corner spring-damper systems were investigated under impulsive loading using finite element analysis in ANSYS APDL. The primary objective was to evaluate how geometric discontinuities, represented by a central cut-out, influence the vibration behavior and energy transmission characteristics of the plate structure.

The transient analysis results demonstrated that the impulsive excitation applied at one corner of the plate propagated throughout the structure, producing distinct displacement responses at each corner. The impact point exhibited the highest amplitude, while the opposite corners responded with smaller and delayed oscillations due to the damping and stiffness effects of the connecting springs. When comparing the two geometries, the cut-out model showed slightly faster amplitude decay and reduced vibration persistence, indicating that the removal of material modifies the global stiffness and mass distribution, leading to altered energy dissipation patterns.

Overall, the presence of a central cut-out was found to reduce the structural rigidity and affect the coupling between corner supports. Despite these differences, both configurations reached steady-state equilibrium within comparable time intervals, reflecting the stabilizing influence of the damping mechanism. The study highlights that even moderate geometric modifications can have a measurable impact on vibration propagation in plate structures supported by discrete spring-damper elements. These findings provide a foundation for future investigations focusing on optimizing geometric design and support configurations to achieve targeted dynamic performance in lightweight and flexible structural systems.

REFERENCES

- ANSYS. (2015). *ANSYS® Academic Research, Release 16.0*. Canonsburg, PA: ANSYS, Inc.
- Aydın, İ., Bahçepinar, A. İ., & Ayvaz, M. (2022). Hydroxyapatite coating processes with EPD method and investigation of mechanical properties of coatings. *International Advanced Researches and Engineering Journal*, 6(3), 197–204. <https://doi.org/10.35860/iarej.1102381>
- Ayvaz, S. İ., Aydın, İ., & Bahçepinar, A. İ. (2023). Effect of shot blasting on the boriding kinetics of AISI 316L stainless steel. *Protection of Metals and Physical Chemistry of Surfaces*, 59, 671–678. <https://doi.org/10.1134/S2070205123700624>
- Dai, L., Chen, Y., Wang, Y., & Lin, Y. (2020). Experimental and numerical analysis on vibration of plate with multiple cutouts based on primitive cell plate with double cutouts. *International Journal of Mechanical Sciences*, 183, 105758. <https://doi.org/10.1016/j.ijmecsci.2020.105758>
- Cox, H. L., & Boxer, J. (1960). Vibration of rectangular plates point-supported at the corners. *The Aeronautical Quarterly*, 11(1), 41–50. doi:10.1017/S0001925900001682
- Kalita, K., & Haldar, S. (2016). Free vibration analysis of rectangular plates with central cutout. *Cogent Engineering*, 3(1), 1163781. doi:10.1080/23311916.2016.1163781
- Kim, C. S., & Dickinson, S. M. (1987). The flexural vibration of rectangular plates with point supports. *Journal of Sound and Vibration*, 117(2), 249–261. doi:10.1016/0022-460X(87)90537-2
- Li, Y., Zhou, M., & Li, M. (2020). Analysis of the free vibration of thin rectangular plates with cut-outs using the discrete singular convolution method. *Thin-Walled Structures*, 147, 106529. doi:10.1016/j.tws.2019.106529
- Majhi, B., Sahu, P. K., Bhoi, U. N., & Mohanty, B. (2024). Free vibration of laminated composite plate with square cut-outs: A numerical and experimental analysis. *Romanian Journal of Acoustics and Vibration*, 21(2), 142–152. Retrieved from <https://rjav.sra.ro/index.php/rjav/article/view/438>
- Reed, R. E. (1965). *Comparison of methods in calculating frequencies of corner-supported rectangular plates* (NASA Technical Note D-3030). Washington, DC: National Aeronautics and Space Administration. Retrieved from <https://ntrs.nasa.gov/api/citations/19650024240/downloads/19650024240.pdf>

CHAPTER 2

Advances and Mechanization in Capture Fisheries

F. Ozan Düzbastılar¹

¹ Prof. Dr., Ege University, Faculty of Fisheries, Fishing and Processing Department,
ORCID: 0000-0002-5376-7198

1. INTRODUCTION

Technological developments since the mid-twentieth century have fundamentally changed the global fisheries sector, with catches increasing from more than 20 million tonnes in 1955 to almost 90 million tonnes in 2015 (Valdemarsen, 2001; FAO, 2025a). This increase was due to more efficient vessels, mechanisation, advances in synthetic fibres, materials technology, and modern systems for navigation, detection, and handling of fishing gears (Morgan, 2006; Thermes, Van Anrooy, Gudmundsson, & Davy, 2023; FAO, 2025b). Recent innovations have focused on sustainability, increasing the selectivity of capture technologies to reduce bycatch, improving operational safety, and optimising post-harvest operations (Valdemarsen, 2001; Ward & Hindmarsh, 2007; FAO, 2014; Wiyono, Raharjo & Permana, 2018; IMO, 2024b).

However, increased mechanisation has also led to problems such as overcapacity, overfishing, increased competition, and the marginalisation of small operators (Eigaard, Marchal & Gislason, 2014; Andersson & Ngazi, 1998). The study looks at the long-term effects of mechanisation and commercialisation on artisanal fisheries. The findings suggest that, in the absence of community development programmes, these processes may adversely affect food security and employment. The study also highlights the need to strike a balance between economic growth, poverty reduction, and environmental sustainability (Ramenzoni, 2017). The capture and handling of fish on board continue to depend on both the fish and the vessel's design and equipment, with highly automated systems for locating fish contrasting with a largely manual process, thereby necessitating integrated computerised systems (Choudhury & Bublitz, 1997). Technological progress, combined with market demands and regulatory requirements, is increasingly driving the adoption of automation to optimise productivity, efficiency, and sustainability in modern fisheries.

2. THE EVOLUTION AND DEVELOPMENT OF MECHANIZATION IN FISHERIES

Since prehistoric times, humans have harvested fish using simple tools such as spears, hooks, traps, and nets with minimal environmental impact. Technological progress and population growth gradually transformed subsistence fishing into an organised economic activity, and industrialisation in the late 1700s introduced more efficient methods (Pitcher, 2012). The use of steam-powered trawlers in the mid-1800s greatly increased catch efficiency, with steam screw trawlers harvesting several times more than sail-powered vessels, while innovations such as the triplex system further reduced operation times (Engelhard, 2005; Düzbastılar & Hoşsucu, 1997). Fishermen also expanded their crews and gear units to enhance productivity (von Brandt, 1975).

Mechanised and electronically equipped fishing vessels employ integrated navigation and acoustic systems such as radar, Long Range Navigation (LORAN), satellite navigation, gyro and magnetic compasses, autopilots, sonar, echo sounders, and course plotters to precisely locate and monitor fish schools (FAO, 2025b; Purse-seine Observer Guide, 2021). Course plotters combine chart data with GPS, LORAN, radar, and compass inputs, while advanced systems such as *SEANAV*, *Maptech*, *WindPlot*, and *MetMap* provide 3-D bathymetry, weather overlays, and infrared sea-surface data through standardized *NMEA* interfaces (FAO, 2025b).

Vessel Monitoring Systems (VMS) and Alarm and Monitoring Systems (AMS) use microcomputers to monitor fuel consumption, propulsion performance, and up to 150 sensor inputs, continuously comparing data with set thresholds to detect anomalies and support predictive maintenance (Bonnet, 1988; FAO, 2025b).

The capture and retention of fish is affected by the design of the fishing gear and by the behaviour of the fish, which depends on environmental and biological factors (Regenstein & Regenstein, 1991). Although automation has advanced, few systems fully automate capture and storage. Automated longlining, purse seining, and gillnetting rely on mechanical controls, while trawl systems employ computerised winch control and net monitoring to maintain optimal configuration relative to fish and the seabed (Choudhury & Bublit, 1997; Purse-seine Observer Guide, 2021; Trygg Mat Tracking and IMCS Network, 2023; Walsh et al., 2009). Advanced trawl control systems monitor warp tension, speed, and winch rpm, allowing automatic gear shooting, hauling, and tension compensation. Telemetry systems collect multi-parameter data - depth, mesh size, catch, inclination, water flow, salinity, temperature, and solar radiation - by means of ultrasound and sonar sensors, which display the net geometry and the environmental conditions in real time (FAO, 2025b; Tetley & Calcutt, 1991; Trygg Mat Tracking and IMCS Network, 2023).

Once on board, fish are stored in bulk or pre-processed according to species, vessel size, and market requirements, with individual catches ranging from 50-150 t on medium-sized vessels to over 150 t on larger ones (Olsen, 1985). Despite the presence of semi-automated systems on larger fishing vessels, handling, maintenance, and gear operations remain largely manual (Choudhury & Bublit, 1997).

3. ELECTRONIC EQUIPMENT OF FISHING VESSELS

3.1. Electronic Systems for Communication and Navigation

In larger fishing vessels, the bridge serves as the central control centre for navigation, communication, and fishing operations, integrating

displays, autopilot, and winch systems for vessel management, with sensors monitoring the performance of nets and cables (FAO, 2025b). Originally developed for underwater detection, Sound Navigation Ranging (SONAR) scans the water column vertically or horizontally and is widely used for navigation and fishing (Hermansson, 1978; FAO, 2025b; Purse-seine observer guide, 2021), whereas radar and specialized bird radars detect ships, land masses and flocks of seabirds, and Automatic Radar Plotting Aids (ARPA) systems calculate safe courses according to collision rules (Purse-seine guide, 2021; CIRCO, 2025). Navigation is based on gyrocompasses, magnetic compasses, and flux-gate sensors; communication systems use satellite networks, Very High Frequency (VHF), the Global Maritime Distress and Safety System (GMDSS), and the Vessel Monitoring System (VMS) for real-time monitoring and emergency communication, with the Emergency Position Indicating Radio Beacon (EPIRB) and the Global Positioning System (GPS) ensuring precise positioning (FAO, 2025b; IMO, 2024a; Purse-seine Observer Handbook, 2021; Trygg Mat Tracking and IMCS Network, 2023). Integration with computers has enabled Electronic Chart Display and Information Systems (ECDIS) to replace paper maps with interactive digital maps, improving navigation, plotting, and security (FAO, 2025b; SOLAS, 2000).

3.2. Fish-finding Equipment and Technologies

The echo-sounder, a basic hydroacoustic instrument, detects fish and determines the depth and composition of the sea floor, especially in shallow water or in poor visibility, by sending out vertical acoustic pulses and measuring the time between transmission and reception (FAO, 2025b; Hermansson, 1978). Modern multi-element sensors provide directional capability, helping to estimate fish abundance and to identify species with probability (FAO, 2025b). Integrated fishing and gear monitoring technologies enable vessels to locate target species, keep nets at optimum depth, avoid obstacles and improve the efficiency, quality and safety of catches, while also helping to recover lost trawls (Trygg Mat Tracking and IMCS Network, 2021). These systems are often combined with navigational and positioning instruments such as radars, compasses, ECDIS and GPS to improve operational efficiency and situational awareness.

3.3. Monitoring Fishing Gears and Other Equipment

The net sounder, an echo sounder equipped with a sensor mounted on the trawl headline and operated by a dedicated winch (Figure 1), provides the bridge with real-time data on net depth and water temperature during trawling operations

(FAO, 2025b). Additional sensors monitor the upper surface, fish passing through the mouth of the net, and the eyes of the stern trawler assess the concentration of fish above, below, and inside the net as well as the opening of the net from the bottom (Thermes, Van Anrooy, Gudmundsson, & Davy, 2023). Lost trawl or ground gears (footrope, bobbins, rollers, etc.) may be located using driftnets guided by sensor signals and last recorded coordinates (Walsh et al., 2009). In some fisheries, underwater cameras identify fish by size and species, with catch sensors on the codend signalling when the filling level is reached, allowing the catch quantity to be adjusted to the processing capacity. To reduce the bycatch of dolphins and other cetaceans, acoustic deterrent devices (pingers) are applied to the wings and head ropes of trawls (Trygg Mat Tracking and IMCS Network, 2023).

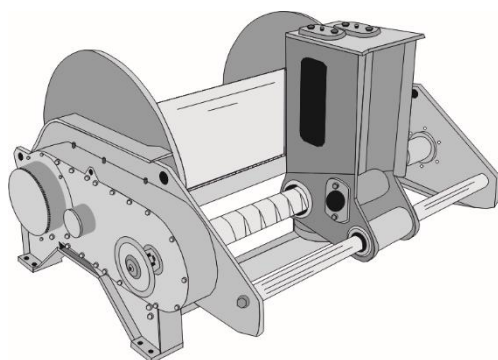


Figure 1. Net sounder winch used for trawling.

4. FISHING GEAR AND EQUIPMENT

4.1. Deck Machinery and Equipment

4.1.1. Purse seine mechanization - Purse seiners

In artisanal and semi-industrial fisheries, purse seiners are equipped with winches, capstans, purse line reels, brailers, power blocks, and occasionally net drums (Figure 2). Hydraulic derricks, introduced in the 1960s, further enhanced mechanisation and operational efficiency (ICCAT, 2008).

There are two main types of vessels. The American type features the bridge and accommodation forward, with the working deck aft, facilitating easy handling of the catch. The European type has the bridge and accommodation aft and the working deck amidships or forward, allowing large-scale net operations and onboard processing (Thermes, Van Anrooy, Gudmundsson, & Davy, 2023).

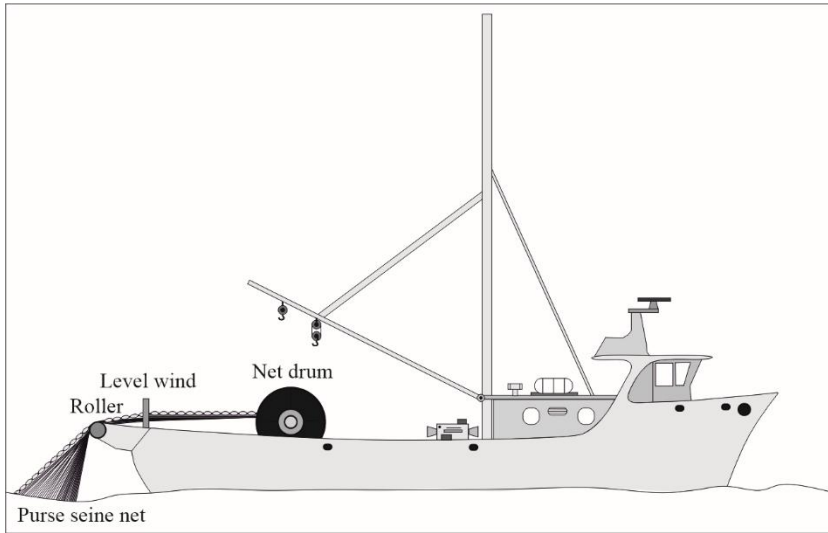


Figure 2. Drum seiner: Net drum, level wind, and roller on the fishing boat for drum seining (*Redrawn and modified from the FAO, 2025d*).

Industrial tuna purse seiners employ hydraulic power units, multi-stage winches, derricks, auxiliary winches, and support skiffs. Occasionally, helicopters are used to support operations. These vessels deploy nets ranging from 2,000 to 3,000 m in length and 250 to 300 m in depth, with a net weight between six and 20 t, depending on the design (FAO, 2025c; He, Chopin, Suuronen, Ferro, & Lansley, 2021). Such dimensions are typically used for large schools of tropical and temperate tunas, ensuring efficient encirclement and hauling. Net support includes bottom chains and rings, floats, and stacks on the stern deck, requiring powerful engines and high-capacity winches. Smaller net boats assist during encircling and transport phases. Fish Aggregating Devices (FADs) and artificial lights attract small pelagic species through phototaxis, forming dense schools. Strategic FAD use can enhance catch rates; however, environmental impacts and ecosystem effects should be carefully considered and managed (FAO, 2025c; Trygg Mat Tracking and IMCS Network, 2021).

Power block: The power block is a mechanical pulley used to haul nets, such as purse seines, onto the deck of the vessel (Figure 3).

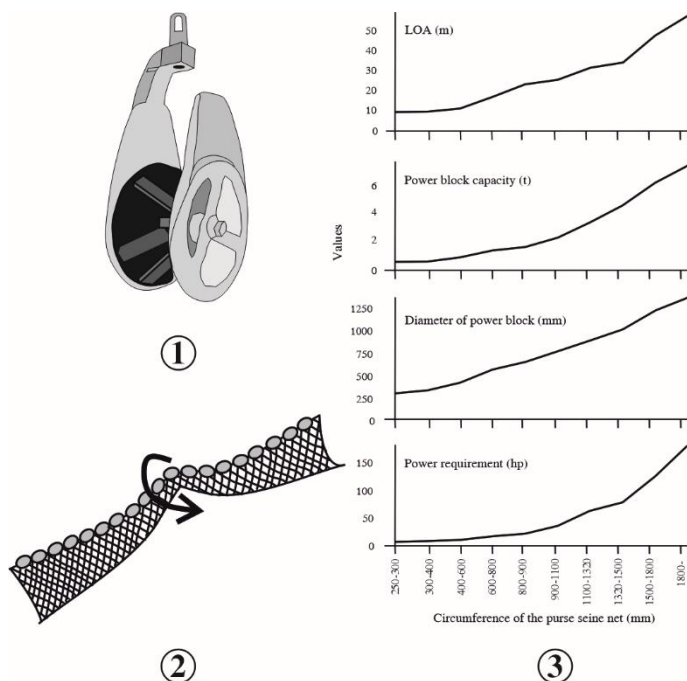


Figure 3. Selecting and calculating the appropriate power block capacity (1-Power block used in purse seine fishing, 2-Measurement of perimeter of the net, 3-Certain criteria are taken into account when choosing a power block (*Adopted and derived from Ben-Yami (1994)*)).

Recent innovations have improved efficiency and reduced labour requirements, and the co-generation of energy-intensive vessels is a key development in the purse-seine fishery (FAO, 2025b). During fishing, the net is lifted from the stern using the power block, with the operation varying depending on the fishing area and the net size (ICCAT, 2008; Ben-Yami, 1994).

Most vessels employ hydraulic power blocks suspended from the boom, allowing the seine to pass through and descend for manual loading. The correct positioning of the net is facilitated by the vertical and lateral movement of the main boom (Trygg Mat Tracking and IMCS Network, 2021). Purse seiners generally feature a mast with one or more booms: the largest carries the power block, while smaller side booms assist with gear handling and fish retrieval. When not in use, the aluminium power block with a hard rubber rim is stored on the stacked net, and its main wheel rotates to lift and lower the net during operations (Guide to the Purse-Seine Observer, 2021).

Power blocks are manufactured in different diameters, either open or closed, with the size determining the pulley capacity. For example, a 26-inch (~660 mm) pulley weighing 150 kg can handle 1.5 t and pull a net with a 60 mm mesh, suitable for vessels of 18-27 m in the purse seine segment. Typically, a power

block handles 20-50 % of the net weight in air, with smaller vessels operating at approximately 30 mmin⁻¹ and larger vessels at 80 mmin⁻¹ (Ben-Yami, 1994).

The selection of the appropriate power block depends on the net size and mesh configuration. Although exact calculations are challenging due to varying mesh sizes, approximate results can be obtained using standard formulas. For instance, a sardine or anchovy purse seine measuring 350 m in length and 72 m in depth (6,000 meshes, 210d 6, PA, 12 mm) may employ different strands for the main and anchor sections. Using denier-tex conversion, 210d (6) is equivalent to 1,260 g per 9,000 m, yielding an R_{tex} value of 155 for perimeter calculation (Ben-Yami, 1994).

The net perimeter, as illustrated in Figure 3, guides the selection of the pulley diameter. For a calculated circumference of 900-1,100 mm, the drum diameter should range from 700 to 840 mm, with a load capacity of 1.8-2.5 t. Power requirements and vessel *LOA* are set out in the same table. While theoretical calculations are useful, direct measurements - including the use of floats - can also determine the net perimeter using tape or rope.

$$C = 450 (6.10^{-5} R_{\text{tex}} + 0.02) \sqrt{N}$$

$$C = 450 (6.10^{-5} \cdot 155 + 0.02) \sqrt{6000} \quad (1)$$

$$C = 1021.3 \text{ mm}$$

Triple roller: On purse seiners, the hauling system often operates as a multi-power mechanism using a synchronized triple-roller (Triplex) winch, designed for high-performance net retrieval with improved grip, even traction, and reduced wear (Figure 4) (FAO, 2025b; Lin, Chen, Xu, & Wang, 2016). The winch combines mechanical and hydraulic components, including the frame, multiple drums, transmission system, baffles, and auxiliary parts, all installed on the deck.

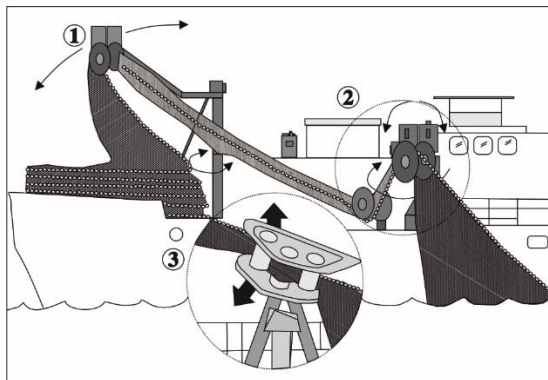


Figure 4. Hauling and arranging system in large-scale purse seiners (1- Net crane stacker, 2- Net winch and net drum, 3- Triple roller) (*Redrawn from Petrel International Commercial Fishing Deck Equipment*) (Petrel International, 2025).

Triplex winches, available in capacities ranging from 6 to 50 t, simultaneously grip the net with three rollers. The low-level mounting and the offset centre roller enhance stability and efficiency during hauling. Additional features, such as an independently declutchable aft roller, oil-bath lubrication, stainless-steel components, tilt-out functions, and compact deck installation, ensure safe and efficient operation while facilitating maintenance (Petrel International, 2025).

These design features improve operational reliability, reduce netting wear, and enable purse seiners to handle large or heavy nets efficiently. The combination of mechanical strength and hydraulic control allows for precise control of the tension, which contributes to the quality of the catch as well as to the safety of the crew.

Net (crane) stackers: A net stacker is a robust, adjustable system designed to handle purse seine nets efficiently. Its reinforced structure and hydraulic drive provide high torque and operational stability. Parallel arms maintain the roller head in a level position, ensuring precise net mounting. The optimised geometry allows a broad reach and lifting capability, while a vulcanised rubber roller head with inserts enhances traction, reduces slip, and minimises wear (Petrel International, 2025) (Figure 4). Different stacker types assist in orderly net handling: a head line stacker on a telescopic arm manages the upper portion of the net, while a cork line stacker on an articulated arm guides the lower sections, facilitating smooth, controlled hauling during purse seine operations.

Crane: Cranes serve as universal lifting devices for loading, unloading, and other operations on board (Figure 5). On purse seine vessels, cranes often replace the spreading mast for lifting nets, with power blocks mounted directly on the crane.



Figure 5. Cranes are installed on a purse seine vessel for operating power blocks, hoisting fishing gear, and performing auxiliary lifting tasks. Spreader and boom masts are not fitted on the decks of this type of vessel.

Cranes perform three primary movements - lift, roll, and slide - each powered by a dedicated engine. Some deck cranes, such as provision or hose cranes, are designed for specific tasks and not general cargo handling. Twin cranes may be mounted on a single platform, capable of 360° rotations, and can operate independently or in combination from the same tractor cab in two modes (WÄRTSILÄ, 2025). These design elements improve operational flexibility, reduce manual labour, and increase safety when handling heavy nets and equipment.

Net winch: The purse seine winches (Figure 6) are used to haul the line and bring the net to the vessel. They operate on a transmission principle, transferring driving force through friction between the drive drum and the net (FAO, 2025b; Lin, Chen, Xu, & Wang, 2016). Small vessels typically employ simple winches with lifting capacities of 2-6 t. Modern compact or multi-drum winches handle 5-45 t and often feature automatic level-wind systems, braking mechanisms, and innovations such as ring bolts to enable faster and safer net closure (Purse-seine Observer Guide, 2021; ICCAT, 2008).

Large purse-seine vessels use more complex hydraulic systems incorporating net stackers, winches, drums, and triple rollers. These systems offer variable speed, adjustable articulation, and high pull capacity (18.1-45.8 t). Stainless steel or rubber contact surfaces, along with remote control options, ensure accurate and efficient net handling while supporting the drying and management of the catch (Szczepanek, 2017; Lin, Chen, Xu, & Wang, 2016; Petrel International, 2025).

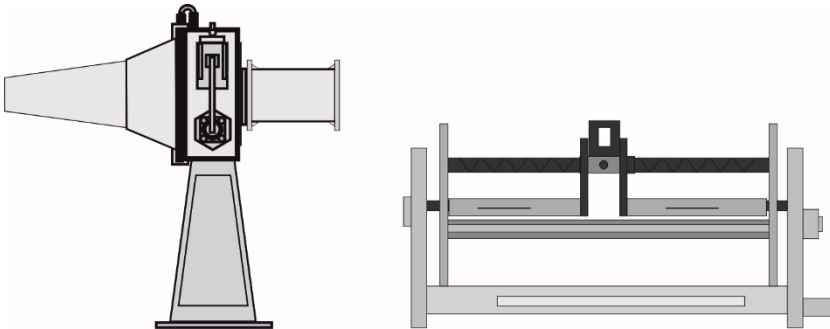


Figure 6. Purse winch and purse line winches are equipped with an automatic level-wind system.

Brailer (Grap): The brailer is used to transfer the catch from the deep-sea seine to the vessel once it reaches the deck. It can be operated manually or through a combination of manual and mechanical power (FAO, 2025b; Purse-seine Observer Guide, 2021).

Rapid loading is necessary, especially in high sea temperatures, to maintain the quality of the catch. Brailers with high capacity are usually kept on board and

can carry around 10 t of fish per haul from the chafer to the deck (ICCAT, 2008; Trygg Mat Tracking and IMCS Network, 2021).

Fish pump: The fish pump (Figure 7) transfers the catch from the nets to the vessel or dock, replacing the labour-intensive brailers used in medium-sized and large purse seiners. Transfers often occur directly from the purse seine of the catching vessel to another vessel (Digre, Tveit, Solvang-Garten, Eilertsen, & Aursland, 2016; Düzbastılar et al., 2023; Trygg Mat Tracking and IMCS Network, 2021). Historically, manual handling of fish was slow and labour-intensive, often taking more than 12 hours on busy days. Today, centrifugal or reciprocating pumps can empty the tank within minutes (Carlson, 1975).

These pumps, centrifugal and mounted either on the engine or on the surface of the water, use a radial piston hydraulic engine that operates at high speeds to cope with thermal shocks and frosty temperatures (FAO, 2025b; Petrel International, 2025). Their marine-grade aluminium construction allows safe passage of fish, while galvanised cages and rubberised edges prevent entanglement with nets. Fish pumps are used in purse seine and trawl fisheries, as well as in aquaculture, for counting, sorting (separation based on species or general characteristics), and transporting fish (FAO, 2025b).

Capacities range from 500 to 4,000 th^{-1} , with discharge heights of 0-14 m and power requirements of 65-85 kW. Optimal operation is typically at 5-6 m discharge height, yielding around 2,000 th^{-1} , depending on species, fish size, net density, and fan speed (Düzbastılar & Hoşsucu, 1997).

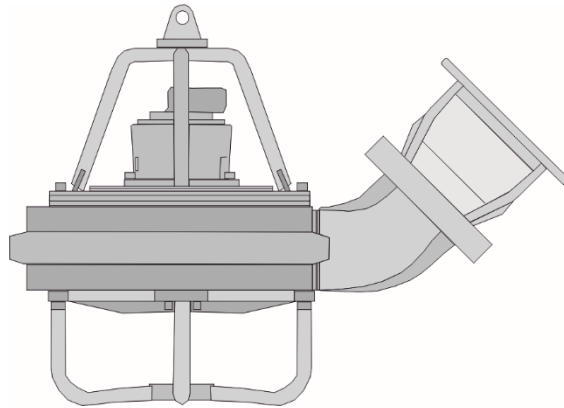


Figure 7. A fish pump used in purse seine fishing

Capstan: A capstan is a winch with a hydraulically or electrically driven horizontal or vertical warp head that can pull a rope from any horizontal direction without guide bars, as long as the rope is aligned with the shaft (Figure 8) (Petrel International, 2025; FAO, 2025b). Warping heads - cylindrical devices mounted on the terminal end of winch or windlass shafts - are made of cast iron for soft

ropes or steel for cables, and can be customised to meet specific operational requirements (WÄRTSILÄ, 2025; Petrel International, 2025).

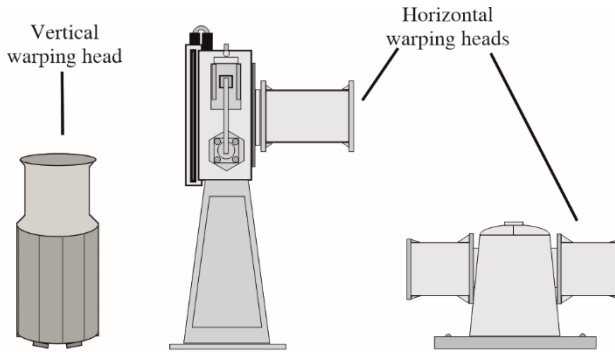


Figure 8. Capstans with vertical and horizontal warping heads.

Capstans operate at slow speeds with high torque, handling hawsers, anchor lines, or lifting fish via booms. High-torque motors enable hauling speeds of 18-90 mmin^{-1} and pulling capacities of 1.5-3.5 t. Operation is controlled through hydraulic valves for two-way motion, making capstans suitable for heavy-duty purse seine and trawl operations (Beverly, Chapman, & Sokimi, 2003).

Purse davit and davit roller: On purse seine vessels, the purse davit is a specialised crane or lifting device installed on deck to handle the purse seine net and associated tools during lifting, shooting, and positioning. A davit roller, mounted on the vessel's rail, assists in closing the net on the port side to form a chafer (bag), facilitating orderly and efficient net handling (ICCAT, 2008).

Excluder grate: In purse seine fishing, an excluder grate is installed at the end of the net system, typically at the pump outlet, to prevent bycatch from entering the collection system. It allows target species such as sardines, mackerel, and tuna to pass, while excluding larger or non-target species like sharks, rays, and juvenile fish. This device protects the fishing gear, reduces the mortality of non-target species, and ensures compliance with bycatch regulations (Naisbett-Jones, Mudrak, & Raborn, 2025; Raborn, Mudrak, & Naisbett-Jones, 2025).

Sieve system: The sieve system returns undersized live fish after sorting and separates fish according to size and species (Figure 9).

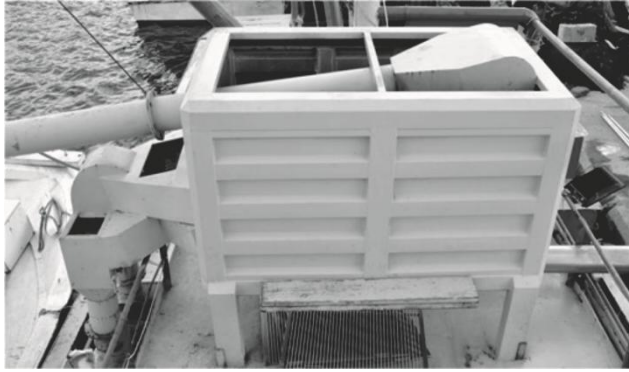


Figure 9. A sieve system is employed to separate and grade fish after they are pumped through the fish transfer line, with undersized individuals passing through the grids and being released back into the sea.

Although local fishermen often target small fish to increase income, the fate of discarded fish remains uncertain. The vibrating sieve system, adapted from a tea sieve with two or three grids, has been used on Aegean purse-seine vessels for the past ten years, provided there is sufficient deck space. This system improves the survival of multi-species fisheries, as the absence of such sorting contributes to the high mortality of overexploited stocks through indiscriminate discarding (Düzbastılar et al., 2023).

4.1.2. Mechanisation of seine fishing

Seine nets, including beach and boat seines, may be either cone-shaped with long wings and a codend or single-piece nets without a codend, capturing fish by encircling and herding (He, Chopin, Suuronen, Ferro, & Lansley, 2021). Seiners differ regionally: bridges are usually located aft with working decks amidships, though newer vessels often feature forward bridges and aft decks. These vessels commonly target bottom species such as cod, haddock, whiting, and flatfish.

Deployment techniques vary, including anchored “anchor seining” in Denmark and free-floating “fly dragging” in Scotland. Older vessels rely on manual rope hauling and coiling, whereas modern seiners use warps stored on reels and compact, high-powered winches with aft-mounted power blocks. Onboard processing includes sorting, weighing, grading (classification based on size or quality), gutting, freezing, and storage at 0°C. Larger industrial vessels may also be equipped with filleting and packing lines, as well as onboard refrigeration systems (Thermes, Van Anrooy, Gudmundsson, & Davy, 2023).

4.1.3. Mechanisation of dredging

A dredge is a cage-like device equipped with a scraper blade or teeth at the bottom, and towed along the seabed to collect animals in a cage or bag (He, Chopin, Suuronen, Ferro, & Lansley, 2021). Dredges primarily target molluscs

such as scallops, mussels, oysters, and clams. The gear is towed using up to four outriggers, and large dredgers may deploy multiple dredges per side. Some vessels employ heavy mechanical units mounted on the bow. Deck layout varies, with bridges and accommodations positioned either aft or forward, depending on the vessel's design. Hydraulic dredges require powerful pumps and can be integrated with winches to enhance operational efficiency (Thermes, Van Anrooy, Gudmundsson, & Davy, 2023).

4.1.4. Mechanisation of lift netting

A lift net is a framed net lowered and raised to capture fish congregating above it, typically deployed from the vessel's side using beams or outriggers supported by masts (He, Chopin, Suuronen, Ferro, & Lansley, 2021). Vessel sizes range from small open boats of about 10 m to larger semi-industrial vessels designed for short trips targeting fish for human consumption. Most baitfish are caught using pole-and-line or long-line techniques. Larger lift netters employ hydraulic derricks, winches, and powerful lights to handle lines, outriggers, and booms efficiently. In contrast, smaller vessels may operate non-mechanically or use hand-operated winches (Thermes, Van Anrooy, Gudmundsson, & Davy, 2023).

4.1.5. Trawl mechanization

A trawl is a cone-shaped net, usually with a codend, towed by one or two vessels to herd and capture fish. Trawls likely evolved from coastal seines (He, Chopin, Suuronen, Ferro, & Lansley, 2021; Sahrhage & Lundbeck, 2012). They are classified as bottom, midwater, or semipelagic, and their efficiency has historically been enhanced by trawl doors (Gabriel, Lange, Dahm & Wendt, 2005).

Modern trawlers, including side, stern, and outrigger beam types, are equipped with high-capacity hydraulic, electric, or mechanical systems such as winches, cranes, drums, refrigeration/freezing devices, and electronic equipment. Pelagic trawlers may also use fish pumps to clear cod (Thermes, Van Anrooy, Gudmundsson, & Davy, 2023; Hammer, 1974). Vessel sizes range from small nearshore boats with ice or seawater storage to large factory and freezer trawlers capable of extended voyages. Large vessels include onboard processing, packaging, and freezing facilities. Deck arrangements vary according to fishing techniques, with outriggers, beams, and working decks optimally positioned for efficient handling and towing (Thermes, Van Anrooy, Gudmundsson, & Davy, 2023).

Trawl doors: Otter boards (trawl doors) (Figure 10) are used in most trawl types, except pair trawl and beam trawl (ICES, 2019). They are essential for efficient fishing as they create horizontal net openings using hydrodynamic resistance, which significantly affects fishing efficiency and effort (Aydın &

Düzbastılar, 2011). The size, weight, shape, aspect ratio, attack angle, and other design features depend on the target species, tow speed, sea conditions, and vessel dimensions. Vessel design and size are often influenced by the specifications of the trawl doors themselves (ICES, 2019).

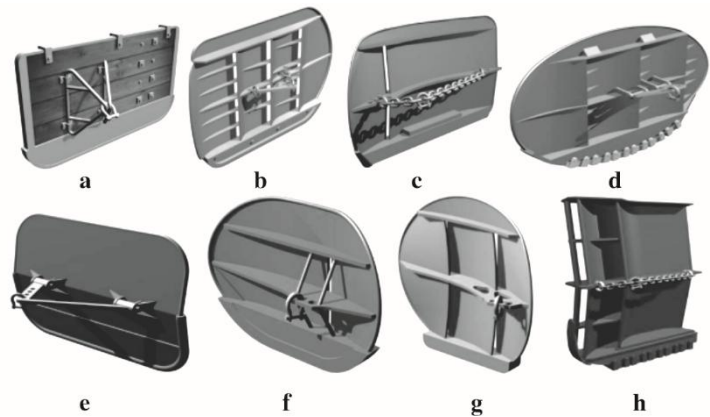


Figure 10. Types of trawl doors: a-Flat door (wood/steel); b-Flat cambered door; c-Cambered V-door; d-Foiled curved door; e-V-shape door; f-Oval cambered door; g-Round foiled door; h-Multi-foiled door (*Redrawn and modified from the ICES Scientific Reports*) (ICES, 2019).

Davits: Davits are lifting devices on fishing vessels designed to ensure the secure operation and handling of equipment and auxiliary gear. In trawl fishing, they are primarily employed to launch, retrieve, or support trawl doors, thereby reducing manual labour and enhancing operational safety. Davits can be fixed, swinging, or hydraulically/electrically powered, with handling capacity generally increasing from fixed to powered types (FAO, 2025b).

Net drum: Mechanical nets are large, hydraulically driven spools used to haul trawls or surrounding nets (Figure 11). Nets can be stored either on board or on a drum/roller mounted above deck. Rollers are especially useful in mid-water trawls, which often use wider and longer nets. Warps are typically wound onto large drums on either side of the vessel to facilitate efficient deployment and retrieval (Thermes, Van Anrooy, Gudmundsson, & Davy, 2023).

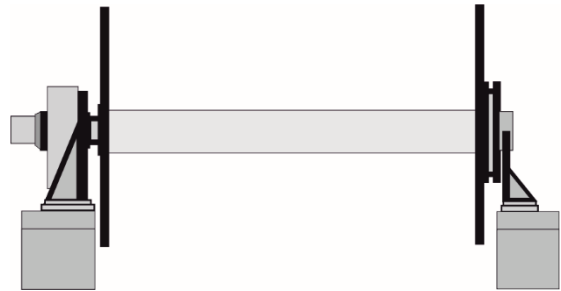


Figure 11. Net drum in trawl fishing.

Trawl winch: The efficient operation of the trawl winch requires careful design of all components, taking into account factors such as type and amount of load applied, level of maintenance during operation, and type of propulsion system used (Velu, 1972). Winch design parameters - including drum capacity, traction power, and hauling speed - are directly influenced by the fishing area, target species, and depth, which in turn dictate the type of vessel and gear required (Carral, Carral, Lamas, & Rodríguez, 2015).

Specialised trawl winches often work in conjunction with net drums. Some are automatic, deploying the warp according to predefined lengths and timings, while the first 50 m is applied manually before the system continuously adjusts tension. Recent control methods have improved cable tension stability and reliability, as simulations show reduced stress variations under various sea conditions (Liu, Wang, & Xu, 2025).

The transfer to winches depends on the size of the vessel: small vessels usually use auxiliary engines with coupling or chain drives, while larger vessels rely on main engine-belt or chain drives. The main engine power, winch capacity, drum size, warp radius, and pulling speed are all variable depending on operational requirements and vessel specifications (Table 1). For example, a hydraulically controlled single trawl winch designed to handle 1,680 m of 14 mm warp requires a winch power of approximately 65 hp, taking into account transmission losses (Nomura, 1985).

Table 1. Power requirement and winch specifications (Prado, 1990)

BHP (hp)	Winch power (hp)	Warp length (m)	Warp diameter (mm)	Hauling speed (ms ⁻¹)	*Towing capacity (kg _f)
100	25	700	10.5	1.0	900
200	40	1000	12.0	1.2	1600
300	60	1250	13.5	1.35	2500
400	80	1350	15.0	1.4	3500
500	120	2100	16.5	1.5	4500
700-800	165	2000	19.5	1.5	6500

* Drum half-full

The drum is empty

Hauling speed (v): 44 mmin⁻¹ (~0,73 ms⁻¹)

Towing capacity (F): 6600 kg_f

The drum is half full

Hauling speed (v): 99 mmin⁻¹ (~1,65 ms⁻¹)

Towing capacity (F): 2930 kg_f

The drum is full

Hauling speed (v): 130 mmin⁻¹ (~2,16 ms⁻¹)

Towing capacity (F): 2120 kg_r

$$P_{\text{winch}} = \frac{F \cdot v}{k \cdot \eta}$$

$$P_{\text{winch}} = \frac{6600 \cdot (0.73)}{75.1} \quad (2)$$

$$P_{\text{winch}} = 64.24 \text{ hp}$$

Engine power terminology is critical for understanding winch operation. Indicator horsepower (*IHP*) measures theoretical engine output, whereas brake horsepower (*BHP*) accounts for friction losses in bearings and connecting rods. Shaft horsepower (*SHP*) subtracts transmission and auxiliary losses from *BHP*, and effective horsepower (*EHP*) represents usable power after further frictional adjustments, typically around 23% of *SHP*. In practice, the pull force of the trawl winch is taken as equal to or up to twice the *EHP*. For instance, with an *EHP* of 65 hp, the corresponding *SHP* and *BHP* are 283 hp and 301 hp, respectively. Using these relationships, a winch requiring 60 hp corresponds approximately to 300 hp *BHP*, and after adjustment, the winch power (P_{winch}) falls between 60 and 75 hp, consistent with other calculations (Nomura, 1985).

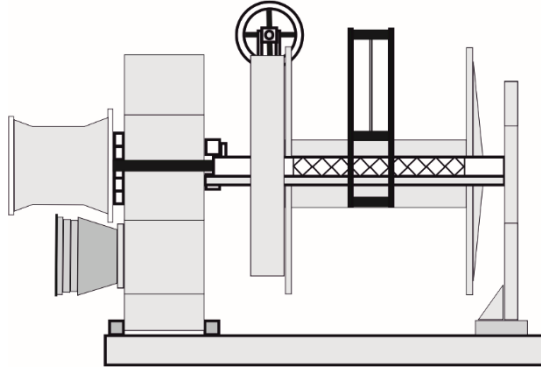


Figure 12. Single trawl winch with level-wind system.

Roller: Rollers are cylindrical or grooved devices installed on trawler decks to guide and support nets and ropes during fishing operations. They reduce friction, prevent damage to the gear, and enhance operational efficiency during both lifting and deployment. Proper selection and positioning of rollers are crucial to ensure smooth net handling and to minimize wear on ropes and nets.

4.1.6. Mechanisation of gillnetting/trammel net fishing

The net hauler (Figure 13), operated manually or hydraulically, is commonly used in small-scale fisheries to improve efficiency and reduce labour. Gillnets and entangling nets consist of long mesh panels that capture fish by gilling,

snagging, or entangling (He et al., 2021). These nets are used by small coastal vessels and also by large industrial driftnets operating in the high seas. Combined gear types, such as trammel nets and gillnets, are also widely used. Most gillnetters utilize hydraulic or mechanical haulers and net drums to speed up hauling, increase safety, and reduce physical strain on the crew (Thermes et al., 2023).

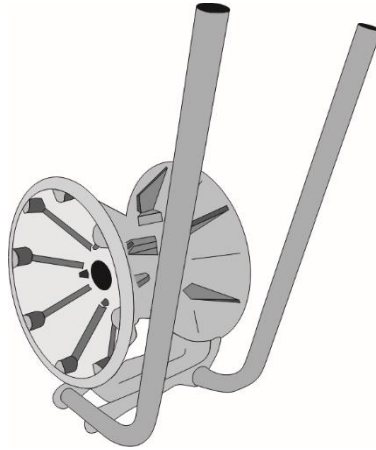


Figure 13. Net hauler, commonly used in small-scale fisheries. Drum support includes guiding arms to facilitate proper net handling.

4.1.7. Longline mechanization

A longline is a hook-and-line gear consisting of a mainline and branch lines with hooks at regular intervals. Partial mechanisation began in the late 1970s, with fully automated systems introduced in the early 1990s (He, Chopin, Suuronen, Ferro, & Lansley, 2021; Nygaard, 1988; Thermes, Van Anrooy, Gudmundsson, & Davy, 2023). Longliners range from small artisanal boats deploying 300-400 hooks per day to large freezer vessels using 2,500-3,500 hooks per day on midwater or drifting lines up to 80 km. Mechanised longlines incorporate hook haulers, line coilers, baiting equipment, and storage units (Beverly, Chapman, & Sokimi, 2003). These systems increase efficiency and safety by reducing manual labour, controlling line tension, and preventing tangles. However, in small-scale fisheries, their use is limited due to high costs, reliance on electricity, mechanical failures, and safety risks. Furthermore, improved efficiency may increase fishing pressure if not managed sustainably (Nygaard, 1988).

4.1.8. Line fishing mechanisation

Handlines and pole-and-lines are hook-and-line gears operated directly by fishers (He, Chopin, Suuronen, Ferro, & Lansley, 2021). Line vessels use hooks with or without bait or lures. The minimum equipment required consists of a

landing platform, anchor buoy, a live bait tank with a water spray system and a central holding area for the fish. Trollers target pelagic species such as tuna, dolphinfish, mackerel, and billfish using multiple baited lines or lures. While mechanisation is not essential, purpose-built trollers often have 2-4 outriggers with lines hauled using hydraulic or electric reels to improve efficiency and reduce physical effort (Thermes, Van Anrooy, Gudmundsson, & Davy, 2023).

4.1.9. Mechanisation of trap fishing

Traps are stationary structures used to guide or attract fish, either by utilizing currents or by bait (He, Chopin, Suuronen, Ferro, & Lansley, 2021). Trap setters deploy a variety of gear, including pots, pound nets, fyke nets, stow nets, and other stationary barriers. Larger vessels are equipped with derricks, cranes, or other lifting devices to haul pots efficiently. Smaller vessels often rely on hydraulic or mechanical pot haulers to reduce manual labour and improve operational efficiency (Thermes, Van Anrooy, Gudmundsson, & Davy, 2023).

4.2. Fishing Gear Modifications

Bycatch Reduction Devices (BRDs), commonly installed in shrimp trawl codends, are designed to exclude unwanted species, juveniles, and endangered animals. Turtle Excluder Devices (TEDs) specifically allow turtles to escape (Figure 14) (FAO, 2025b). Additional mitigation tools include Sea Lion or Seal Exclusion Devices (SLEDs), Seabird Scaring Devices (SSDs), Dolphin Deterrent Devices (DDD), Fisheye escape openings, NAFTAED and Super Shooter systems, and square mesh or Radial Escape Sections (RES). These modifications allow non-target species to leave while maintaining target species, improving selectivity and reducing bycatch (FAO, 2025b; Metin et al., 2005; Trygg Mat Tracking and IMCS Network, 2023).

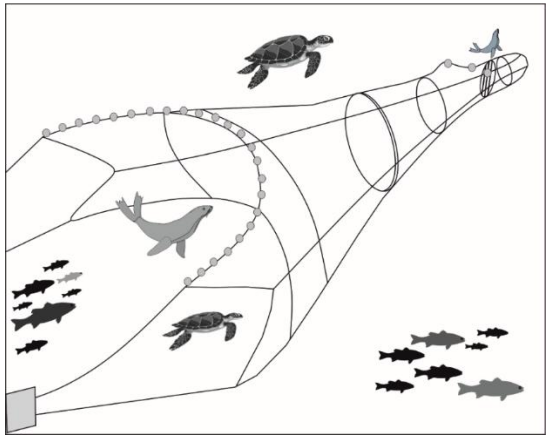


Figure 14. Exclusion devices for mammals and sea turtles (Redrawn from MCS Practitioners Guide to: Trawling).

Trawl doors are an essential part of trawl nets and may be made of steel, aluminium, wood or plastic. Their characteristics - material, construction, size, configuration, angle and orientation - have a significant impact on the efficiency of the catch (Aydın & Düzbastılar, 2011).

The relationship between brake horsepower (*BHP*) and trawl door surface area (*A*) can be determined using established formulas (Nomura, 1985). For instance, Formula (3) calculates the maximum trawl door area ($A_{\max}$) and weight ($W_{\max}$) based on the main engine's *BHP*. If the *BHP* is 300 hp, the maximum area and weight of a single trawl door are 2.58 m² and 458 kg, respectively.

Formula (4) can alternatively be used to estimate the trawl door area using the vessel's main engine power (*HP*). Considering that *BHP* typically corresponds to 80-90% of the nominal *HP*, a vessel with an *HP* of 375 corresponds to a trawl door area of 2.82 m². This value is very close to the result obtained from Formula (3). Table 2 provides a summary of trawl door dimensions, weight, and the corresponding vessel power requirements, allowing appropriate selection based on engine specifications.

$$\begin{aligned} A_{\max} &= (0.0945) BHP^{0.58} & W &= (0.00478) BHP^{0.8} \\ A_{\max} &= (0.0945) 300^{0.58} & W &= (0.00478) 300^{0.8} \\ A_{\max} &= 2.58 \text{ m}^2 & W &= 0.458 \text{ t} = 458 \text{ kg} \end{aligned} \quad (3)$$

$$\begin{aligned} A &= (3.47) \log HP - (6.11) \\ A &= (3.47) \log 375 - (6.11) \\ A &= 2.82 \text{ m}^2 \end{aligned} \quad (4)$$

Table 2. Selecting the appropriate power for a rectangular flat door (Prado, 1990).

<i>BHP</i> (hp)	Length (m)	Width (m)	Area (m ²)	Maximum weight (kg)
50-75	1.3	0.65	0.85	45
100	1.5	0.75	1.12	100-120
200	2.0	1.00	2.00	190-220
300	2.2	1.10	2.42	300-320
400	2.4	1.20	2.88	400-420
500	2.5	1.25	3.12	500-520
600	2.6	1.30	3.38	600-620
700-800	2.8	1.40	3.92	800-900

Galvanic Time Release (GTR) devices prevent ghost fishing caused by lost or abandoned pots. They consist of plated wire eyes (cathodes) inserted into precisely machined cylinders made of a highly reactive metallic alloy (anode), which corrode over time to release the trap (FAO, 2025b).

Medina panels are fine mesh sections attached to the top of the net in purse seine fisheries to exclude dolphins, reduce bycatch and improve conservation.

5. POST-HARVEST TECHNOLOGIES AND EQUIPMENT

The first fish processing vessels, such as the wooden *Santa Flavia* and *Mazama* in the 1920s and the steel *International* in 1929, laid the foundation for offshore fish processing and canning (Nickum, 1975). Modern factory vessels are larger and highly automated, equipped with advanced onboard processing and freezing systems. These allow for integrated cold chains for fresh, chilled, and frozen products (Thermes, Van Anrooy, Gudmundsson, & Davy, 2023). Automation has significantly increased processing speeds and capacities. In the 1990s, pelagic processing plants handled around 150 kg of fish per worker per day; today, this can exceed 1,500 kg, and in some cases reach 15,000 kg per worker per day (Einarsdóttir, Guðmundsson, & Ómarsson, 2022).

The fish processing sector is slow to adopt technological innovation due to the diversity of raw materials, regional conditions and reliance on manual labour. The introduction of canning, freezing, and surimi production has gradually led to mechanisation, improving product quality and operational efficiency. However, technological progress lags behind other food sectors, with automation largely confined to filleting, canning, and surimi processing. Challenges such as irregular fish geometry, variable catches, and the lack of automated detection systems continue to limit full automation (Choudhury & Bublitz, 1997).

Onboard handling - including transfer, sorting, gutting, washing, chilling, and storage - is crucial for maintaining product quality. Key factors include product purity, temperature control, water quality, handling methods, and staff hygiene (Mishra, Kulkarni, & Ghosh, 2021). Studies on whitefish indicate that freezing extends market availability beyond the fishing season (Erikson, Uglem, & Greiff, 2021). Immediate pre-cooling with crushed ice or a seawater-ice mixture is essential, especially in warm climates. Imaging and modelling techniques further support optimisation of handling and classification (Thanasarnsakorn, Taradol, Thimkrup, & Siriraksophon, 2005).

Although traditional on-board processing methods continue to be common, machine-vision technologies are increasingly used for the automation of sorting, grading and processing (Gümüş, Balaban, & Unlüsay, 2011). These systems enhance efficiency and support advancements in fisheries, aquaculture, and processing operations (Mathiassen, Misimi, Østvik, & Aursand, 2012).

Sorting and grading systems: After being caught, the fish are transported to processing plants, either by vessel or by land. In wild fisheries, a first classification by species is often required and the fish can also be classified according to size, weight and quality (Einarsdóttir, Guðmundsson, & Ómarsson, 2022). Sorting remains labour-intensive, but automation and artificial intelligence (AI) can substantially reduce costs while improving efficiency and

product quality. Despite progress, fully effective automated sorting systems at a commercial scale are not yet widely implemented (Ibrahim & Sultana, 2006).

Recent research shows that computer vision and robotics are mature enough to perform classification, inspection and treatment tasks in the pelagic sector (Einarsdóttir, Guðmundsson, & Ómarsson, 2022; Mathiassen et al., 2006). A prototype onboard sorting system using computer vision and pneumatic controls achieved high accuracy under realistic marine conditions (Strachan, 1994). The industry has increasingly adopted AI and computer vision technologies to develop new automation applications (Einarsdóttir, Guðmundsson, & Ómarsson, 2022). For example, a YOLO-based deep learning model has recently been developed for onboard, real-time detection, tracking, and classification of Gazami crab (*Portunus trituberculatus*), integrating attentional and feature-fusion mechanisms to improve accuracy and robustness (Li et al., 2025).

Traditionally, the grading of fish was done manually, on the basis of visual inspection, experience or the use of measuring instruments. Mechanical systems such as roller graders were also used to separate sizes. Modern technological advances - including optical sensors and computer vision - have improved both the accuracy and the consistency of grading in comparison with manual methods (Einarsdóttir, Guðmundsson, & Ómarsson, 2022; Ibrahim & Sultana, 2006). An automatic grading line for large yellow croakers uses machine vision to classify fish by size and freshness based on sensory characteristics. This system allows rapid, non-destructive, and stable grading by estimating fish size through image analysis of features such as body shape, iris, and coloration (Wu, Zhuang, & Cui, 2019).

Cooling and Freezing Systems: Modern cooling and freezing technologies are essential to maintain fish quality, reduce post-harvest losses, and improve operational efficiency (Shawyer & Medina Pizzali, 2003). Effective use requires balancing economic feasibility, technical reliability, and environmental performance. Chilling lowers fish temperatures to 0-4 °C to inhibit microbial growth and enzymatic activity without freezing tissue water (Shawyer & Medina Pizzali, 2003).

Compact onboard ice machines have made sea ice production practical for fishing vessels. Commercial ice-makers produce ice at -9 to -20 °C with varying salt content. Constructed from corrosion-resistant materials such as stainless steel and aluminium, these systems may be installed onboard or in fish holds, while compressors and condensers are typically placed in the engine room. Seawater ice allows on-demand production and longer storage, especially where freshwater is limited, but it may increase costs and cause salt absorption (Shawyer & Medina Pizzali, 2003).

Fish preservation methods include embedding in crushed ice, storage in water-ice mixtures, or chilling in refrigerated seawater. Chilled Seawater (CSW) maintains fish at 0 to -2 °C, while Refrigerated Seawater (RSW) systems use onboard refrigeration and circulation pumps for continuous temperature control, either immersing the catch in chilled seawater tanks or spraying water over shelved fish. These systems enhance chilling uniformity and operational efficiency but involve higher installation and energy costs (Thanasarnsakorn, Taradol, Thimkrap, & Siriraksophon, 2005; Shawyer & Medina Pizzali, 2003).

Different ice types are used depending on operational needs: flake ice cools rapidly due to high surface area, block ice is preferred for transport and storage, and slush ice, composed of fine crystals in water, is pumpable and ideal for CSW tanks or holds (Shawyer & Medina Pizzali, 2003).

Vapour-compression refrigeration systems consist of a compressor, condenser, expansion valve, and evaporator (Figure 15). The compressor circulates refrigerant, the condenser removes heat, the expansion valve reduces pressure, and the evaporator absorbs heat to produce cooling (Thanasarnsakorn et al., 2005). Since refrigeration accounts for nearly half of a vessel's total power consumption, energy-efficient designs are critical (Ruiz, 2012). Alternative systems - adsorption, absorption, and ejector cooling - can improve efficiency and utilize waste heat. For example, activated carbon-methanol adsorption systems increased the coefficient of performance (COP) by 60%, while ammonia-water absorption systems achieved 90% heat recovery with a COP of 0.3 (Wang et al., 2006; Poku & Ogbonnaya, 2018). Computational studies on small vessels have demonstrated COP values above 4 and optimized temperature distribution for efficient onboard cooling (Budiyanto, Azharrisman, & Utama, 2023).

Environmental sustainability is of growing importance. Following the phase-out of Chlorofluorocarbons (CFCs), natural refrigerants such as ammonia (R-717) and carbon dioxide (R-744) have become efficient, eco-friendly alternatives to R-22 (Söylemez et al., 2022; Widell, Svendsen, & Nordtvedt, 2023). Waste-heat-driven systems can save up to 1,600 kg of fuel annually and reduce CO₂ emissions by ~3 t (Palomba, Aprile, Motta, & Vasta, 2017). Further optimizations show R-32 offers high exergetic efficiency, and CO₂-based trigeneration systems provide simultaneous refrigeration, hot water, and propulsion power, improving engine efficiency and reducing fuel use (Karakurt & Güneş, 2016; Liang et al., 2023).

In industrial fisheries, rapid cooling and freezing is essential for the quality of the product. Purse-seined fish are often stored whole in RSW or brine tanks until frozen. Larger vessels use blast freezers down to -35 °C, while smaller boats rely on motherships or refrigerated carriers for storage and transfer. Continuous

improvements in onboard refrigeration, particularly in Norwegian fleets, have significantly enhanced fish quality and storage capacity (Trygg Mat Tracking and IMCS Network, 2021; Digre et al., 2016).

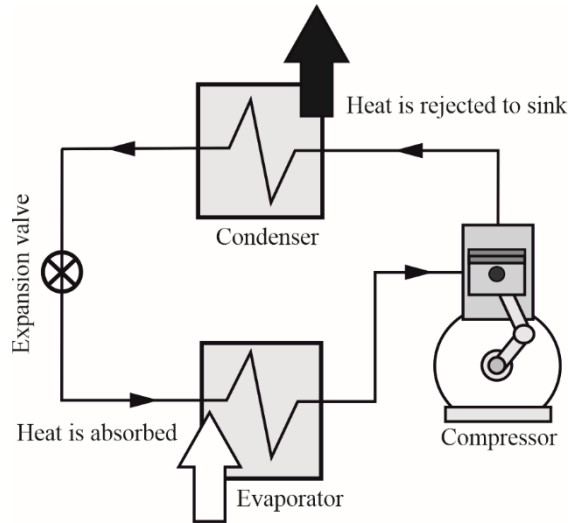


Figure 15. Schematic diagram of the refrigeration system on board a fishing vessel.

Automatic filleting and processing lines: Fish processing involves sequential steps to ensure product quality and safety, beginning with transport, stunning, bleeding, and chilling, followed by preparation operations such as gutting, descaling, and heading. Subsequent species- and product-specific operations include filleting, trimming, skinning, portioning, inspection, packing, weighing, and grading. Additional steps - mincing, freezing, thawing, and storage handling - preserve product integrity, supported by routine cleaning and knife maintenance (Baader, 2025).

The processing is centred on the production of fillets, with automation used for sorting, grading, gutting, sorting and filleting. Systems such as the *Scanvaegt* Fish Monitoring System (FMS) and *Scanvaegt* Image System use computer vision and artificial intelligence to classify fish according to species, size, and weight (Scanvaegt Systems, 2025). Modern filleting machines, including *Baader* models, incorporate gutting and filtration with computerised motors that adjust the instruments to the size of the fish, increasing accuracy and throughput (Choudhury & Bublit, 1997). The post-filtration steps, including the detection of parasites, remain partly manual, although optical and electromagnetic methods are showing promise for rapid industrial automation (Choudhury & Bublit, 1997).

Canning, used for species such as salmon, tuna, sardines and mackerel, involves the cleaning, salting, cooking, filling, searing and drying of the fish.

Automated systems such as Digicon 2000 optimise the weight of the fill by means of statistical control (Choudhury & Bublit, 1997). The production of surimi, predominantly from Alaska Pollock (*Theragra chalcogramma*), involves the extraction of muscle, washing, refining, mixing with cryoprotectants and freezing. For greater efficiency and consistency, automation and sensors control the flow of materials, water consumption, and timing (Choudhury & Bublit, 1997).

The smoked products - salmon, herring, mackerel and trout - are filleted and smoked in microprocessor-controlled systems, which regulate the temperature, humidity and duration of the smoke while monitoring for defects (Choudhury & Bublit, 1997). By-products, which account for 30-80% of the catch depending on the species, are transformed into fishmeal and oil by grinding, boiling, pressing and drying, with programmable logic controllers allowing centralised control (Choudhury & Bublit, 1997).

Overall, automation in fish processing remains partial and mainly focuses on data collection and control of signals. Fully computerised management - from harvesting to packaging - is not yet feasible and requires the use of sophisticated sensors and software. Integrated, user-friendly systems and stronger collaboration between industry and academia could speed up automation, supported by a steady supply of raw materials from the fisheries and aquaculture sectors (Choudhury & Bublit, 1997; Einarsdóttir, Guðmundsson, & Ómarsson, 2022).

6. ADVANTAGES OF MECHANIZATION

Mechanisation in fisheries means the use of machines, engines and automated systems for catching, handling, processing and transporting fish. This technological development has brought significant economic, operational, and safety benefits to the fishing sector.

Mechanised vessels and equipment - such as trawlers, purse seiners and hydraulic winches - allow for larger catch sizes while simultaneously reducing the number of people needed to carry out each operation. Integrated sonar, network sensors, and advanced surveillance systems are further increasing the efficiency and accuracy of the monitoring process (FAO, 2025b; ICCAT, 2008; Therms, Van Anrooy, Gudmundsson & Davy, 2023). Moreover, efficient engines extend the operational reach of fishing fleets, providing access to deeper and more remote areas, while reducing pressure on coastal stocks (Düzbastılar & Soykan, 2024).

In addition to operational improvements, mechanisation significantly reduces the physical strain, fatigue, and risk of injury, and enables smaller crews to carry out complex operations in an efficient manner (Nygaard, 1988). Hydraulic

handling systems also ensure a smooth transport of fish, while automated sorting and refrigeration systems on board maintain the uniform quality of the product and prolong the freshness of the product (Einarsdóttir, Guðmundsson & Ómarsson, 2022; Huse, 1975).

In addition, GPS-guided nets and sonar technologies improve the accuracy of the gear deployment and help to minimise bycatch and over-exploitation (FAO, 2025b; Valdemarsen, 2001). From an economic point of view, mechanisation promotes greater efficiency in catching, lower labour costs, and improved fuel efficiency, which in turn increases competitiveness and export potential (Düzbastılar & Soykan, 2024; Olsen, 1985). It also maintains large-scale fisheries by allowing for continuous, high-volume production, which feeds the global demand for seafood and strengthens the link between harvesting, processing, and logistics (Einarsdóttir, Guðmundsson, & Ómarsson, 2022; Regenstein & Regenstein, 1991; Wiyono, Raharjo & Permana, 2018).

Equally important, ergonomic and remote-controlled systems increase safety by reducing exposure to rough seas and minimising the risk of accidents (Valdemarsen, 2001; Ward & Hindmarsh, 2007; FAO, 2014; Wiyono et al., 2018; IMO, 2024b, 2025).

In short, mechanisation improves productivity, product quality, and crew safety while improving economic viability. However, these benefits depend on adequate investment, qualified operations, and good environmental management to ensure sustainable development.

7. CHALLENGES AND LIMITATIONS OF MECHANIZATION IN FISHERIES

Although mechanisation has transformed fishing into more efficient and productive activities, it has also created many technological, economic, environmental, and social challenges that require careful evaluation.

First, the purchase and maintenance of mechanised vessels, engines, winches, and hydraulic systems require significant investments, which are not readily available to small-scale fishermen (Düzbastılar, 2025; Nyborg, 1988; Trygg Mat Tracking & IMCS Network, 2021). Moreover, the heavy dependence on diesel engines makes mechanised fishing vulnerable to fluctuations in fuel prices and supply disruptions, which in turn lead to higher operating costs (Düzbastılar & Soykan, 2024).

From an environmental point of view, intensive mechanised fishing may damage benthic habitats, increase by-catches and discards, and contribute to overfishing, all of which pose a threat to marine biodiversity (Düzbastılar et al., 2023). Moreover, the efficient use of advanced technologies requires specialised

knowledge and training; without which inefficiencies, mechanical failures, and accidents are more likely to occur (Hermansson, 1978; Prado, 1990).

The social dimension of mechanisation is also of great importance. As machines replace manual labour, the employment opportunities of traditional fishermen may be reduced, which would increase economic inequality and favour the industrial operators (Ramenzoni, 2017). Moreover, the harsh marine environment speeds up the wear and tear of equipment, and maintenance problems in remote areas may lead to delays in operation and safety risks (Düzbastılar & Soykan, 2024).

Effective governance frameworks are therefore necessary for the regulation of mechanised fishing fleets, the enforcement of quotas, and the prevention of overfishing (Andersson & Ngazi, 1998; Düzbastılar, 2025; Eigaard, Marchal & Gislason, 2014). Another practical limitation is the adaptability of the gear - mechanised gears often fail in small, shallow, or structurally complex fisheries, where the artisanal methods remain more efficient and environmentally sound (Düzbastılar, 2025). Despite the reduction in physical workload, mechanical failures, collisions and accidents continue to pose a safety risk for crews (Atacan and Düzbastılar, 2023).

It is clear that mechanisation increases operational capacity and efficiency, but it also creates new economic, environmental and social bottlenecks. Sustainable mechanisation, therefore, requires comprehensive planning, effective regulation, targeted training, and balanced policy development to reconcile technological progress with environmental and social responsibility.

8. CONCLUSION

By increasing productivity, efficiency, and operational capacity, mechanization has significantly changed the fishing industry. The incorporation of advanced fishing gear, hydraulic systems, and mechanised vessels has improved product quality and processing efficiency, increased catch volumes, and opened up previously unreachable fishing grounds. Mechanization has improved industrial and large-scale fisheries from an economic standpoint, boosting national economic growth and competitiveness.

However, these advances are accompanied by continuing challenges, such as high investment and maintenance costs, fuel dependency, environmental pressures, and the need for specialised technical skills. In many regions, mechanisation has also increased socio-economic disparities by favouring capital-intensive operators and reducing the employment opportunities of small-scale fishermen. Effective governance, regulatory oversight, and targeted training are therefore essential to ensure that mechanisation develops sustainably and inclusively.

Finally, mechanisation has undeniable operational and economic benefits, but its long-term sustainability depends on a balanced integration of technology with environmental protection and social justice. Achieving this balance requires coordinated policies, adaptive management and continuous innovation to reconcile technological progress with the responsibility for the environment and social well-being.

REFERENCES

- Andersson, J., & Ngazi, Z. (1998). Coastal communities' production choices, risk diversification, and subsistence behavior: Responses in periods of transition (A case study from Tanzania). *Royal Swedish Academy of Sciences*, 27(8), 686-693. Retrieved from <https://www.jstor.org/stable/i399329>
- Atacan, C., & Düzbastılar, F. O. (2023). Determination of risk perception in small-scale fishing and navigation. *Ege Journal of Fisheries and Aquatic Sciences*, 40(1), 1-14. doi:10.12714/egejfas.40.1.01
- Aydın, C., & Düzbastılar, F. O. (2011). Otter boards' performances and design criteria. *Ege Journal of Fisheries and Aquatic Sciences*, 28(4), 127-136.
- Baader. (2025). Retrieved October 30, 2025, from <https://www.baader.com/fish>
- Ben-Yami, M. (1994). *Purse Seining Manual*. Wiley. Retrieved September 22, 2025, from @inproceedings{BenYami1994PurseSM
- Beverly, S., Chapman, L., & Sokimi, W. (2003). *Horizontal Longline Fishing Methods and Techniques A Manual for Fishermen*. 130. Secretariat of the Pacific Community. Retrieved September 25, 2025, from <http://coastfish.spc.int/Sections/Development/FDSPublications/FDSManuals/HLL/HLL.pdf>
- Bonnett, D. E. (1988). Vessel vital signs monitoring. *World Symposium on Fishing Gear and Fishing Vessel Design* (pp. 380-384). St. John's: Newfoundland and Labrador Institute of Fisheries and Marine Technology.
- Budiyanto, M. A., Azharrisman, F., & Utama, A. F. (2023). Cooling performance of modular fish hold for 30 gross-tonnage fishing vessel. *Journal of Marine Engineering & Technology*, 22(1), 38-44. doi:10.1080/20464177.2022.2094655
- Carlson, C. (1975). Recent Developments in Deck Gear. In J.-O. Traung, & F. a. Organization (Ed.), *Fishing Boats of the World* (pp. 475-480). Surrey, England: Fishing News (Books) Limited.
- Carral, J., Carral, L., Lamas, M., & Rodríguez, M. J. (2015). Fishing grounds' influence on trawler winch design. *Ocean Engineering*, 102, 136-145. doi:10.1016/j.oceaneng.2015.04.055
- Choudhury, G. S., & Bublitz, C. G. (1997). Computer-Based Controls in Fish Processing Industry. In G. S. Mittal, *Computerized Control Systems in the Food Industry* (pp. 513-538). CRC Press.
- Digre, H., Tveit, G. M., Solvang-Garten, T., Eilertsen, A., & Aursand, I. G. (2016). Pumping of mackerel (*Scomber scombrus*) onboard purse seiners, the effect

on mortality, catch damage and fillet quality. *Fisheries Research*, 176, 65-75. doi:10.1016/j.fishres.2015.12.011

- Düzbastılar, F. O. (2025). Addressing Challenges in Small-Scale Fisheries: Insights from the Aegean Sea. In H. Dereli, & Z. Tosunoğlu, *Beyond Sustainable Fisheries: Conservation and Sustainability, Small-Scale Fisheries and Creating Value Chains* (pp. 99-126). İzmir, Türkiye: Duvar Yayınları.
- Düzbastılar, F. O., & Hoşsucu, H. (1997). The mechanical systems of purse seine vessel. *Mediterranean Fisheries Congress* (s. 119-126). İzmir: Ege University, Faculty of Fisheries.
- Düzbastılar, F. O., & Soykan, O. (2024). Alternative Energy Approaches in Fishing Vessels. In B. Kılıç, *Advances in Engineering Sciences* (pp. 55-87). Ankara, Türkiye: Platanus Publishing. doi:10.5281/zenodo.14577732
- Düzbastılar, F. O., Tosunoğlu, Z., Ceyhan, T., Kaykaç, M. H., Aydın, C., Güleç, Ö., & Metin, G. (2023). A variation in the mortality of European Anchovy and European Pilchard after sieving and discarding from a purse seine fishery in the Eastern Mediterranean. *Turkish Journal of Fisheries and Aquatic Sciences*, 23(1), 13. doi:10.4194/TRJFAS21516
- Eigaard, O. R., Marchal, P., & Gislason, H. (2014). Technological development and fisheries management. *Reviews in Fisheries Science & Aquaculture*, 22(2), 156–174. doi: 10.1080/23308249.2014.899557
- Einarsdóttir, H., Guðmundsson, B., & Ómarsson, V. (2022). Automation in the fish industry. *Animal Frontiers*, 12(2), 32-39. doi:10.1093/af/vfac020
- Engelhard, G. H. (2005). Catalogue of Defra historical catch and effort charts: six decades of detailed spatial statistics for British fisheries. *Sci. Ser. Tech. Rep.*, Cefas.
- Erikson, U., Uglem, S., & Greiff, K. (2021). Freeze-chilling of whitefish: Effects of capture, on-board processing, freezing, frozen storage, thawing, and subsequent chilled storage-A review. *Foods*, 10(2661), 1-26. doi:10.3390/foods10112661
- FAO. (2014). *The State of World Fisheries and Aquaculture: Opportunities and challenges*. Rome: FAO. Retrieved July 20, 2024, from <https://www.fao.org/3/i3720e/i3720e.pdf>
- FAO. (2025a). *Fishery and Aquaculture Statistics – Yearbook 2022*. FAO Yearbook of Fishery and Aquaculture Statistics. FAO, Rome. doi:10.4060/cd4312en
- FAO. (2025b). *Fishing Technology Equipments. Electronic instruments. Technology Fact Sheets In: Fisheries and Aquaculture*. Updated 2008-11-28. Retrieved September 23, 2025, from <https://www.fao.org/fishery/en/equipment/electronicinstruments/en>

- FAO. (2025c). Fishing Gear types. Purse seines. Technology Fact Sheets. In: Fisheries and Aquaculture. Retrieved from <https://www.fao.org/fishery/en/geartype/249/en>
- FAO. (2025d). Fishing Vessel types. Drum seiners. Technology Fact Sheets. In: Fisheries and Aquaculture. Updated 2024/08/22 [Cited Tuesday, November 4th 2025]. <https://www.fao.org/fishery/en/vesseltype/733>
- Gabriel, O., Lange, K., Dahm, E., & Wendt, T. (2005). Von Brandt's Fish Catching Methods of the World (4 ed.). Wiley-Blackwell.
- Gümüş, B., Balaban, M. Ö., & Ünlüsay, M. (2011). Machine vision applications to aquatic foods: A review. *Turkish Journal of Fisheries and Aquatic Sciences*, 11, 171-180. doi:10.4194/trjfas.2011.0124
- Hammer, B. W. (1974). Small Trawler Hydraulic Power System. Technical Report Series of the Industrial Development Branch.
- He, P., Chopin, F., Suuronen, P., Ferro, R. S., & Lansley, J. (2021). Classification and illustrated definition of fishing gears. Rome: FAO Fisheries and Aquaculture Technical Paper, FAO. doi:10.4060/cb4966en
- Hermansson, B. (1978). FAO Fishing Manuals: Training Fishermen at Sea (Vol. 144376). Farnham, Surrey, England: Food and Agriculture Organization of the United Nations by Fishing News Books Ltd.
- Huse, H. V. (1975). Hydraulic Deck Equipment. In J.-O. Traung, & F. a. Nations (Ed.), *Fishing Boats of the World* (pp. 488-493). Surrey, England: Fishing News (Books) Limited.
- Ibrahim, M. Y., & Sultana, S. (2006). Study on fresh fish sorting techniques. 2006 IEEE International Conference on Mechatronics, (pp. 462-467). Budapest. doi:10.1109/ICMECH.2006.252571
- ICCAT. (2008, January 28). Purse seine. Description of fishing with purse seine nets. ICAAT Manuel, 3.1.1. Retrieved October 13, 2025, from https://www.iccat.int/Documents/SCRS/Manual/CH3/CHAP%203_1_1_PS_ENG.pdf
- IMO. (2024a). GMDSS Manual: Global Maritime Distress and Safety System Manual (2024 Edition). International Maritime Organization. doi:10.62454/KJ970E
- IMO. (2024b). International Regulations for the Safety of Fishing Vessels. Retrieved June 26, 2024, from <https://wwwcdn.imo.org/localresources/en/About/Conventions/Documents/Consolidated%20text%20of%20the%20Agreement.pdf>
- IMO. (2025). Convention on the International Regulations for Preventing Collisions at Sea, 1972 (COLREGs). September 29, 2025 tarihinde

<https://www.imo.org/en/about/conventions/pages/colreg.aspx> adresinden alındı

- Karakurt, A. S., & Güneş, Ü. (2016). Performance analysis of multipurpose refrigeration system (MRS) on fishing vessel. *Polish Maritime Research*, 23(2), 48-56. doi:10.1515/pomr-2016-0020
- Li, P., Zhang, S., Zheng, H., Fan, X., Shi, Y., Wu, Z., & Zhang, H. (2025). YOLO-DFAM-based onboard intelligent sorting system for *Portunus trituberculatus*. *Fishes*, 10(364), 1-25. doi:10.3390/fishes10080364
- Liang, Y., Zhu, Y., Sun, Z., Ye, K., Wu, J., & Lu, J. (2023). Feasibility assessment of a CO₂ based power, cooling, and heating system driven by exhaust gas from ocean-going fishing vessel. *Journal of Cleaner Production*, 406(137058), 1-15. doi:10.1016/j.jclepro.2023.137058
- Lin, L., Chen, Z., Xu, Z., & Wang, Z. (2016). Design and experimental study of the four-drum net winch. 6th International Conference on Advanced Design and Manufacturing Engineering (ICADME 2016) (pp. 82-88). Zhuhai: Published by Atlantis Press.
- Liu, Q., Wang, Y., & Xu, M. (2025). Fishing vessel trawl winch tension control: A BP neural network PID feedforward control method based on NARX neural network prediction. *Processes*, 13(2001), 1-17. doi:10.3390/pr13072001
- Mathiassen, J. R., Janssøn, S., Veliyulin, E., Njaa, T., Lønseth, M., Bondø, M., . . . Skavhaug, A. (2006). Automatic weight and quality grading of whole pelagic fish. *NFTC 2006 Nor-Fishing Technology Conference*, (pp. 1-8). Trondheim, Norway.
- Mathiassen, J. R., Misimi, E., Østvik, S. O., & Aursand, I. G. (2012). Computer vision in the fish industry. In D.-W. Sun, *Computer Vision Technology in the Food and Beverage Industries* (pp. 352-378). Woodhead Publishing. doi:10.1533/9780857095770.3.352
- Metin, C., Özbilgin, H., Tosunoğlu, Z., Gökçe, G., Aydın, C., Metin, G., . . . Tokaç, A. (2005). Effect of square mesh escape window on codened selectivity for three fish species in the Aegean Sea. *Turkish Journal of Veterinary Sciences*, 29, 461-468. Retrieved from <https://journals.tubitak.gov.tr/veterinary/vol29/iss2/40/>
- Mishra, M., Kulkarni, G. N., & Ghosh, S. K. (2021). Fresh fish: Handling, transportation and preservation. *Biotica Research Today*, 3(5), 409-412.
- Miyake, P. M. (2004). A brief history of the tuna fisheries of the world. In W. H. Bayliff, J. I. de Leiva Moreno, & J. Majkowski (Eds.), *Second meeting of the Technical Advisory Committee of the FAO Project "Management of Tuna*

- Fishing Capacity: Conservation and Socio-economics*", Madrid, Spain, 15–18 March 2004 (FAO Fisheries Proceedings No. 2, pp. 23–50). Rome: FAO.
- Morgan, P. W. (2006). Brief history of fibers from synthetic polymers. *Journal of Macromolecular Science: Part A - Chemistry*, 15(6), 1113-1131 . doi:10.1080/00222338108066456
- Naisbett-Jones, L., Mudrak, P., & Raborn, S. (2025). Characterization of Bycatch for the Menhaden Purse Seine Fishery Occurring off the Coast of Louisiana. LGL Ecological Research Associates. Texas: Gulf States Marine Fisheries Commission. Retrieved October 12, 2025, from chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.wlf.louisiana.gov/assets/Resources/Publications/Saltwater_Fish/CharacterizationOfBycatch_July_8_Presentation_LGL.pdf
- Nickum, G. C. (1975). Pacific Coast Processing Vessels. In J.-O. Traung, & F. a. Nations (Ed.), *Fishing Boats of the World* (pp. 513-534). Surrey, England: Fishing News (Books) Limited.
- Nomura, M. (1985). *Fishing Techniques III - Discussion on net and rope material, basic gear behaviour in fishing operation, model net experiment system and fishing instruments*. Tokyo, Japan: Japan International Cooperation Agency (JICA).
- Nygaard, W. (1988, December). Australian Fisheries. Developments in longlining, 47(11), 30-33. Retrieved September 24, 25, from <https://nla.gov.au/nla.obj-749290267/view?sectionId=nla.obj-754169357&partId=nla.obj-749295795#page/n0/mode/1up>
- Olsen, K. B. (1985). European Fish Handling and Holding Methods. In J. F. Roache (Ed.), *New Directions in Fisheries Technology-Proceedings of the Atlantic Fisheries Technology Conference* (pp. 46-56). Ottawa, Ontario: Canadian Special Publication of Fisheries and Aquatic Sciences 85. Retrieved October 30, 2025, from https://publications.gc.ca/collections/collection_2016/mpo-dfo/Fs41-31-85-eng.pdf
- Palomba, V., Aprile, M., Motta, M., & Vasta, S. (2017). Study of sorption systems for application on low-emission fishing vessels. *Energy*, 134, 554-565. doi:10.1016/j.energy.2017.06.079
- Petrel International. (2025, October 5). Retrieved from Commercial Fishing Dech Equipment: <https://www.petrelequipment.com/>
- Pitcher, T. J. (2012). Fisheries. In I. Spellerberg, & D. E. Vasey, *Berkshire Encyclopedia of Sustainability: The Future of Sustainability* (Vol. 10, pp. 108-115). Great Barrington, MA, USA: Berkshire Publishing Group. doi:10.2307/jj.9561398.24

- Poku, R., & Ogbonnaya, E. A. (2018). Thermodynamics analysis of recovering heat for refrigeration system of a fishing vessel. *Nigerian Journal of Technology (NIJOTECH)*, 37(4), 936 – 944. doi:10.4314/njt.v37i4.11
- Prado, J. (1990). *Fishermen's Workbook*. Fishing News Books Oxford (Food and Agriculture Organization of the United Nations by . .). Retrieved June 2, 2025, from chrome-extension://efaidnbmnnnibpcajpcgclefindmkaj/https://openknowledge.fao.org/server/api/core/bitstreams/0e97b755-e4ba-437d-936b-30feac9fd8c8/content
- Purse-seine Observer Guide. (2021). 140. (S. O. Programme, Ed.) Noumea, New Caledonia: Pasific Community. Retrieved October 9, 2025, from https://www.spc.int/digitalibrary/get/4aya
- Raborn, S., Mudrak, P., & Naisbett-Jones, L. (2025). Characterization of Bycatch for the Menhaden Purse Seine Fishery Occurring off the Coast of Louisiana. LGL Ecological Research Associates. Texas: Gulf States Marine Fisheries Commission. Retrieved October 11, 2025, from chrome-extension://efaidnbmnnnibpcajpcgclefindmkaj/https://www.wlf.louisiana.gov/assets/Resources/Publications/Saltwater_Fish/Characterization-of-Bycatch-for-the-Menhaden-Purse-Seine-Fishery-Occurring-off-the-Coast-of-Louisiana-LGL.pdf
- Ramenzoni, V. C. (2017). Reconstructing the history and the effects of mechanization in a small-scale fishery of Flores, Eastern Indonesia (1917–2014). *Frontiers in Marine Science*, 4(65), 1-20. doi:10.3389/fmars.2017.00065
- Regenstein, J. M., & Regenstein, C. E. (1991). *Introduction to Fish Technology*. USA: Springer.
- Ruiz, V. (2012). Analysis of existing refrigeration plants onboard fishing vessels and improvement possibilities. *Second International Symposium on Fishing Vessel Energy Efficiency E-Fishing 2012*, (pp. 1-8). Vigo, Spain.
- Sahrhage, D., & Lundbeck, J. (2012). *A History of Fishing*. Springer Science & Business Media. doi:10.1007/978-3-642-77411-9
- Scanvaegt systems. (2025). Retrieved October 30, 2025, from https://www.scanvaegt.com/en
- Shawyer, M., & Medina Pizzali, A. F. (2003). The use of ice on small fishing vessels. Rome: FAO. Retrieved from https://www.fao.org/4/y5013e/y5013e00.htm
- SOLAS. (2000). Retrieved September 26, 2025, from IMO-Electronic Nautical Charts (ENC) and Electronic Chart Display and Information Systems (ECDIS): https://www.imo.org/en/ourwork/safety/pages/electroniccharts.aspx

- Söylemez, E., Widell, K. N., Gabriell, C. H., Ladam, Y., Lund, T., & Hafner, A. (2022). Overview of the development and status of carbon dioxide (R-744) refrigeration systems onboard fishing vessels. *International Journal of Refrigeration*, 40, 198-212. doi:10.1016/j.ijrefrig.2022.05.007
- Strachan, N. J. (1994). Sea trials of a computer vision based fish species sorting and size grading machine. *Mechatronics*, 4(8), 773-783. doi:10.1016/0957-4158(94)90052-3
- Szczepanek, M. (2017). Demand for Electricity for Hydraulic Power Systems at Fishing Cutters. In T. 5. 2017 (Ed.), (pp. 368-373). Prague.
- Tetley, L., & Calcutt, D. M. (1991). *Electronic Aids to Navigation: Position Fixing*. Baltimore: Edward Arnold.
- Thanasarnsakorn, S., Taradol, A., Thinkrap, T., & Siriraksophon, S. (2005). *On Board Fish Handling and Preservation Technology*. Phrasamutchedi, Samut-Prakarn, Thailand: Southeast Asian Fisheries Development Center - Training Department.
- Thermes, S., Van anrooy, R., Gudmundsson, A., & Davy, D. (2023). Classification and definition of fishing vessel types second edition. *FAO Fisheries and Aquaculture Technical Paper*, (267), 1-45,47-53,55-57,59-85, I, V, X. Retrieved from <https://www.proquest.com/scholarly-journals/classification-definition-fishing-vessel-types/docview/2873570443/se-2>
- Tosunoglu, Z., Ceyhan, T., Güleç, Ö., Düzbastılar, F. O., Kaykaç, M. H., Aydın, C., & Metin, G. (2021). Effects of lunar phases and other variables on CPUE of European Pilchard, *Sardina pilchardus*, caught by purse seine in the Eastern Mediterranean. *Turkish Journal of Fisheries and Aquatic Sciences*, 21, 283-290. doi:10.4194/1303-2712-v21_6_03
- Trygg Mat Tracking and IMCS Network. (2021). *A MCS PRACTITIONERS INTRODUCTORY GUIDE TO: PURSE SEINE FISHING*. 23. (T. M. (TMT), Ed.) Oslo, Norway. Retrieved October 8, 2025, from www.imcsnet.org/library-search
- Trygg Mat Tracking and IMCS Network. (2023). *A MCS PRACTITIONERS INTRODUCTORY GUIDE TO: TRAWL FISHING*. (T. M. (TMT), Ed.) Oslo, Norway. Retrieved October 15, 2025, from <https://imcsnet.org/resource/mcs-practitioners-introductory-guide-trawl-fishing>
- Valdemarsen, J. W. (2001). Technological trends in capture fisheries. *Ocean & Coastal Management*, 44, 635-651. doi:10.1016/S0964-5691(01)00073-4
- Velu, M. (1972). Design considerations for trawl winches. *Fishery Technology*, 9(2), 159-168. Retrieved from <http://hdl.handle.net/1834/33602>

- von Brandt, A. (1975). Vessels and Gear. In J.-O. Traung, & F. a. Nations (Ed.), *Fishing Boats of the World* (pp. 474-475). Surrey, England: Fishing News (Bokks) Limited. Retrieved September 25, 2025
- Walsh, S. J., Hickey, W. H., Porter, J., Delouche, H., & McCallum, B. R. (2009). *NAFC Survey Trawl Operations Manual*. Version 1.0, 195. Newfoundland Region, St. John's.: Fisheries and Oceans, Northwest Atlantic Fisheries Centre.
- Wang, L. W., Wang, R. Z., Wu, Y., Xu, Y. X., & Wang, S. G. (2006). Design, simulation and performance of a waste heat driven adsorption ice maker for fishing boat. *Energy*, 31(2-3), 244-259. doi:10.1016/j.energy.2005.03.006
- Wang, S. G., & Wang, R. Z. (2005). Recent developments of refrigeration technology in fishing vessels. *Renewable Energy*, 30(4), 589-600. doi:10.1016/j.renene.2004.03.020
- Wang, X., Wang, L., Song, Y., Deng, J., & Zhan, Y. (2021). Optimal design of two-stage ejector for subzero refrigeration system on fishing vessel. *Applied Thermal Engineering*, 187(116565), 1-12. doi:10.1016/j.applthermaleng.2021.116565
- Ward, P., & Hindmarsh, S. (2007). An overview of historical changes in the fishing gear and practices of pelagic longliners, with particular reference to Japan's Pacific fleet. *Reviews in Fish Biology and Fisheries*, 17, 501-516. doi:10.1007/s11160-007-9051-0
- WÄRTSILÄ. (2025, October 1). *Encyclopedia of Marine and Energy Technology*: <https://www.wartsila.com/encyclopedia> adresinden alındı
- Widell, K. N., Svendsen, E. S., & Nordtvedt, T. S. (2023). Possibilities of ice slurry systems onboard fishing vessels. 10th Conference on Ammonia and CO2 Refrigeration Technologies, (pp. 1-8). Ohrid. doi:10.18462/iir.nh3-co2.2023.0017
- Wiyono, E. S., Raharjo, S. S., & Permana, S. M. (2018). Fishermen acceptance on introduction of fishing technology: perception and its development strategies. *AACL Bioflux*, 11(3), 666-667. Retrieved from <http://www.bioflux.com.ro/aac>
- Wu, Y., Zhuang, R., & Cui, Z. (2019). Automatic sorting system of large yellow croaker based on machine vision. *International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS)*, (pp. 233-237). Shenzhen. doi:10.1109/HPBDIS.2019.8735486
- Yagi, K. (1975). Japanese Long Lining. In J.-O. Traung, & F. a. Nations (Ed.), *Fishing Boats of the World* (pp. 503-504). Surrey, England: Fishing News (Books) Limited.

CHAPTER 3

Microplastics in Cosmetics and Personal Care Products

Sebnem Gun¹

¹ Research Assistant, Istanbul University – Cerrahapasa, 0009-0000-5240-1322

1. INTRODUCTION

Micropollutants are compounds present in water and wastewater at concentrations of micrograms or less, yet are believed or have been established to exert detrimental effects on humans and the environment, even at these levels (Bhatt, Bhandari, & Bilal, 2022). Micropollutants can encompass both organic and inorganic contaminants. Micro-pollutants present in aquatic ecosystems are predominantly anthropogenic in origin. Examples of these substance categories encompass pharmaceuticals for human and veterinary use, personal care products, industrial chemicals, solvents, agricultural chemicals (pesticides), nanomaterials, and disinfection by-products (Doğruel et al., 2022).

Micropollutants can infiltrate aquatic ecosystems via multiple channels. Traditional wastewater treatment facilities are not engineered for the elimination of micropollutants. Consequently, owing to their resilient structure, several micropollutants are released from wastewater treatment facilities untreated. The efficiency of micropollutant removal in wastewater treatment facilities fluctuate based on the plant's performance and technology, wastewater flow rate, hydraulic retention time, and climatic circumstances. Recently, sophisticated oxidation techniques and membrane technologies have been utilized for the elimination of micropollutants from aquatic ecosystems (Üstün Odabaşı et al., 2020).

The cosmetics industry is characterized as a global sector encompassing the manufacture and distribution of personal care and beauty products. It provides a diverse array of items, encompassing cosmetics, skincare, haircare, and fragrances. The Turkish Medicines and Medical Devices Agency's Cosmetic Products Consumer Information Guide defines cosmetic products in Article 2 of Law No. 5324 as "preparations or substances intended for application to various external parts of the human body, including the epidermis, nails, hair, and teeth, as well as the oral mucosa, with the sole and primary purpose of cleansing, scenting, altering appearance, correcting, protecting against body odor, and maintaining these parts in optimal condition" (Özçelik & Bebekli, 2015).

The cosmetics industry is a significant contributor to micropollutants. Micropollutants in cosmetic goods manifest as microbeads (Leslie, 2014). Microbeads are generally composed of plastic particles or other synthetic substances utilized to improve product efficacy or deliver cosmetic benefits. Microscopic plastic particles in cosmetic goods are a threat to both environmental integrity and human health, irrespective of their concentration. Consequently, it is essential to investigate the occurrence of micropollutants in cosmetic products and to promote awareness regarding this matter (Giustra et al., 2024).

This section discusses micropollutants derived from the cosmetics sector. The examination focuses on alternatives to microbeads, their environmental consequences, and the associated regulatory regulations in cosmetic items.

2. MICROBEADS

Microplastics are mainly categorized into two types: main and secondary. Secondary microplastics originate from the degradation of bigger plastic materials over time, whereas primary microplastics are purposely manufactured plastic particles, notably utilized as microbeads in cosmetic items. Secondary microplastics originate from the decomposition of bigger plastic materials over time (Zhou et al., 2023). Microplastic particles present in personal care and cosmetic items have been recognized as a major source and conduit for enduring and potentially detrimental plastics that can infiltrate marine ecosystems. The quotidian application of personal care and cosmetic goods directly discharges plastic particles into wastewater, as these products are predominantly washed or rinsed during or subsequent to usage. Recent research indicates that there are 5.25 trillion plastic particles on the ocean's surface, weighing around 269,000 tonnes (Bhattacharya, 2016).

Microplastics originating from cosmetics, referred to as microbeads, are extensively utilized in personal care products for multiple purposes, including binding agents, volume enhancers, emulsifiers, film formers, viscosity regulators, opacifying agents, and glitter (Zhou et al., 2023). Additionally, they provide qualities like teeth polishers for skin and oral care, gelling agents for denture adhesives, moisturizers, and stabilizers for sunscreens, rendering them a multifaceted element in cosmetic and care product compositions (Giustra et al., 2024).

Plastic microbeads are commonly utilized in a range of products intended for body and facial washing, especially in formulations aimed at providing a specialized exfoliating action or eliminating persistent stains. These components are typically favored for their exfoliating properties, with polyethylene microbeads being the most prevalent variant. Furthermore, other materials including polyurethane and ethylene-vinyl acetate (EVA) copolymer are utilized for this purpose. One advantage of utilizing plastic microbeads for exfoliation and scrubbing is their moderate exfoliating impact, which is linked to favorable skin tolerability. Moreover, it has been observed that these microbeads are chemically inert, devoid of volatile components, do not induce sensitivity, and do not irritate the skin (Giustra et al., 2024). Owing to their characteristics, they are extensively utilized as an ingredient in the development of skincare products (Rahim et al., 2022).

The main microplastics used in cosmetic products are as follows:

- **Polyethylene (PE):** Polyethylene is a prevalent synthetic polymer utilized as a film-forming agent and viscosity modifier in cosmetic formulations. PE microparticles have been used into formulations, especially in facial and body exfoliating products, serving as an exfoliating agent. Since 2018, they have been prohibited in cosmetic items due to their detrimental effects on marine ecosystems (Guzik et al., 2023).
- **Polypropylene (PP):** Polypropylene, synthesized through the polymerization of 1-propene gas sourced from petroleum and natural gas, is a prevalent synthetic polymer utilized in the cosmetics industry. This polymer, known for its exceptional physical and chemical resilience, is often favored to achieve the appropriate consistency and smoothness in product compositions. It functions as a viscosity regulator, especially in cosmetics like foundation and nail polish; furthermore, it serves as an exfoliant in some formulations and aids in refining the skin's surface. Its insoluble structure and low biodegradability render it a possible micro-pollutant from an environmental standpoint (Guzik et al., 2023).
- **Polymethyl methacrylate (PMMA):** PMMA is a transparent thermoplastic polymer synthesized via the polymerization of methyl methacrylate monomer. It is extensively utilized in cosmetic formulations, including nail polish, lip gloss, and mascara, owing to its film-forming characteristics. Additionally, it has been indicated that PMMA microbeads serve as an absorbent medium for the delivery of active ingredients and in modulating the gloss or matte characteristics of products, thus aiding in the diminishment of age-related aesthetic flaws (Becker et al., 2011; Guzik et al., 2023).
- **Polytetrafluoroethylene (PTFE):** This homopolymer, synthesized via the polymerization of tetrafluoroethylene, serves as an ingredient in cosmetic compositions. This compound is especially favored in products like foundation, face powder, and makeup setting powder to attain a matte look. Moreover, it is regarded as a useful element in formulations owing to its compatibility with many products and its attributes, including non-stickiness, moisture resistance, and durability (Giustra et al., 2024).
- **Polyethylene Terephthalate (PET):** Obtained from terephthalic acid and ethylene glycol. It serves as a mattifying agent in creams, lotions, and cosmetics, while its microplastic particles operate as a carrier

material for the regulated dispersion and release of active compounds. Moreover, it serves as an addition in the composition of body paints, nail polish, lip gloss, and hair dye products (Becker et al., 2014).

- **Polyurethane (PU):** Polyurethane microbeads are utilized to provide a peeling effect, enhance product texture, and serve as a binding agent. Moreover, these microbeads possess the capacity to absorb surplus sebum on the skin's surface, rendering them a favored option in formulas designed for oil regulation. Moreover, polyurethane derivatives are regarded as additions in cosmetic formulations, including nail polish, lip gloss, and eye shadow, to augment color intensity and promote product efficacy (Guzik et al., 2023).
- **Polyethylene glycols (PEGs) and polypropylene glycols (PPGs):** PEGs and PPGs are extensively utilized in pharmaceutical formulations for drug delivery purposes. In cosmetic formulations, they are commonly utilized in creams, lotions, and serums as humectants or emulsifiers. The principal function of polyethylene glycol (PEG) derivatives in these applications is to enhance formulation stability, provide uniform distribution of active components, and facilitate moisture retention in the skin (Giustra et al., 2024).

Other types of microplastics commonly used in cosmetic products include polybutylene terephthalate (PBT), polystyrene (PS), various nylon derivatives, acrylic polymers, and polyurethane cross-linked polymers (Giustra et al., 2024).

2.1. Environmental Impacts of Microbeads

Microbeads can improve the efficacy of cosmetic products; nevertheless, they also have environmental drawbacks. The inaugural study to emphasize the environmental consequences of microplastics in cosmetic products was executed by Zitko and Hanlon (1991). Numerous studies indicate that exposure to elevated concentrations of microplastic beads results in observable toxicity symptoms across a range of organisms, from primary producers to marine invertebrates, as well as mammalian and human tissue systems (Miller et al., 2020; Marcharla et al., 2024).

The ecological effects of microbeads are predominantly evident in aquatic habitats. These particles can be consumed by various animals, ranging from plankton to fish, and may lead to detrimental impacts like digestive obstructions, energy depletion, heightened oxidative stress, and diminished reproductive

efficacy (Vo & Pham, 2021). Moreover, the surface architecture of microbeads facilitates the adsorption of chemicals, including heavy metals and persistent organic pollutants. Consequently, they serve a crucial function in the environment, acting both as physical carriers and as carriers of chemical pollutants (Ranjan et al., 2024).

Recent years have seen microplastic exposure emerge as a substantial problem for human health. Microbeads can infiltrate the body via drinking water, seafood, and air; laboratory research indicates that these particles can induce cellular and DNA damage (Prata et al., 2020; Zhang et al., 2025). Nonetheless, the specific amounts and situations under which these effects result in enduring health hazards are inadequately comprehended. In response to mounting scientific evidence, the European Union has progressively prohibited the use of microplastics in cosmetic items through the REACH legislation, effective in 2023, and has encouraged producers to adopt biodegradable alternatives (European Commission, 2023).

2.2. Alternatives to Microbeads

The incorporation of microplastics in cosmetic products is presently seen as a significant environmental issue. The extensive manufacture, promotion, and utilization of cosmetic goods containing microplastics globally complicates efforts to control or entirely eradicate these chemicals. Consequently, safer and more sustainable replacements to these potentially detrimental components must be devised. Global plastic production has exhibited a persistent increase for more than fifty years; it rose from around 1.9 million tonnes in the 1950s to hundreds of millions of tonnes in the 2010s (Bhattacharya, 2016).

Multiple investigations are currently underway to mitigate the detrimental impacts of microplastics, particularly those derived from personal care and cosmetic items, on marine ecosystems and other settings. The primary objective of these initiatives is to substitute conventional synthetic microplastics with eco-friendly, sustainable, and biodegradable materials. The cosmetics industry is aggressively seeking alternatives to conventional microplastic components, which are classified as natural, semi-synthetic, and synthetic (Giustra et al., 2024).

Prior to the incorporation of microbeads in cosmetic formulations, natural substances such as cocoa beans, pulverized almonds, apricot kernels, sea salt, pumice, beeswax, walnut shells, oatmeal, and powdered sugar were prevalently utilized in personal care and beauty products (Perschbacher, 2016). Nonetheless, it has been noted that these natural constituents may induce tiny lacerations in the

skin, resulting in irritation, edema, and indications of premature aging. Consequently, microbeads have increasingly emerged as the favored option owing to their resilience and reduced manufacturing expenses. Nonetheless, the growing recognition of their detrimental environmental effects has expedited initiatives to create eco-friendly and biodegradable substitutes for microplastics (Rahim et al., 2022).

2.2.1. Natural Alternatives

Heightened consumer awareness regarding the adverse environmental effects of synthetic polymers is driving the advancement of biopolymer manufacturing from natural sources. Due to their eco-friendly and safe characteristics, natural polymers have secured a significant role in contemporary cosmetic compositions. Biopolymers, including starch, cellulose, alginate, chitosan, lignin, pectin, xanthan gum, agar, hyaluronic acid, guar gum, gelatin, collagen, and keratin, are transformed into microparticles and employed in diverse cosmetic formulations (Morganti, Morganti, & Coltelli, 2023). These compounds serve as stabilizers, viscosity modifiers, or structural modifiers in skincare, haircare, and cosmetic products. Their secure designs, biocompatibility, and eco-friendliness render them appealing to both producers and consumers. Moreover, semi-synthetic materials are produced by incorporating small molecules or polymers into natural constituents; specifically, derivatives of cellulose and starch are often altered to improve skin compatibility and product efficacy (Aziz, Farid, & Ullah, 2022). These methodologies facilitate the creation of cosmetics that are sustainable in terms of both environmental impact and functionality.

Starch is a natural polysaccharide derived from plant sources like corn, rice, wheat, and barley, and is recognized as a versatile cosmetic component (Apriyanto, Compart, & Fettke, 2022). It is extensively utilized in skincare, haircare, and personal care products. Starch beads are garnering considerable attention in cosmetic formulations owing to their biodegradability, derivation from renewable sources, and capacity to substitute non-biodegradable microplastics. Additionally, starch beads serve as texture enhancers in cosmetic formulations, providing smoothness, effortless spreadability, and a sumptuous sensation. These characteristics render starch a more valued, eco-friendly, and efficient alternative in the cosmetics sector (Giustra et al., 2024).

Cellulose is a natural polymer present in plant cell walls and is derived from plentiful sources including wood, hemp, cotton, and flax. Cellulose-derived macro- and nanofibres exhibit biodegradable, non-toxic, and eco-friendly characteristics (Aziz et al., 2022). Due of these attributes, they provide a mild

exfoliating action without causing environmental harm. Cellulose has multiple functions in cosmetic formulations, including as an emulsifier, film-forming agent, and thickening agent. It is especially favored in facial peels and washing formulations to provide the requisite viscosity and textural characteristics (Mitura et al., 2020). Cellulose is regarded as a significant ingredient in cosmetic products because of its natural structure and eco-friendly attributes (Morganti, et al., 2023).

Chitosan is a natural polysaccharide present in the exoskeletons of crustaceans, including shrimp, crab, and lobster, as well as in the cell walls of fungi. Chitosan has emerged as a significant ingredient in cosmetic formulations for skincare and haircare products. Chitosan serves as a thickening agent, emulsifier, film-forming agent, and moisturizing ingredient in cosmetic goods (Aranaz et al., 2018). Research by Wisuitiprot and associates has shown that chitosan enhances skin elasticity, brightens skin tone, and diminishes wrinkles (Wisuitiprot et al., 2022). These findings underscore chitosan's potential as a novel and efficacious component in cosmetic applications.

2.2.2. Synthetic Alternatives

Synthetic polymers sourced from renewable materials, like polyhydroxyalkanoates and polylactic acid, are garnering attention as substitutes for petrochemical-based plastics because of their biodegradability and non-toxicity (Naser, Deiab, & Darras, 2021).

Polyhydroxyalkanoates (PHAs) are biopolymers made of hydroxyalkanoate (HA) polyesters, spontaneously synthesized by diverse Gram-positive and Gram-negative bacteria (Giustra et al., 2024). The biodegradable and compostable characteristics of PHAs render them an ecologically sustainable choice, facilitating their application in various goods, including cosmetics. PHAs are utilized in cosmetic products including beauty masks and sanitary pads, in addition to diverse applications including heat-sensitive adhesives, smart gels, and surfactants (Giustra et al., 2024). These multifaceted characteristics render PHAs distinguished as both a sustainable and functional material.

Poly(lactic acid) (PLA) is a polyester derived from renewable resources, including corn, rice, and sugarcane. It is regarded as an eco-friendly substance because of its biodegradable composition. It is extensively utilized in cosmetic packaging because to its mechanical strength and stiffness. Moreover, owing to its biocompatible composition, it is deemed an appropriate ingredient for skin, hair, and nail applications in diverse cosmetic formulations, including makeup products, soaps, creams, gels, and lotions (Abu Hajleh et al., 2020).

Poly(butylene succinate) (PBS) is a biodegradable polyester derived from renewable resources. PBS is regarded as an eco-friendly polymer due to its degradation into non-toxic byproducts, including water, carbon dioxide, and biomass (Shen, 2023). It possesses extensive applications owing to its mechanical strength, hardness, chemical resistance, and great compatibility with various additives.

2.3. Legal Regulations

Due to heightened awareness of the detrimental impacts of microbeads on the environment and human health, numerous nations have implemented legal prohibitions on their usage in cosmetic goods. The passage of microbeads through wastewater treatment systems into marine environments, coupled with their persistent non-biodegradable nature, has prompted action from regulatory authorities (UNEP, 2015). In 2015, the United States initiated legal measures with the Microbead-Free Waters Act, prohibiting the incorporation of microbeads in cosmetic and personal care products (U.S. Congress, 2015). Consequently, numerous nations, such as Canada, the United Kingdom, France, Sweden, and South Korea, have implemented such prohibitions (ECHA, 2022). In 2023, the European Union implemented additional limitations under the REACH (Registration, Evaluation, Authorisation and Restriction of Chemicals) laws, aimed at phasing out items containing microbeads. These findings signify a prevailing policy trend advocating for the global prohibition of microbeads and the endorsement of biodegradable alternatives.

In 2016, the United Nations Environment Assembly (UNEA) recognized the detrimental impacts of microplastics and urged nations to implement immediate measures. In later years, several documents advocating ways to mitigate MP were released, including the UN plan (2018) and the European Green Deal (2019). In 2021, Greenpeace released a report analyzing 664 cosmetic goods, revealing that 25% included microplastics and presenting a list of harmful compounds along with their respective quantities in each product. Various NGOs are overseeing and addressing marine debris, including the Honolulu Strategy. In 2019, the European Chemicals Agency (ECHA) recommended defining microplastics as 'particles composed of solid polymers' and advocated for their restriction. Countries have initiated the implementation of new rules and prohibitions against the use of microplastics in commercial products, especially cosmetics (Giustra et al., 2024).

2.3.1. National Legislation and Current Situation in Turkey

The criteria employed in Turkey to evaluate water quality and mitigate water pollution are fundamentally categorized into two primary groups. The initial category comprises discharge standards, which have been implemented for an extended period and are predominantly utilized in practice. The alternative category is designated as "environmental quality standards," which seek to assess the overall quality of water based on the amounts of pollutants determined in samples collected from aquatic bodies. Discharge standards are designed to hold the facility or activity responsible for releasing pollutants into aquatic environments and to impose sanctions when pollutant levels surpass established limits; environmental quality standards concentrate on assessing the current condition of water bodies and safeguarding ecosystems (Doğruel et al., 2022).

While Turkey lacks specific law on micro-pollutants, rules delineate requirements for wastewater discharge into receiving areas. In this regard, discharge criteria are principally delineated within the parameters of the Water Pollution Control Regulation (WPCR) and the Urban Wastewater Treatment Regulation. These regulations primarily target the management of traditional water pollution metrics, including Biochemical Oxygen Demand (BOD), Chemical Oxygen Demand (COD), suspended solids, nitrogen, and phosphorus (Republic of Turkey Ministry of Environment and Urbanization, 2004). Nevertheless, the industrial wastewater standards outlined in the WPCR establish limit values for specific heavy metals regarded as micro-pollutants, including Pb, Cd, Cr, and Hg, in addition to organic constituents such as hydrocarbons and phenolic compounds. Nonetheless, existing national legislation does not have a defined regulatory method or threshold value for significant types of micro-pollutants, including pharmaceutical residues, hormones, pesticides, microplastics, and nanomaterials (OECD, 2019; EEA, 2021). This scenario highlights a substantial necessity for legislation for the incorporation of micro-pollutants into treatment and monitoring protocols in our nation (Doğruel et al., 2022).

Quality criteria for the safeguarding of water resources and pollution control in European Union legislation are principally established by the Water Framework Directive (2000/60/EC) and the Environmental Quality Standards Directive (2008/105/EC). The Water Framework Directive seeks to define the requisite quality requirements for priority substances and specific contaminants in the receiving environment. The European Commission identified particular pollutants and priority compounds for monitoring in the aquatic environment in its Monitoring List (2015/495/EU) released in 2015. In Turkey, to align with EU

legislation, the Surface Water Quality Regulation and the Regulation on the Monitoring of Surface Waters and Groundwaters were developed, thereby integrating the water quality classification and monitoring principles of the Water Framework Directive into national law. These legislation have facilitated the creation of a water quality management framework in our nation that aligns with EU requirements (Üstün Odabaşı et al., 2020).

3. CONCLUSION

Micropollutants are substantial contaminants that pose severe challenges to environmental and ecological systems, even at minimal quantities. Given the advantages that micro-pollutants offer in industrial and consumer products, along with their diverse applications, complete eradication does not seem attainable in the near term. The consequences of pollution resulting from unregulated processes are significant and cannot be overlooked (Doğruel et al., 2022). Micro-pollutants in personal care and cosmetic goods are becoming a significant concern for both environmental and public health. Microplastics, synthetic polymers, preservatives, UV filters, and residues of drugs and hormones enter aquatic environments mainly unmitigated by wastewater treatment facilities, resulting in enduring impacts on ecosystems. The bioaccumulation potential of these contaminants suggests they may indirectly impact human health via the food chain.

In recent years, substantial progress has been made in microbead restrictions, the implementation of sustainable production methods, and the creation of biodegradable alternative materials. Nonetheless, these metrics do not encompass the complete spectrum of micro-pollutants. Consequently, laws should be augmented to provide explicit limit values for compounds designated as micro-pollutants, and monitoring and reporting mechanisms must be enhanced. Moreover, enhancing producer accountability, advocating for eco-labeled products, and propagating sophisticated treatment methods (e.g., membrane filtration, advanced oxidation, adsorption, etc.) are critically significant. In the long run, the extensive implementation of eco-friendly production and consumption practices is regarded as essential for creating a sustainable cosmetics sector and safeguarding aquatic ecosystems (Guerranti et al., 2019).

REFERENCES

- Abu Hajleh, M. N., Al-Samydai, A., & Al-Dujaili, E. A. S. (2020). Nano, micro particulate and cosmetic delivery systems of polylactic acid: A mini review. *Journal of Cosmetic Dermatology*, 19(11), 2805-2811.
- Apriyanto, A., Compart, J., & Fettke, J. (2022). A review of starch, a unique biopolymer — Structure, metabolism and in planta modifications. *Plant Science*, 318, 111223.
- Aranaz, I., Acosta, N., Civera, C., Elorza, B., Mingo, J., Castro, C., Gandía, M. D. I. L., & Heras Caballero, A. (2018). Cosmetics and cosmeceutical applications of chitin, chitosan and their derivatives. *Polymers*, 10(2), 213.
- Aziz, T., Farid, M. A., & Ullah, S. (2022). A review on the modification of cellulose and its derivatives: Applications and perspectives. *Polymers*, 14(13), 2649.
- Becker, L. C., Bergfeld, W. F., Belsito, D. V., Hill, R. A., Klaassen, C. D., Liebler, D. C., Marks, J. G., Shank, R. C., Slaga, T. J., Snyder, P. W., & Andersen, F. A. (2011). Final Report of the Cosmetic Ingredient Review Expert Panel safety assessment of polymethyl methacrylate (PMMA), methyl methacrylate crosspolymer, and methyl methacrylate/glycol dimethacrylate crosspolymer. *International Journal of Toxicology*, 30, 54S-65S.
- Becker, L. C., Bergfeld, W. F., Belsito, D. V., Hill, R. A., Klaassen, C. D., Liebler, D. C., Marks, J. G., Shank, R. C., Slaga, T. J., Snyder, P. W., & Andersen, F. A. (2014). Safety assessment of modified terephthalate polymers as used in cosmetics. *International Journal of Toxicology*, 33, 36S-47S.
- Bhatt, P., G. Bhandari, and M. Bilal, (2022). “Occurrence, toxicity impacts and mitigation of emerging micropollutants in the aquatic environments: Recent tendencies and perspectives”, *J. Environ. Chem. Eng.*, 10, 107598,
- Bhattacharya, P., (2016). “A Review On The Impacts Of Microplastic Beads Used In Cosmetics”, *Acta Biomedica Scientia*, 3(1):47-52.
- Doğruel, S., Gürel, M., Mantaş, P., E., (2022), “Mikrokirleticiler: Tanım, Mevzuat ve Ülkemizde Atıksularda ve Yerüstü Sularında Mevcudiyetleri”, *Çevre, İklim ve Sürdürülebilirlik*, 23(2), 133–144.
- European Chemicals Agency (ECHA). (2022). Annex XV Restriction Report: Proposal for a restriction on intentionally added microplastics. Helsinki: ECHA.
- European Commission. (2023). Commission Regulation (EU) 2023/2055— Restriction of intentionally added microplastics under REACH. Brussels.
- European Environment Agency (EEA). (2021). Emerging chemical risks in Europe — ‘PFAS’ and other substances of concern. EEA Report No 22/2021, Copenhagen.
- Giustra, M., Sinesi, G., Spena, F., De Santes, B., Morelli, L., Barbieri, L., Garbujo, S., Galli, P., Prosperi, D., Colombo, M., (2024) “Microplastics in Cosmetics: Open Questions and Sustainable Opportunities”, *ChemSusChem*, 1-18.

- Guerranti, C., Martellini, T., Scopetani, C., Cincinelli, A., (2019). "Microplastics in cosmetics: Environmental issues and needs for global bans", *Environmental Toxicology and Pharmacology* 68, 75-79.
- Guzik, M., Czerwińska-Ledwig, O., & Piotrowska, A. (2023). Compositions of abrasive cosmetics from Polish manufacturers. *Cosmetics*, 10(2), 67.
- Leslie HA. (2014). "Review of Microplastics in Cosmetics". Institute of Environmental Studies, VU University of Amsterdam, the Netherlands
- M. Rahim, N., A., S., Islahudin, F., Tahrir, N., A., Jasamai, M., (2022). "Microplastics in Cosmetics and Personal Care Products: Impacts on Aquatic Life and Rodents with Potential Alternatives", *Sains Malaysiana* 51(8), 2495-2506.
- Marcharla, E., Vinayagam, S., Gnanasekaran, L., Soto-Moscoso, M., Chen, W. H., Thanigaivel, S., & Ganesan, S. (2024). Microplastics in marine ecosystems: A comprehensive review of biological and ecological implications and its mitigation approach using nanotechnology for the sustainable environment. *Environmental Research*, 256, Article 119181.
- Miller, M. E., Hamann, M., & Kroon, F. J. (2020). Bioaccumulation and biomagnification of microplastics in marine organisms: A review and meta-analysis of current data. *PLOS ONE*, 15(10).
- Mitura, S., Sionkowska, A., Jaiswal, A. K., & Griveau, S. (2020). Biopolymers for hydrogels in cosmetics: A review. *Polymers*, 12(1), 204.
- Morganti, P., Morganti, G., & Coltelli, M.-B. (2023). Natural polymers and cosmeceuticals for a healthy and circular life: The examples of chitin, chitosan, and lignin. *Cosmetics*, 10(2), 42.
- Naser, A.Z., Deiab, I. & Darras, B.M. (2021). Poly(lactic acid) (PLA) and polyhydroxyalkanoates (PHAs), green alternatives to petroleum-based plastics: A review. *RSC Advances* 11(28): 17151-17196.
- Organisation for Economic Co-operation and Development (OECD). (2019). *Pharmaceutical Residues in Freshwater: Hazards and Policy Responses*. OECD Publishing, Paris.
- Özçelik H., Bebekli Ö., (2015). "Kozmetik Sektörüne Genel Bakış", *Anamas Dergisi*, 4, 3-12.
- Perschbacher, E. (2016). History and evolution of the microbeads. <https://www.ijc.org/en/history-and-evolution-microbead>. Accessed on 8 June 2021.
- Prata, J. C., da Costa, J. P., Lopes, I., Duarte, A. C., & Rocha-Santos, T. (2020). Environmental exposure to microplastics: An overview on possible human health effects. *Science of the Total Environment*, 702, Article 134455.
- Ranjan, V. P., Joseph, A., Srivastava, S., Bhakta Sharma, H., Biswas, B., Goel, S., & Kumar, S. (2024). From cosmetics to contamination: Microplastics in personal care products as vectors for chromium in aquatic environments. *Waste Management Bulletin*, 2(3), 229–240.

- Üstün Odabaşı, S., Altın, S. H., Büyükgüngör, H., (2020). “Sucul Ortamdaki Bazı Mikrokirleticilerin Oluşumu, Durumu ve İleri Oksidasyon Prosesleri ile Giderilmesi”, NÖHÜ Müh. Bilim. Derg., 9(1): 57-71.
- Shen, K. (2023). Synthesis, properties and applications of poly(butylene succinate). MATEC Web of Conferences, 386, Article 01005.
- Republic of Turkey Ministry of Environment and Urbanization. (2004). Water Pollution Control Regulation. Official Gazette, December 31, 2004, No. 25687.
- U.S. Congress. (2015). Microbead-Free Waters Act of 2015 (Public Law No: 114–114). Washington, D.C.: Government Publishing Office.
- United Nations Environment Programme (UNEP). (2015). Plastic in cosmetics: Are we polluting the environment through our personal care, UNEP Publication, Nairobi.
- Vo, H. C., & Pham, M. H. (2021). Ecotoxicological effects of microplastics on aquatic organisms: A review. Environmental Science and Pollution Research, 28(33), 44716–44725.
- W. Wisutitprot, K. Ingkaninan, S. Jones, N. Waranuch, J. (2022). “Effect of green tea extract loaded chitosan microparticles on facial skin: A split-face, double-blind, randomized placebo-controlled study” Cosmet. Dermatol. 21, 4001–4008.
- Zhang, X., Yu, C., Wang, P., & Yang, C. (2025). Microplastics and human health: unraveling the toxicological pathways and implications for public health. Frontiers in Public Health, 13, 1567200.
- Zhou, Y., Ashokkumar, V., Amobonye, A., Bhattacharjee, G., Sirohi, R., Singh, V., Flora, G., Kumar, V., Pillai, S., Zhang, Z., Awasthi, K., M., (2023). “Current research trends on cosmetic microplastic pollution and its impacts on the ecosystem: A review”, Environmental Pollution 320, 121106.
- Zitko, V., Hanlon, M., (1991). Another source of pollution by plastics: skin cleaners with plastic scrubbers. Mar. Pollut. Bull. 22, 41–42.

CHAPTER 4

Türkiye's Mobile Communication Technology Journey: An Anatomy of Digital Transformation from 3G to 5G

Ali Özhan Akyüz¹ & Onur İnan²

¹ Prof. Dr., Burdur Mehmet Akif Ersoy University, Bucak Emin Gulmez Vocational School of Technical Sciences, 0000-0001-9265-7293

² Dr. Öğr. Üyesi, Burdur Mehmet Akif Ersoy University, Bucak Emin Gulmez Vocational School of Technical Sciences, 0000-0002-9683-344X

1. INTRODUCTION

Türkiye's digital transformation journey is embodied in the continuous and rapidly evolving development of its mobile communication networks. Mobile technology, which acts as a driving force in numerous areas—from a country's economic growth and the dynamics of social life to education and the delivery of public services—is not only a communication tool for Türkiye but also a fundamental indicator of its global competitiveness. The prevalence of mobile broadband access, the volume of data usage per capita, and network speeds are critical milestones in Türkiye's goal of achieving an information society (Can, 2021). The story of mobile technology in Türkiye began in the early 1990s with Analog (1G) and then quickly progressed to Digital (2G) networks. The Second Generation (2G), also commonly known as GSM (Global System for Mobile Communications), clearly defined the core purpose of mobile communication: to offer reliable, encrypted, and portable voice communication, along with Short Message Service (SMS). This technology revolutionized social and commercial life by enabling users to speak regardless of their geographical location (Toğa and Toğa, 2019; Akkurt, 2021; Özakta and Ardiyok, 2024). However, the GSM network, by its nature, was not designed for high-speed data transfer. Data speeds, even with the most advanced 2G standards (GPRS and EDGE), were only a few hundred kilobits per second. These speeds were inadequate even for opening a simple website, let alone using mobile internet or rich multimedia content. Mobile devices during this era were fundamentally just phones (Halonen et al., 2004; Saily et al., 2011). It is at this exact point that Third Generation (3G) technology ignited a revolution that fundamentally changed the mobile communication paradigm (Mohr and Konhauser, 2002). Licensed and launched in Türkiye in 2009, 3G (UMTS/HSPA) entered the history of mobile communication not just as a speed increase, but as the birth of the mobile internet. The core philosophy of 3G was to effectively merge voice and data traffic, providing users with broadband speeds at any time and from anywhere (Yavuz et al., 2009). This new network opened up a new world for mobile devices. It Broke the Speed Barrier: Speeds starting theoretically at 2 Mbps and reaching up to 42 Mbps with HSPA+ unlocked a new world for mobile devices, making it possible to download email attachments, use map applications, and even watch live video on a mobile device. It Launched the Smartphone Era: The higher speed and lower latency brought by 3G allowed smartphones, such as Apple's iPhone and Google's Android operating system, to showcase their full potential, transforming smartphones from mere talking devices into personal computers that organized daily life. Finally, it Gave Rise to the Application Economy: Applications like mobile banking, real-time tracking of social media feeds (Facebook, Twitter), instant messaging (WhatsApp), and mobile games gained mass adoption thanks

to the bandwidth provided by 3G networks (Steinbock, 2005; Rao et al., 2009; Goggin, 2021).

2. 3G ERA: THE BIRTH OF MOBILE INTERNET

2.1. Transition Process to 3G and Licensing

The transition to 3G technology, one of the most significant leaps in mobile communication in Türkiye, was formalized with a tender process in 2008. The tender was conducted by the then Information and Communication Technologies Authority (BTK), and mobile network operators (then named Turkcell, Vodafone Türkiye, and Avea) were granted licenses to use the UMTS (Universal Mobile Telecommunications System) technology. This process established the legal basis for offering mobile broadband services in a competitive environment in Türkiye. Following the completion of the licensing processes, mobile operators commercially launched 3G services in July 2009; this date marks a turning point, when mobile devices in Türkiye transitioned from being solely a voice communication medium to a data-centric communication platform. Operators were obligated to achieve specific coverage area targets as per the license conditions, which rapidly led to the expansion of the 3G network not only in major cities but also in smaller settlements (Özaktaş and Ardiyok, 2004; Evci et al., 2004; Alderfer et al., 2009; Ayan and Şenol, 2016; Can, 2021).

2.2. Technological Features: HSPA and HSPA+

3G technology is technically based on UMTS standards; however, the sub-technologies that truly provided widespread adoption and high speeds to users in Türkiye and globally were HSPA and HSPA+. HSPA (High-Speed Packet Access) introduced a series of enhancements that significantly increased the data speed of the UMTS network: HSDPA (Downlink) offered the potential to increase download speeds from 1.8 Mbps up to 14.4 Mbps, radically changing the user experience for web browsing and file downloading; while HSUPA (Uplink) increased upload speeds, paving the way for two-way applications such as sending emails, uploading photos, and video calling from mobile devices. HSPA+ (Evolved HSPA) became the final and most advanced evolution of 3G, enabling theoretical download speeds starting from 21 Mbps up to 42 Mbps through enhancements like data compression and multiple antenna (MIMO) technologies, and serving as the last fast mobile internet solution before the transition to 4G. In terms of Services Offered, thanks to the low latency and increased bandwidth brought by 3G, video calling, mobile television (a widespread service at the time), fast mobile email access, and real-time use of map/navigation applications became standard services (Holma and Toskala, 2010; Hytönen, 2015; Çemen, 2023).

2.3. Impact on the Market: The Widespread Adoption of Mobile Internet Usage

The impact of 3G on the market was transformative, not only for operators but for the entire technology ecosystem. Under the "Smartphone Revolution" heading, 3G provided the technological infrastructure for the rapid proliferation of Apple iPhones and Android devices in the Turkish market. Mobile devices have become a central part of internet access, creating a user expectation of staying constantly online. There was an Explosion in Data Consumption; monthly mobile data usage, which was symbolic during the 2G era, saw a geometric increase with the advent of 3G. Users began following social media, instantly sharing photos and watching short videos on their mobile devices. This shift increased mobile operators' primary revenue source from voice to data. Finally, New Business Models emerged; application developers began offering new mobile services and content, leveraging the high speeds, and e-commerce sites in Türkiye rapidly expanded their user bases due to the ease of mobile Access (Toğa and Toğa, 2019; Evci et al., 2004; Chiplunkar and Goldberg, 2022).

2.4. User Experience: Early Mobile Applications and Web Browsing

3G significantly enriched the daily user experience compared to the 2G era. In Instant Communication, messaging, previously limited to text-based interactions, can now include rich media such as photos, voice recordings, and location sharing. The rise of applications like WhatsApp in Türkiye is directly indebted to the reliable data transmission provided by 3G networks. Regarding Mobile Email and Work, critical email access for the business world became nearly as fast and seamless as a desktop experience with 3G, providing great convenience for mobile workers. Simple Web Browsing was accelerated; loading a web page, which used to take minutes on 2G, was now completed in seconds, establishing the habit of mobile search and simple news reading. Finally, Location-Based Services emerged; the combination of GPS and 3G data made mobile map and navigation applications (such as Google Maps) indispensable parts of urban life (Acar, 2022; Tözer, 2023).

In conclusion, the 3G era transformed Türkiye into a consumer of mobile internet, laying the groundwork for the next stages: 4.5G and 5G.

2.5. The 3G Era in Türkiye: Timeline and Key Transformations

Table 1 summarizes the milestones in Türkiye's transition to 3G and the fundamental changes brought by this technology (Özaktaş and Ardiyok, 2004; Tözer, 2023; Göken et al., 2022)

3. 4.5G (LTE-A): INCREASED SP

EED AND CAPACITY

3.1. Transition to 4.5G: Adopting a Strategic Model

When transitioning to fourth-generation mobile communication technology (4G) in 2016, Türkiye adopted the LTE-Advanced (LTE-A) standard, which it referred to as 4.5G, rather than the basic 4G (LTE) widely adopted globally. There were strategic reasons behind this decision: A Greater Need for Speed and Capacity was paramount, as the speeds offered by 3G had begun to fall short against the rapidly increasing demand for mobile data (especially video and social media usage); since LTE-Advanced offered much higher speeds and capacity than basic 4G, the aim was to meet the intense data needs of the Turkish market directly. The choice of the name 4.5G served a Marketing and Perception goal, strengthening the perception that the launched technology was superior and more future-oriented than standard 4G, thereby providing a significant competitive advantage for mobile operators. Finally, Technological Preparedness played a role, as operators had already begun installing network equipment—which forms the basis of LTE-A technology—during the 4G tender, granting them the flexibility to launch the most advanced technology immediately. 4.5G services in Türkiye were commercially launched as of April 1, 2016 (AA, 2025; UAB; 2016; Çalış and Özdemir, 2013; Eryılmaz, 2022; Mustari et al., 2023)

3.2. Technological Differences: Carrier Aggregation and LTE-A

The most significant technological innovations distinguishing 4.5G (LTE-Advanced) from basic 4G are as follows. Carrier Aggregation (CA) is LTE-A's most critical feature, enabling mobile operators to combine separate fragments of spectrum across different frequency bands (e.g., 800 MHz, 1800 MHz, and 2600 MHz) to form a single, wider virtual channel. Its Significance: Much like dividing vehicle traffic into multiple lanes, this technology allows data to be transmitted over a wider path, greatly increasing the theoretical speed. In Türkiye, this enabled speeds theoretically ranging from 100 Mbps to 300 Mbps, with gigabit speeds being approached in later stages. Advanced MIMO (Multiple-Input Multiple-Output) involves deploying more advanced versions (e.g., 4x4 MIMO) of the technology, which simultaneously sends data to both the receiver and the transmitter using multiple antennas. This increased spectral efficiency enables faster service to more users. Finally, Higher Order Modulation (256 QAM) enhances the quality of the data signal, allowing more data packets to be transmitted within the same time slot, which directly boosts speeds (Akkurt, 2021; Çalış and Özdemir, 2013; Halonen et al., 2004; Mustari et al., 2023; Geylani et al., 2016; Bouras et al., 2020; Jiang and Han, 2024).

Table 1. Summary of 3G Technology's Rollout and Consequences in Türkiye

Category	Feature / Milestone	Description and Impact
Timeline	Licensing Tender	Conducted by the BTK in 2008.
	Service Launch	July 2009. Türkiye officially entered the mobile broadband era.
Technological Focus	Core Standard	UMTS (Universal Mobile Telecommunications System).
	Speed Technology	HSPA (High-Speed Packet Access) and HSPA+.
	Theoretical Speeds	Up to 14.4 Mbps with HSPA, and up to 42 Mbps with HSPA+.
Difference from Previous Generation	2G (GSM)	Focus: Voice and SMS. Data Speed: Low (≈ 200 kbps).
	3G (UMTS)	Focus: Voice + Data. Data Speed: High. Video calling became possible with low latency.
Market Impact	Device Transformation	The infrastructure was prepared for the mass proliferation of smartphones (iPhone, Android).
	Usage Trend	Geometric increase in mobile data consumption. Operator revenue focus shifted from voice to data.
User Experience	New Applications	Mobile usage of social and instant messaging applications, such as WhatsApp, Facebook, and Twitter, has exploded.
	Services	Mobile map/navigation applications, mobile banking, and simple web browsing became standard.

3.3. Licensing and Frequency Distribution

In 4.5G licensing, significantly wider and more diverse frequency bands were allocated than in 3G, which is crucial for enhancing network performance and coverage (Table 2) (Toğa and Toğa, 2019; Çalış and Özdemir, 2013; Ceylan et al., 2021).

Table 2. Türkiye's 4.5G (LTE-A) Frequency Bands and Their Impact

Frequency Band	Characteristic	Impact
800 MHz	Low Frequency	Provides wide coverage in rural areas and inside buildings. Covers a large geography with fewer base stations.
1800 MHz	Medium Frequency	A balancing band. Offers both good coverage and high capacity.
2100 MHz	Medium Frequency	A commonly used band inherited from 3G and also utilized for 4.5G.
2600 MHz	High Frequency	Provides high capacity and speed for city centers and crowded areas. Requires dense base stations due to shorter range.

The combined use of different bands (Carrier Aggregation) formed the basis for Turkish operators to be able to offer both wide coverage in rural areas and high speeds in cities.

3.4. The Impacts of 4.5G: Video, Content, and Gaming

4.5G radically changed mobile internet usage habits: The increased speeds enabled HD (High Definition) and even 4K video streaming on mobile devices, leading to an astronomical rise in viewing rates for services like YouTube, Netflix, and local series/movie platforms on mobile devices, causing an explosion in video consumption. High-resolution content sharing has become standard on social media, with the instant uploading of HD photos and videos becoming increasingly common, thereby enhancing the quality of user-generated content. Low latency and high speed allowed for the proliferation of Online Multiplayer (MMO) mobile games and competitive e-sports on mobile platforms, as fast and seamless connections are vital for players, thus boosting Mobile Gaming. Finally, uploading and downloading large files from mobile devices to cloud storage services (such as Dropbox and Google Drive) became practical, leading to the Widespread Adoption of Cloud Services (Shimray et al.,2015; Saylor and Saylor, 2025).

3.5. Infrastructure Development: Fiber Backbone and Densification

The high speeds and capacity offered by 4.5G necessitated strengthening the mobile network's "backbone": the Fiber Optic Backbone was critical as data traffic to the base stations (backhaul) increased significantly. To handle this traffic, a large portion of the base stations needed to be connected to the main network centers using fiber-optic cables. The speed of fiberization became critically important in Türkiye with the introduction of 4.5G. Additionally, due

to the short range of high-frequency bands (especially 2600 MHz), operators were forced to install a higher Base Station Density in city centers. This densification also paved the way for the transition to the small cell architecture required for the shift to 5G (Tözer, 2023; Ceylan et al., 2021; Routray and Sharmila, 2016).

In short, the 4.5G era moved Türkiye to the upper ranks of the global mobile broadband league and raised user expectations to a peak for 5G.

3.6. 4.5G (LTE-A) Transformation in Türkiye: Strategy, Technology, and Impact

Table 3 summarizes why Türkiye directly adopted 4.5G and the core innovations this technology brought to network architecture and user experience (Mustari et al., 2023; Ardiyok, 2024; Demircioğlu et al.,2024).

Table 3. Anatomy of Türkiye's Mobile 4.5G Transition

Category	Feature Innovation /	Description and Impact
Transition Strategy	Why 4.5G?	Adoption of the more advanced and higher-capacity LTE-Advanced model instead of the standard 4G (LTE). Focused on marketing and the need for higher speeds.
	Service Launch	April 1, 2016. Provided a revolutionary increase in mobile internet speed in Türkiye.
Technological Core	Carrier Aggregation (CA)	The most critical innovation. It combined signals from different frequency bands (800, 1800, and 2600 MHz) into a single virtual channel, increasing speed compared to 3G.
	MIMO and Modulation	Advanced antenna technologies and higher-order modulation, such as 256-QAM, increased spectral efficiency and capacity.
Frequency Distribution	800 MHz	Coverage band. Used for penetration inside buildings and in rural areas.
	2600 MHz	Capacity band. Met the high-speed and capacity needs in dense city centers.
Market Impact	Data Usage	Mobile data consumption has increased exponentially due to HD/4K video streaming and the sharing of high-resolution content.
	Application Area	Provided low latency and high speed for Mobile Gaming and high-speed cloud services.
Infrastructure Requirement	Fiber Backbone	The necessity for fiber optic (backhaul) connections became prominent to support the increased data traffic at base stations.

Category	Feature Innovation	/	Description and Impact
	Base Station Densification		The number of base stations increased in city centers to boost capacity and compensate for the short range of high frequencies.

4. 5G: THE NETWORK OF THE FUTURE

5G is built upon a Software-Defined Network (SDN) concept, completely changing not only the speed but also the fundamental architecture of the network. This positions 5G as an infrastructure platform, not just a generation of technology (Lehr et al., 2021).

4.1. The Critical Importance of 5G: The Triple Service Pillars

The main driving force behind the design of 5G is the need to respond to vertical industrial demands where 4G's single focus (speed) proved insufficient (Table 4) (Bega et al., 2017; Xu et al., 2021).

Table 4. Core 5G Service Categories (eMBB, mMTC, uRLLC)

Scenario	Technical Focus	Potential Application Areas
eMBB (enhanced Mobile Broadband)	Capacity/Speed (10 Gbps Peak Speed)	8K video streaming, Holoportation (live 3D hologram), and real-time cloud rendering.
mMTC (massive Machine Type Communications)	Connection Density (10 ⁶ devices/km ²)	Smart homes, remote reading of city water/energy meters, container and pallet tracking in logistics.
uRLLC (ultra-Reliable Low Latency Communications)	Latency (1 ms) and Reliability (99.999%)	Industrial Automation (PLC-Robot) control, precision agriculture, and control of remote surgery robots.

4.2. Breakthrough Technological Architectural Innovations

The core technological architecture enabling 5G to deliver on its promises includes:

A. Spectrum Usage and Signal Processing

5G utilizes a much broader spectrum than previous generations and introduces breakthroughs in signal processing methods. Through Expanded Spectrum Usage, 5G focuses on two main frequency bands: the Sub-6 GHz (Mid-Band), particularly the 3.5 GHz (C-band), which is the primary breakout band for 5G, offering better speeds than 4G while providing similar coverage, and is where

initial commercial 5G deployments worldwide and in Türkiye are concentrated. On the other hand, mmWave (High-Band), operating at 24 GHz and above, can offer a ~ 100 times wider bandwidth, but its range is limited to a few hundred meters and it is highly sensitive to physical obstacles; this challenge is overcome using Ray Tracing algorithms and Beamforming technology. This technology underpins 5G's "smart antenna" feature, known as Massive MIMO (Massive Multiple-Input Multiple-Output). Massive MIMO utilizes a large number of antennas to direct energy toward the targeted user with pinpoint accuracy (Beamforming). This method not only enhances signal quality but also increases network capacity by allowing the same radio frequency resource to serve multiple users in different directions simultaneously (Spatial Multiplexing) (Alexandropoulos et al., 2022; Abd et al., 2022; Ibrahim et al., 2023; Sathishkumar et al., 2025).

B. Network Virtualization and Software-Defined Networks (SDN/NFV)

The foundation of 5G's promised flexibility and its ability to serve diverse vertical industries is rooted in Network Slicing and Autonomy. This feature ensures that each virtual slice of the network possesses its own dedicated resources (processing power, bandwidth, security protocols); for instance, one slice can be statically allocated for industrial control requiring 1 ms latency, while another slice can be used for best-effort consumer internet. This adaptability allows 5G to be tailored to diverse business models. Furthermore, the Software-Defined Core Network (5G Core - 5GC) establishes its core network using a cloud-based architecture, relying on virtual machines or containerization. This architectural flexibility grants operators a Time-to-Market advantage, allowing them to deploy new services (such as edge computing and IoT platforms) much faster (Ahokangas et al., 2019; Aranda et al., 2021; Banda et al., 2022).

4.3. Türkiye's 5G Vision: Local Capacity and Regulation

Türkiye is planning its 5G transition by focusing on the concerns highlighted during the 4.5G tender, namely domestic production and cybersecurity. Regarding Regulatory and Licensing Challenges, the timing and structure of the 5G tender are largely tied to frequency allocation and technological readiness. Main regulatory topics include the allocation of the 3.5 GHz band and making supplementary bands, such as White Spaces (TV spectrum), available for 5G use. The vision for Deepening Domestic and National Technology (ULAK and Beyond) is evidenced by the ULAK base station project aimed at increasing the domestic content ratio in the Radio Access Network (RAN), and with 5G, this goal is expanded to encompass not only base station hardware but also Core Network software, management systems, and the AI/Machine Learning-based network optimization tools required by 5G, ultimately aiming to keep Türkiye's critical 5G infrastructure under national control. Furthermore, Edge Computing

is a key focus of Türkiye’s 5G preparations, where the 5G network brings servers and applications as close as possible to the base station (the edge) to minimize latency, which is the only way to meet the critical 1 ms latency requirements for Industry 4.0 and autonomous vehicles (Acar, 2022; Peterson and Sunay, 2020; Yıldırım and Tunalı, 2022; Ateş, 2024).

4.4. Expected Applications and Sectoral Economic Transformation

5G is expected to contribute trillions of dollars to the global economy over a decade. Turkish sectors will benefit from this transformation through several key areas, including Intelligent Transportation Systems (ITS), which leverages V2X (Vehicle-to-Everything) communication to enable instant data sharing between vehicles. This reduces accident risk and increases cities' energy efficiency by optimizing traffic flow through the central intelligence of the 5G network. Digital Factories (Industry 4.0) will see boosted efficiency and flexibility through the integration of wirelessly controlled robots, production lines connected via 5G, and field technicians receiving maintenance/repair instructions via Augmented Reality (AR) glasses. Finally, the Metaverse and Mobile AR/VR will be revolutionized, as 5G's high bandwidth and low latency enable virtual and augmented reality applications to function seamlessly on mobile devices, creating new interactive experiences in sectors such as education, real estate, and retail (Mustari et al., 2023; Ateş, 2024; Şevket et al.,2022; Uğur and Aksoy,2022).

In summary, 5G is a strategic national infrastructure project built upon the foundation laid by 4.5G in Türkiye, designed to prepare not only the telecommunications sector but also the national industry and service sectors for global competition.

4.5. 5G Technology: Three-Dimensional Platform and Türkiye Strategy

Table 5 summarizes the three main 5G usage scenarios, their technical components, and the application and preparation areas in Türkiye.

Table 5. 5G's Main Service Categories and Türkiye Application Areas

Category		Core Focus Point	Key Features / Technologies	Impact Areas in Türkiye
5G Usage Scenarios (The Triangle)		eMBB (enhanced Mobile Broadband)	Peak speeds up to 20 Gbps, high capacity.	4K/8K video streaming in crowded areas, VR/AR-based education and entertainment.
		mMTC (massive Machine Type Communications)	1 million devices/km ² density, low energy consumption.	Smart Cities, meter reading, smart agriculture, and sensor tracking in logistics.

Category	Core Focus Point	Key Features / Technologies	Impact Areas in Türkiye
Technological Innovations	uRLLC (ultra-Reliable Low Latency Communications)	1 ms latency, 99.999 % reliability.	Industrial automation (Robot Control), autonomous vehicles (V2X), remote surgery (Tele-Health).
	Spectrum Usage	Millimeter Wave (mmWave) (High-capacity band) and Sub-6 GHz (Balance of coverage and capacity).	Allocation of the 3.5 GHz band (Mid-Band) and mmWave tests.
	Network Architecture	Massive MIMO (Signal focusing with smart antennas) and Network Slicing (Dividing the network into virtual slices).	Special allocation of network resources for enterprise and critical applications.
Türkiye's Preparations	Domestic Capacity	Development of domestic base stations like ULAK and core network software.	Contribution to cybersecurity and technological independence goals.
	Regulation Focus	Frequency planning and tender preparations.	Allocation of spectrum compliant with international standards and the implementation of Edge Computing infrastructure.

5. CONCLUSION

Mobile communication technologies have been the most critical infrastructure component driving radical changes in Türkiye's socio-economic structure over the last two decades. The mobile internet revolution, which began with 3G, reached its peak with 4.5G, and now attention is focused on 5G technology, seen as the engine of national digital transformation.

Türkiye has demonstrated success in adopting and disseminating mobile communication technologies at a pace exceeding the global average. The Rapid Adaptation and Proliferation—transitioning to 3G in 2009 and 4.5G in 2016—ensured high nationwide penetration rates within a short period. Despite a high subscriber density, the 4.5G (LTE-Advanced) network has successfully delivered world-class data speeds to its users, driving mobile data consumption to record levels. Global Standard Network Performance has been demonstrated, particularly through the effective use of Carrier Aggregation (CA), where 4.5G

performance in major cities has surpassed the 4G average in many developed countries, creating a solid foundation for intensive video consumption, mobile gaming, and cloud-based services. Regarding the Development of Domestic Technology, local base station projects, such as ULAK, initiated during the 4.5G process, represent concrete steps toward Türkiye's strategic goal of independence in communications technology, thereby strengthening the domestic technology ecosystem for 5G and beyond. Additionally, the Competitive Market structure involving three major operators has resulted in both affordable services for users and continuous investment in network quality. Currently, Türkiye continues to utilize all capabilities of 4.5G technology. Still, with the acceleration of global 5G deployment, a rapid transition to 5G is an unavoidable necessity to maintain economic and industrial competitiveness.

Türkiye faces several significant challenges that It must carefully manage to achieve its 5G goals. A major cost factor is the Intensive Base Station Need. The high-frequency bands, such as mmWave and 3.5 GHz, that deliver 5G's promises have a short range, requiring the installation of exponentially more, smaller cell (Small Cell) base stations in cities than in existing 4G networks, resulting in very high CAPEX (Capital Expenditure) for operators. Furthermore, a Fiber Optic Backbone is essential; the enormous data traffic generated by 5G base stations cannot be managed with existing low-capacity connections. For 5G to realize its full potential, every base station must be connected to the core network via fiber-optic cable (fiber backhaul). The prevalence of fiber infrastructure in Türkiye, especially in the context of mobile needs, is a critical challenge that requires a rapid increase.

Frequency Band Diversity is a fundamental requirement for 5G, necessitating the use of low-, mid-, and high-band frequencies. The efficient management of available spectrum, particularly the allocation of the 3.5 GHz band and other bands such as White Spaces, is a critical regulatory task. Delays in tender processes risk postponing the expected economic benefits of 5G for Türkiye. Moreover, regulatory mechanisms for passive infrastructure sharing are needed to encourage the effective sharing of passive components (such as base stations and fiber) among operators, thereby reducing infrastructure investment costs.

The connection of a massive number of devices (IoT) and critical industrial processes via 5G creates New Attack Surfaces, exponentially increasing cyber risk. The national 5G network needs robust protection against international cyber threats. The Domestic Production Pressure is a requirement imposed to advance the goals of technological independence and cybersecurity. Mandating the use of domestic technology in the 5G core and access networks strengthens the local ecosystem, but also creates the urgency for local firms to deliver mature and scalable solutions promptly, competing with global players.

Türkiye must proactively manage these challenges to build on its past success in rapid adoption and transition to 5G. 5G will not only be the future of mobile communication but also a fundamental catalyst for Türkiye's industrial and economic leap.

REFERENCES

- AA; 4.5G teknolojisinin geleceği. (2025). <https://www.aa.com.tr/tr/ekonomi/turkiyenin-4-5g-ile-tanismasinin-uzerinden-9-yil-gecti/3525055>
- Abd, A. K., Rasool, J. M., Rahman, Z. A. S., & Al-Yasir, Y. I. (2022). Design and analysis of novel reconfigurable monopole antenna using DIP switch and covering 5G-sub-6-GHz and C-band applications. *Electronics*, 11(20), 3368.
- Acar, M. (2022). An assessment of localization policies in Turkey: The case of mobile telecommunication sector (Master's thesis, Middle East Technical University).
- Ahokangas, P., Matinmikko-Blue, M., Yrjölä, S., Seppänen, V., Hämmäinen, H., Jurva, R., & Latva-Aho, M. (2019). Business models for local 5G micro operators. *IEEE Transactions on Cognitive Communications and Networking*, 5(3), 730–740.
- Akkurt, R. (2021). Geçmişten günümüze kablosuz mobil iletişim teknolojileri. *Avrupa Bilim ve Teknoloji Dergisi*, (28), 115–119.
- Alderfer, H., Flynn, S., Birchmeier, B., & Schulz, E. (2009). Information policy country report: Turkey.
- Alexandropoulos, G. C., Islam, M. A., & Smida, B. (2022). Full-duplex massive multiple-input, multiple-output architectures: Recent advances, applications, and future directions. *IEEE Vehicular Technology Magazine*, 17(4), 83–91.
- Aranda, J., Sacoto Cabrera, E. J., Haro Mendoza, E. D., & Astudillo Salinas, F. (2021). 5G networks: A review from the perspectives of architecture, business models, cybersecurity, and research developments. *Novasinerгия*, 4.
- Ardıyok, F. D. Ö. S. (2024). Policies and legislation. In *The economics and regulation of digitalisation: The case of Türkiye*.
- Ateş, M. (2024). Enabling intelligent transport in smart cities: A 5G Valley in Türkiye. *Journal of Transportation and Logistics*, 9(2), 219–233.
- Ayan, M., & Şenol, R. (2016). Bulanık Mantık Tabanlı – Uzaktan Erişimli Sera Otomasyonu. *Duzce University Journal of Science and Technology*, 4(2), 734-746.
- Banda, L., Mzyece, M., & Mekuria, F. (2022). 5G business models for mobile network operators—A survey. *IEEE Access*, 10, 94851–94886.
- Bega, D., Gramaglia, M., Bernardos Cano, C. J., Banchs, A., & Costa-Perez, X. (2017). Toward the network of the future: From enabling technologies to 5G concepts. *Transactions on Emerging Telecommunications Technologies*, 28(8), e3205

- Bouras, C., Gkamas, A., Diles, G., & Andreas, Z. (2020). A comparative study of 4G and 5G network simulators. *International Journal on Advances in Networks and Services*, 13(1).
- Can, K. (2021). The evolution of communication technologies in Turkey's modern economy. *Journal of Policy Options*, 4(3), 11–17.
- Ceylan, H., Demir, G., & Elri, Z. (2021). Internet speed issue of Turkey. *International Journal of Engineering and Innovative Research*, 3(2), 121–132.
- Chiplunkar, G., & Goldberg, P. K. (2022). The employment effects of mobile internet in developing countries (No. w30741). National Bureau of Economic Research.
- Çalış, K., & Özdemir, S. (2013). Yeni nesil mobil genişbant teknolojileri ve Türkiye. *Akademik Bilişim*, 935–941.
- Çemen, M. F. (2023). A UPF module for 5G networks with quality of service support: Software implementation and realistic evaluation (Master's thesis, Middle East Technical University).
- Demircioğlu, E., Ertok, H. H., & Çankaya, İ. (2024). The non-terrestrial communication network applications in Türkiye on the way to 6G—Universal Project Phase 3. In *2024 32nd Signal Processing and Communications Applications Conference (SIU)* (pp. 1–4). IEEE.
- Eryılmaz, E. (2022). 1G'den günümüze hücresel ağlar. *Aselsan Dergi*, 111, 6–16.
- Evcı, C., Ciliz, K., Anarım, E., & Sankur, B. (2004). Wireless networks in Turkey: A jewel in the crowd. *Alcatel Telecommunications Review*, (3), 320–326.
- Geylani, M., Çıbuk, M., Çınar, H., & Ağgün, F. (2016). Geçmişten günümüze hücresel haberleşme teknolojilerinin gelişimi. *Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi*, 18(54), 606–623.
- Goggin, G. (2021). *Apps: From mobile phones to digital lives*. John Wiley & Sons.
- Göken, Ç., Atak, O., Bereketli, A., & Tunalı, A. İ. (2022). 3GPP tarihçesi, standardizasyon süreci ve ASELSAN'ın 3GPP katılımı. *Aselsan Dergi*, 111, 16–26.
- Halonen, T., Romero, J., & Melero, J. (Eds.). (2004). *GSM, GPRS and EDGE performance: Evolution towards 3G/UMTS*. John Wiley & Sons.
- Holma, H., & Toskala, A. (2010). *WCDMA for UMTS: HSPA evolution and LTE*. John Wiley & Sons.
- Hytönen, V. (2015). *Enhancing the performance of HSDPA communication with multipoint transmission techniques* (Doctoral dissertation, University of Jyväskylä).

- Ibrahim, S. K., Singh, M. J., Al-Bawri, S. S., Ibrahim, H. H., Islam, M. T., Islam, M. S., ... & Abdulkawi, W. M. (2023). Design, challenges and developments for 5G massive MIMO antenna systems at sub 6-GHz band: A review. *Nanomaterials*, 13(3), 520.
- Jiang, W., & Han, B. (2024). Evolution to fourth-generation (4G) mobile cellular communications. In *Cellular communication networks and standards: The evolution from 1G to 6G* (pp. 107–124). Springer.
- Lehr, W., Queder, F., & Haucap, J. (2021). 5G: A new future for mobile network operators, or not? *Telecommunications Policy*, 45(3), 102086.
- Mohr, W., & Konhauser, W. (2002). Access network evolution beyond third generation mobile communications. *IEEE Communications Magazine*, 38(12), 122–133.
- Mustari, N., Karabulut, M. A., Shah, A. S., & Türeli, U. (2023). 1G'den 6G'ye hücresele evrim üzerine kapsamlı bir derleme. *Politeknik Dergisi*, 1–1.
- Özaktaş, F. D., & Ardiyok, Ş. (2024). Policies and legislation: A historical overview and current perspectives on digitalization in Türkiye. In *The economics and regulation of digitalisation* (pp. 43–64).
- Peterson, L., & Sunay, O. (2020). *5G mobile networks: A systems approach*. Morgan & Claypool Publishers.
- Rao, A. M., Weber, A., Gollamudi, S., & Soni, R. (2009). LTE and HSPA+: Revolutionary and evolutionary solutions for global mobile broadband. *Bell Labs Technical Journal*, 13(4), 7–34.
- Routray, S. K., & Sharmila, K. P. (2016). 4.5G: A milestone along the road to 5G. In *2016 International Conference on Information Communication and Embedded Systems (ICICES)* (pp. 1–6). IEEE.
- Saily, M., Sebire, G., & Riddington, E. (Eds.). (2011). *GSM/EDGE: Evolution and performance*. John Wiley & Sons.
- Sathishkumar, N., Shree, S. H., Ilakkiya, B., & Janani, G. (2025). Design and development of a sub-6 GHz tri-band microstrip patch antenna for 5G communication. In *2025 Second International Conference on Cognitive Robotics and Intelligent Systems (ICC-ROBINS)* (pp. 105–109). IEEE.
- Saylor, M. J., & Saylor, M. (2025). *The mobile wave: How mobile intelligence will change everything*. Hachette+ ORM.
- Shimray, S. R., Keerti, C., & Ramaiah, C. K. (2015). An overview of mobile reading habits. *DESIDOC Journal of Library & Information Technology*, 35(5), 343–354.

- Steinbock, D. (2005). The mobile revolution: The making of mobile services worldwide. Kogan Page Publishers.
- Şevket, M., Ergin, S. B., Bişkin, S. B., & Eregiz, E. (2022). Dikey sektörlerde 5G. *Aselsan Dergi*, 111, 76–82.
- Toğa, G., & Toğa, M. (2019). Diffusion and substitution effect on telecommunication technologies in Turkey. *Academic Platform–Journal of Engineering and Science*, 7(3), 496–506.
- Tözer, A. (2023). Analysis of mobile telecommunications sector in Türkiye on the eve of 5G transition (Doctoral dissertation, Middle East Technical University).
- UBD; Türkiye 4.5G'ye geçiş yaptı. (2016). <https://www.uab.gov.tr/haberler/turkiye-4-5g-ye-gecis-yapti>
- Uğur, İ., & Aksoy, M. S. (2022). Ulaşımın geleceği: Bağlantılı araçlar. *Aselsan Dergi*, 111, 86–88.
- Xu, Y., Gui, G., Gacanin, H., & Adachi, F. (2021). A survey on resource allocation for 5G heterogeneous networks: Current research, future trends, and challenges. *IEEE Communications Surveys & Tutorials*, 23(2), 668–695.
- Yavuz, M., Meshkati, F., Nanda, S., Pokhariyal, A., Johnson, N., Raghothaman, B., & Richardson, A. (2009). Interference management and performance analysis of UMTS/HSPA+ femtocells. *IEEE Communications Magazine*, 47(9), 102–109.
- Yıldırım, A., & Tunalı, A. İ. (2022). 5G nedir? *Aselsan Dergi*, 111, 26–32.

CHAPTER 5

A Simple Interactive Simulation and Learning Platform for Automotive Braking Systems

Ali Özhan Akyüz¹ & Kazım Kumaş² & Mustafa Ayan³ & Durmuş Temiz⁴

¹ Prof. Dr., Burdur Mehmet Akif Ersoy University, Bucak Emin Gulmez Vocational School of Technical Sciences, 0000-0001-9265-7293

² Öğr. Gör., Burdur Mehmet Akif Ersoy University, Bucak Emin Gulmez Vocational School of Technical Sciences, 0000-0002-2348-4664

³ Öğr. Gör., Burdur Mehmet Akif Ersoy University, Bucak Emin Gulmez Vocational School of Technical Sciences, 0000-0001-6780-0625

⁴ Öğr. Gör. Dr., Burdur Mehmet Akif Ersoy University, Bucak Emin Gulmez Vocational School of Technical Sciences, 0000-0002-7350-7502

1. INTRODUCTION

Braking systems are one of the most vital safety elements of automotive technology and are used to reduce, stop or maintain a vehicle's speed. Since the invention of the automobile, braking systems have continuously evolved and constitute one of the most important topics in automotive history, both in terms of engineering and road safety. The wooden block brakes used in early automobiles slowed the vehicle by rubbing against the wheel, and while these systems were partially adequate at low speeds and for light vehicles, they became inadequate as speed and mass increased (Halderman & Mitchell, 2004; Miller et al., 2013). In the early 20th century, mechanical drum brakes were developed, with pads pressing against a rotating drum inside the wheel, providing better braking but limited performance due to heating and adjustment problems. The introduction of hydraulic braking systems in the 1920s ensured that the force was transmitted evenly and more efficiently to the wheels by means of hydraulic fluid, and this system remains the basic principle of the vast majority of automobiles today. From the 1950s onwards, disc brakes became common on the front wheels and are now standard on almost all modern vehicles, as they provide more stable and powerful braking, especially at high speeds and in wet conditions. After the 1980s, advanced braking technologies such as ABS (Anti-lock Braking System), EBD (Electronic Brakeforce Distribution) and ESC (Electronic Stability Control) integrated with electronic control systems were introduced, allowing braking to be optimized not only mechanically but also electronically (Pi et al., 2019; Lutanto et al., 2024; Ghazaly et al., 2014).

Braking systems are basically divided into two main groups, mechanical and hydraulic; in mechanical brakes the force is transmitted directly through mechanical linkages, while in hydraulic brakes the force is transmitted through a fluid, and hydraulic systems are standard in modern automobiles. In addition, pneumatic brakes are used in heavy vehicles, while electromechanical and regenerative brakes are used in electric vehicles. Today, the success of a braking system is measured not only by its ability to stop a vehicle, but also by its consistency under different road and weather conditions, resistance to heat build-up, driver feel and control, integration with electronic safety systems, and minimum maintenance requirements. These limitations have necessitated the development of hydraulic systems (Hamsini & Kathires, 2021; Fischer & Kolagari, 2025).

The complexity of braking systems and their dependence on a large number of parameters has increased the importance of simulation in engineering, and web-based brake simulations have made it possible for both students and professionals to quickly and visually analyze the effect of different parameters and road conditions on brake performance.

Figure 1 shows the brake system simulated in this study. The braking process starts with the driver applying a force to the brake pedal (Pedal Force). This force is amplified by the leverage ratio provided by the pedal mechanism (Pedal Mechanical Ratio) and then further increased by the brake booster or hydrovac (Brake Booster / Servo Ratio). The increased force is transmitted to the master cylinder, where it is converted into hydraulic pressure depending on the diameter of the master cylinder (Master Cylinder Diameter). This hydraulic pressure is transmitted through the brake pipes (Hydraulic Lines) to the caliper on the wheel (Brake Caliper). The caliper's piston diameter (Caliper Piston Diameter) ensures that the hydraulic pressure is converted into a force that is transmitted to the pads. The caliper pistons press the brake pads against the brake disc, where the Pad Friction Coefficient (Pad Friction Coefficient) between the pads and the disc comes into play and this friction generates braking torque. The diameter of the brake disc (Brake Disc Diameter) plays an important role in converting the applied force into torque transmitted to the wheel. The resulting braking torque slows the wheel and the vehicle's motion energy is converted into heat. The total mass of the vehicle (Vehicle Mass) determines the amount of kinetic energy that must be stopped by braking, while the coefficient of friction between the wheel and the road (Tire/Road Friction Coefficient) limits the maximum force that can be transmitted during braking and therefore the minimum stopping distance. At the end of this whole process, the kinetic energy of the vehicle is converted into heat by the braking system and the vehicle comes to a safe stop (Kinkaid et al., 2003; Kumar et al., 2014; Ouyang et al., 2005; Dhir, 2018).

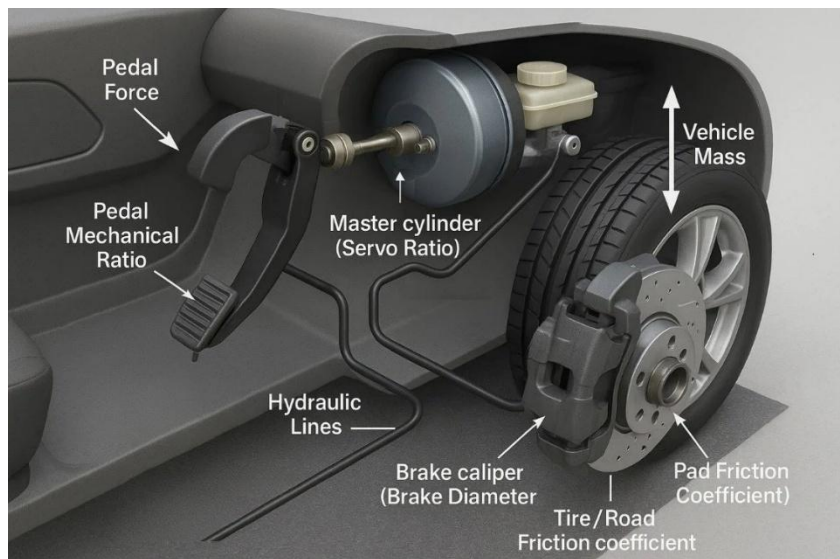


Figure 1. Automotive Disc Brake System Schematic (with Simulation Variable Labels)

In this study, the basic physical principles and components of an automotive disc brake system were modeled with realistic parameters and interactively studied in a simulation environment. In the code, starting from the force applied by the driver to the pedal, all basic processes such as pedal mechanical ratio, hydrovac support, hydraulic pressure generation, caliper and pad forces, disc and wheel dynamics and tire/road friction are calculated step by step; performance outputs such as stopping distance, braking time, brake temperature are presented to the user in real time. Thus, how the braking behavior changes under different vehicle types, road conditions and system parameters has been visually and numerically analyzed, and the working logic of the braking system and its importance in terms of safety has been presented.

2. LITERATURE REVIEW

Kisilowski and Lozia (1986) developed a mathematical model of vehicle braking on uneven road surfaces. Road irregularities were modeled as a random process defined by power spectral density. Six different tire radial property models were examined, and braking performance was analyzed on both flat and uneven roads. The results highlighted the significant influence of road surface roughness on braking efficiency (Kisilowski & Lozia, 1986).

Gerdes and Hedrick (1999) presented reduced-order physical models of brake system dynamics. Models for the vacuum booster and hydraulic brake systems were developed, combining static and dynamic effects. The relationship between pedal force and brake pressure was experimentally validated, showing improved accuracy compared to earlier literature. These models allowed for the design of vehicle control systems and the study of driver–vehicle interactions (Gerdes & Hedrick, 1999).

Day and Roberts (2002) introduced a new ABS model within the HVE (Human Vehicle Environment) simulation framework. ABS operating principles were described, and the developed model was integrated into the SIMON vehicle simulation platform. It showed strong agreement with experimental pressure cycles. The effects of ABS were evaluated for straight-line braking and ISO 3888 lane-change maneuvers, demonstrating that ABS improves not only stopping distance but also stability and maneuverability (Day & Roberts, 2002).

Li et al. (2008) proposed a simplified braking model based on kinetic energy conversion. A single-wheel mathematical model was developed, where the kinetic energy dissipated through tire-road friction and brake pad friction. The model was implemented in MATLAB/Simulink, and simulations showed that braking distances matched known data, verifying the model's reliability. This provided an alternative to conventional complex brake models (Li et al., 2008).

Magargle et al. (2017) emphasized brake system monitoring and predictive maintenance with a simulation-based digital twin. Developed in Modelica and integrated with ANSYS, the digital twin enabled fault injection, thermal monitoring, and predictive diagnostics. This model-based approach allowed for comprehensive system health monitoring and preventive maintenance strategies for automotive braking systems (Magargle et al., 2017).

Kalikate et al. (2018) investigated the potential of MR brakes as an alternative to conventional hydraulic systems. Using MATLAB/Simulink, braking performance was compared against Indian automotive braking standards. Results showed that a single MR brake could not meet the stopping distance requirements, necessitating at least six MR disks. Due to weight and packaging constraints, MR brakes were found impractical for most automotive applications (Kalikate et al., 2018).

Viselga et al. (2021) developed a comprehensive simulation model that included wheels, brake shoes, drums, hydraulic cylinders, vacuum boosters, and pressure regulators. Each parameter's influence on braking distance was evaluated, enabling designers to assess and optimize system safety at the design stage. Compared with physical testing, the simulations provided significant time and cost savings, while also allowing multiple test scenarios to be analyzed virtually (Viselga et al., 2021).

Quan et al. (2023) studied modern automatic braking systems designed to reduce accidents caused by human error and delayed reaction. Ultrasonic sensors were employed for obstacle detection, and a fuzzy logic controller was implemented in MATLAB/Simulink. Simulation results showed noticeable reductions in stopping distance and time, demonstrating the effectiveness of driver-assist braking control (Quan et al., 2023).

Kacprzycki et al. (2025) modeled ABS in MATLAB/Simulink and compared different road adhesion conditions (dry/wet asphalt, snow, ice). The relationship between slip ratio and friction was modeled using the Magic Formula. The performance of ABS, ESP, ASR, and EBD systems was analyzed, and stopping distances were compared with and without ABS, highlighting its critical importance particularly on low-adhesion surfaces (Kacprzycki et al., 2025).

Gowda et al. (2025) focused on the V-model approach for the development of braking control systems. ABS was modeled as an example case, and the simulation process was analyzed across different levels: MIL (Model-in-the-Loop), SIL (Software-in-the-Loop), HIL (Hardware-in-the-Loop), and DIL (Driver-in-the-Loop). This structured methodology demonstrated how V-diagram-based development improves the reliability of testing and validation in

the automotive industry (Gowda et al., 2025). An overview comparison of the cited literature is presented in Table 1.

Table 1. Summary comparison of related literature and the present study

Reference	Literature focus	Difference in this study
Kisilowski and Lozia (1986)	Road irregularities modeled via Power Spectral Density; tire radial property models analyzed.	This study addresses surface friction with scenario-based simple parameters, not PSD irregularity modeling; simplification chosen for educational purposes.
Gerdes and Hedrick (1999)	Vacuum booster and hydraulic brake dynamic models with experimental validation.	In this study, the force–pressure–area chain is modeled in real time, but detailed booster transition dynamics are simplified.
Day and Roberts (2002)	ABS model in HVE/SIMON; replicated experimental pressure cycles and evaluated ISO 3888 maneuvers.	In this work, ABS is modeled with a limit-based approach; quick web-based comparison is possible, but no ISO maneuvers or cycle replication are included.
Li et al. (2008)	Simplified kinetic energy conversion single-wheel model, implemented in Simulink.	The present study follows a similar simplification philosophy but deploys it on the web without installation, with multi-parameter UI, real-time charts, and animations.
Magargle et al. (2017)	Digital twin for predictive maintenance, fault injection, and thermal monitoring (Modelica + ANSYS).	This study focuses on educational and interactive visualization; predictive maintenance and digital twin functions are not implemented.
Kalikate et al. (2018)	MR brakes compared against standards in Simulink; six MR discs required to satisfy regulations.	In this study, interactive exploration is limited to conventional hydraulic brake parameters; MR technologies are excluded but can be added in future extensions.

Viselga et al. (2021)	Comprehensive brake system model (wheel, drum, master, booster, regulators) for design optimization.	The present study emphasizes user-oriented exploration and browser-based accessibility rather than engineering optimization.
Quan et al. (2023)	Autonomous braking with ultrasonic sensors and fuzzy logic controller in Simulink.	This study centers on driver input and does not include ADAS or fuzzy control; however, the interface could serve as a basis for future expansions.
Kacprzycki et al. (2025)	ABS, ESP, ASR, EBD simulated with Magic Formula slip ratio–friction modeling under various surfaces.	In this study, wheel slip is modeled as a simple friction limit; multi-system comparative analysis is not conducted (possible future addition).
Gowda et al. (2025)	V-model (MIL/SIL/HIL/DIL) applied to brake control development and validation.	The present study is not integrated into validation chains but instead offers a lightweight, web-based platform for prototyping and teaching.

This study differs from previous research in terms of both its approach and objectives. While the literature has mainly focused on accuracy, experimental validation, control algorithms, detailed physical modeling, design verification (MIL/SIL/HIL/DIL), and predictive maintenance through digital twins, the present study emphasizes accessibility, usability, and interactive learning. By offering a fully client-side web-based environment, it provides real-time computation and visualization for educational and demonstrative purposes.

From an architectural and technological perspective, the study adopts a lightweight, modular, and sustainable design. Since the simulation runs entirely on the client side, there is no backend dependency, and it can be directly accessed on static hosting platforms such as GitHub Pages or Netlify. HTML5 is used for semantic input structure, CSS3 provides flexible and responsive design, and modern JavaScript (ES6+) is used for functional logic. Parameters are processed through a single update function, while Chart.js ensures real-time and responsive data visualization. Animations are managed via requestAnimationFrame, and Web Audio API is used for sound feedback. Unlike the literature, which typically

relies on specialized engineering tools such as MATLAB/Simulink, Modelica, or ANSYS, this study demonstrates that a purely web-based and installation-free architecture can serve as an effective pedagogical tool.

In terms of modeling depth and simplifications, the present study deliberately follows a simplified yet transparent approach. Physical relations such as pedal force, hydraulic amplification, braking torque, and thermal effects are calculated in SI units and presented in real time. However, details such as the “effective radius” in disc torque relations, slip–friction curves modeled by the Magic Formula, or complex hydraulic dynamics are deliberately omitted. These simplifications are not limitations but rather conscious design choices, ensuring that the simulation remains understandable and suitable for educational use.

Interactivity and accessibility are also among the strong contributions of this work. Sliders, numerical inputs, and checkboxes provide instant feedback; smooth animations, dynamic charts, and audio–visual cues enrich the user experience. The application is fully responsive and optimized for different devices, while accessibility considerations (form–label pairing, semantic tags, and i18n-ready structure) have been implemented. Whereas most literature studies remain confined to reporting experimental results or using platform-dependent simulations that require installation, this study provides an open and browser-based tool for quick interaction and exploration.

The user experience and educational aspects of the simulation stand out as well. Since the application works instantly without installation or licensing requirements, students and practitioners can directly explore cause–effect relationships between parameters and braking performance. By integrating accessibility standards, multilingual support potential, and a low learning barrier, the study enhances applicability in classrooms and workshops, making it especially beneficial in educational contexts.

Nevertheless, the scientific depth of the model is acknowledged with transparency. Detailed tire models, effective radius considerations, advanced heat transfer models, and ABS pressure cycle replications have been simplified. Similarly, systematic validation with benchmark data (e.g., ISO 21994 braking distance tests) has not yet been conducted. For this reason, the model is positioned as an educational and demonstrative tool rather than a high-fidelity engineering validation environment.

Despite these simplifications, the contribution of the present study is significant. It acts as a bridge between complex models and pedagogical demonstrations, quickly revealing parameter sensitivities and enabling comparative analysis of different vehicle–road conditions. As a reusable demonstrator operating in an open web environment, it facilitates the

understanding of the fundamental principles of braking systems without the high barrier of specialized software.

In addition, the study provides a clear roadmap for possible future expansions to bring it closer to high-fidelity models in the literature. These include the integration of slip–friction curves (Pacejka Formula), enrichment of the ABS model with pressure cycles, incorporation of effective radius and detailed heat dissipation models, validation against reference data, inclusion of basic ADAS braking scenarios, and expansion into system diversity (EBD, ASR, ESP, regenerative braking, or MR brakes). This approach demonstrates the modularity and extensibility of the architecture.

In conclusion, the present study distinguishes itself from the literature by offering an accessible, browser-based, real-time, and user-focused braking simulation instead of heavy desktop- or hardware-based ecosystems. By combining architectural simplicity, client-side performance, interactive visualization, and accessibility, it delivers a sustainable tool with broad applicability. Positioned at the intersection of education, demonstration, and rapid parameter exploration, this simulation represents an original and valuable contribution.

3. MATERIAL AND METHOD

The complexity of braking systems and their dependence on a large number of parameters has increased the importance of simulation in engineering. Web-based brake simulations allow both students and professionals to quickly and visually analyze the effect of different parameters and road conditions on brake performance. In this way, the limits and behavior of the system can be understood theoretically without the need for real vehicle tests.

During the development of the simulation, we adopted a modular and sustainable web architecture that works entirely on the client side. The basic structure of the application was built in HTML5, with forms such as `<input type="range">`, `<input type="number">`, and `<input type="checkbox">` for all parameters received from the user. Each parameter is associated with semantically meaningful and accessible tags, thus ensuring both user experience and accessibility standards.

The interface design makes extensive use of the advanced features of CSS3. While a modern look is achieved with linear-gradient in the backgrounds, the components are placed flexibly and aligned thanks to the flexbox structure. Transition effects make user interactions more fluid, and details such as box-shadow and border-radius add visual depth and softness. Responsive design is ensured with media queries, ensuring that the application is consistent and usable on desktop, tablet and mobile devices.

The functional logic of the application was written entirely in JavaScript (ES6+). All physical quantities such as pedal force, hydrovac ratio, system efficiency, cylinder and caliper areas, hydraulic pressure, brake force, wheel/road limit, brake torque, deceleration and stopping distance are calculated in real time. In addition, cumulative and cooling effect models have been implemented for time-dependent variables such as brake temperature. All these operations are performed entirely on the scanner and with high performance, without the need for any backend service.

Chart.js library was preferred for the visual analysis and data presentation layer of the simulation. Parameters such as speed, deceleration, brake force and temperature are visualized with dynamically updated charts that are responsive to user interaction. Chart.js and other external libraries are loaded via CDN, optimizing the loading time of the application and eliminating the need for users to perform an additional installation. The main reason for preferring Chart.js over alternatives such as D3.js or Plotly is its ease of use, lightweight structure, and quick integration capabilities. The library significantly speeds up the learning and implementation process, particularly because it provides simple and commonly used chart types (line, bar, pie, etc.) in a ready-to-use and easily customizable format. Additionally, being Canvas-based makes it efficient in terms of performance, and its small size ensures it does not add extra load to web projects. Thanks to its widespread community support and comprehensive documentation, quick solutions and examples can be found when needed. These advantages make Chart.js a more practical and accessible option compared to D3.js, which is more flexible but also more complex, or Plotly, which offers comprehensive interactions but has a heavier structure.

In addition, vehicle and road animations were managed in synchronization with JavaScript to visually track the braking process. Animations and sound effects triggered by user interaction increased the realism and interactivity of the simulation. By combining all these technological components, a highly accessible and user experience-oriented braking simulation that works entirely on the client side and offers real-time computation and visual analysis has been created.

All of the technical elements used in the simulation and their functions are summarized in Table 2.

Table 2. Technical components used in the simulation

Layer/ Component	Technology / Approach Used		Critical Details and Optimizations	Main Reference in Code
Structural Frame	HTML5, tags	semantic	Each parameter input was associated with a label; meaningful div hierarchy was established for partitioning	<div id="inputs">, <div id="results">
Form Inputs	range, checkbox, select	number,	Instant feedback with min-max, step and oninput bindings; simple logic without input debouncing	<input oninput="updateAll()">
Style Layer	CSS3, Flexbox, Grid, Gradient		Gradient background → GPU friendly; fluid layout with flexbox; 60fps target with transition	body{background:linear-gradient...}
Responsive Design	Media (@media), Layout	Queries Fluid	Breakpoints: 1100 px and below; overflow prevented on small screen with .chart-container{width:95vw}	Bottom @media (max-width: 1100px)
Thematic Visuals	CSS Properties	Custom	Root variables can be added for color palette change (:root{--primary:#3b82f6;})	Design layer
Operational Logic / Computation	JavaScript strict mode	(ES6+),	updateAll() single responsible function - idempotent; physical formulas calculated by converting to SI units	function updateAll(){ ... }
Physical Models	Kinetic Newtonian mechanics	energy,	No drift; optional heat capacity parameter can be added (heatCapacity)	brakeEnergy = 0.5 * mass * v ²
State Management	Global variable + DOM state		currentTemp, animationRunning flag; low complexity → no need for React/vue	Upper script block
Data Visualization	Chart.js (v4, CDN)		4 independent line graphics; update-in-place avoids recreation cost; colors dynamic according to ABS status	if (!chartSpeed)... else { chartSpeed.update('none') }
Animation	requestAnimationFrame, CSS Keyframes		Road (roadMove) CPU-light; vehicle position changing left on JS side; animationRunning lock prevents race conditions	animateVehicle()
Audio /	Web Audio API		Minimal oscillator-	playSound(f, t)

Media			gain chain; ear-friendly with 0.05 gain limit; restart spam protection	
Error Management	Try/Catch, Clauses	Guard	Divide by zero check (if(masterArea_m2>0)); audioContext support check	Various try {} blocks
Performance	Lazy Chart Init, Re-use Canvas		Charts are created on the first call; on subsequent updates only the data field changes; anim removed with update('none')	Chart update section
Accessibility (a11y)	Tag-input matching, ARIA potential		<label> for each input; optional role="slider" can be added	Form HTML
Globalization (i18n)	Text separation, emoji iconization		UI texts in one file; JSON dictionary for different languages can be easily integrated	HTML internal texts
Distribution / Loading	CDN (Chart.js) + Pure Client-side		No server dependency; 0 backend → Ready to go live instantly on a static host like GitHub Pages / Netlify	<script src="https://cdn.jsdelivr.net/npm/chart.js">
Expandability	Modular Functions, Parametre Abstraction		A separate block can be opened in updateAll() to add new brake models (e.g. regenerative)	// TODO comment lines
Safety	CSP applicable, No eval		Fully local computation → XSS surface only static HTML; inserted inline script can be restricted by CSP	Header / meta recommendation

All physical parameters used in the study have been described in detail below.

Pedal Force refers to the force directly applied by the driver to the brake pedal and is the starting point of the braking process. When the driver wants to slow or stop the vehicle, he or she applies a certain force to the brake pedal with his or her foot. This force is transmitted through all components of the braking system and amplified in various proportions, eventually becoming the braking force acting on the wheels. Since Pedal Force is the main parameter that transmits the driver's braking request to the system, the magnitude of the applied force directly affects the braking performance. In the simulation, this value is represented by the variable `pedalForce` and is usually measured in Newtons (N). The more force the driver applies to the pedal, the higher the braking force on the rest of the system. However, the magnitude of this force is also multiplied by other

parameters such as pedal mechanical ratio, hydrovac ratio and system efficiency to create the final braking effect.

Pedal Mechanical Ratio is an important parameter at the very beginning of the braking system that determines how much the force applied by the driver to the pedal is magnified. The brake pedal works on the principle of a lever; the force the driver applies when pressing the pedal is transferred through the long arm of the pedal to a shorter arm, which in turn pushes the piston of the master cylinder. The ratio between the distances to the pivot point of the pedal, the mechanical ratio of the pedal, determines how much the applied force is magnified. In this way, a relatively small force applied by the driver is converted into a much larger force by the pedal mechanism and transmitted to the braking system. For example, in a system with a pedal mechanical ratio of 4, when the driver applies 50 N force to the pedal, a force of 200 N is obtained at the pedal output. In the simulation, this ratio is represented by the `pedalRatio` variable and as its value increases, the force transmitted to the brake system increases with the same pedal force.

Brake Booster Ratio (Hydrovac/Servo Ratio) is a force multiplier used in the braking system to further magnify the force the driver applies to the pedal. The hydrovac, or “servo”, provides an additional boost to the mechanical force applied by the driver, usually by using vacuum power from the engine's intake manifold or a specialized pump. In this way, the driver can achieve a much higher braking force with less pedal force. The force from the brake pedal is first amplified by the pedal mechanical ratio, then the hydrovac kicks in and increases this force again by the “Brake Booster Ratio”. For example, if the hydrovac ratio is 5, the force from the pedal is multiplied five times and delivered to the brake master cylinder. This ratio is represented in the simulation by the `servoRatio` variable, and as its value increases, the driver's braking effort decreases; more brake force is achieved with the same pedal force. The hydrovac (servo) system is an indispensable component that improves driving comfort and safety, especially in modern vehicles. Even when the driver presses the brake pedal lightly, a high force can be transmitted to the braking system through the hydrovac.

System Efficiency is a parameter that expresses the overall performance of the braking system and the losses in force transmission. In an ideal world, the force applied by the driver to the pedal, magnified by the pedal mechanical ratio and the hydrovac ratio, would all be available for braking. In the real world, however, no system is 100% efficient. Various losses occur in the braking system during force transmission. These losses can be caused by friction in the pedal mechanism, hydraulic lines or caliper pistons, slight expansion of hydraulic hoses or other components under pressure, conversion of some of the energy into heat,

or small air bubbles in the hydraulic system. Especially if the brake fluid has not been bled, air bubbles in the system prevent the pressure from being fully transmitted and reduce efficiency. The System Efficiency parameter expresses these losses as a percentage. For example, if the system efficiency is 90%, only 90% of the total force amplified by the pedal and hydrovac is available for braking; the remaining 10% is wasted due to losses in the system. Formulaically, the efficient output force is obtained by multiplying the booster output force by the system efficiency. In simulation, this value is represented by the efficiency variable and is usually used as a percentage (e.g. between 0.70 and 1.00). A higher system efficiency means less energy loss and therefore more effective braking.

Master Cylinder Diameter refers to the inside diameter of the master cylinder located at the beginning of the hydraulic portion of the brake system. The master cylinder is a critical component that converts the mechanical force from the brake pedal into hydraulic pressure. The force applied by the driver to the pedal and amplified by the pedal mechanism and hydrovac pushes the piston inside the master cylinder. This piston compresses the brake fluid inside the cylinder, creating hydraulic pressure. The diameter of the master cylinder directly affects how much pressure is generated by a given force. A smaller diameter allows higher hydraulic pressure to be generated with the same force. This is because the force acts on a smaller area. The master cylinder diameter affects not only the pressure generation, but also the feel of the brake pedal. While a smaller diameter produces higher pressure, the piston may have to travel a longer distance. This can result in a softer or longer pedal travel. A larger diameter, on the other hand, can provide a “harder” pedal feel and a shorter pedal stroke, but produces less pressure with the same force. The master cylinder diameter also affects the hydraulic flow rate in the system. Choosing the right master cylinder diameter is vital to strike a balance between the braking needs of the vehicle and the driver's preferred pedal feel. In the simulation, this value is represented in millimeters by the masterDiameter variable. This parameter plays a key role in braking performance, as it directly determines the hydraulic pressure generation capacity of the brake system.

Caliper Piston Diameter refers to the diameter of the piston inside the caliper, the hydraulic component of the brake system closest to the wheel. During braking, hydraulic pressure from the master cylinder is transmitted through the brake hoses to the caliper, where the piston(s) inside the caliper move under the influence of this pressure. The movement of the piston pushes the brake pads against the brake disc, thus creating friction between the disc and the pad and slowing down the vehicle. The diameter of the caliper piston directly determines the relationship between the hydraulic pressure applied and the brake force transmitted to the wheel. A larger caliper piston diameter will have more surface

area under the same hydraulic pressure, increasing the total brake force transmitted to the wheel. The larger the caliper piston diameter, the greater the force applied to the brake pads, resulting in more powerful braking. However, increasing the caliper piston diameter also increases the volume of hydraulic fluid in the system. In this case, the master cylinder has to send more fluid, which can lead to a slightly longer pedal stroke. Also, pistons that are too large can soften the pedal feel and change the response of the system. Therefore, the choice of caliper piston diameter is important to strike a balance between braking power and pedal feel. In the simulation, this parameter is represented in millimeters (mm) by the `caliperDiameter` variable. The correct caliper piston diameter directly affects the braking performance of the vehicle, pedal feel and the overall response of the system.

Brake Disc Diameter refers to the diameter of the discs that rotate in relation to the wheels of the vehicle and are compressed by the pads during braking. By making contact with the brake pads, the brake discs convert the kinetic energy of the vehicle into heat through friction, thus slowing the vehicle down. The disc diameter is a key parameter that directly affects the amount of torque the brake system can produce and its thermal capacity. A larger brake disc diameter increases the distance of the pads from the center of the disc. This means that the applied brake force is multiplied by a longer arm distance, which increases the brake torque transmitted to the wheel. Torque is calculated by multiplying the force by the arm length, where the arm length is the radius of the disc. So, when using a larger disc diameter with the same braking force, more braking torque can be applied to the wheel, resulting in more effective braking. Also, because larger diameter discs have more surface area, they can better dissipate and cool the heat generated during braking. This helps to prevent brake performance fading, especially at high speed or during prolonged braking. However, too large a disc diameter can also affect the vehicle's suspension and wheel design and increase weight. In the simulation, this parameter is represented in millimeters by the variable `disk Diameter`.

The model in this study directly addresses the relationship between disc diameter and torque, but the concept of 'effective radius', which reflects the area where the brake pads contact the disc, has not been included. The reason for this is to ensure that the designed simulation provides results to the user more simply and understandably, and to reduce the complexity of the formulas. Although including the effective radius would produce more accurate results from an engineering perspective, simplification was chosen because the current work is for educational/demonstration purposes.

Pad Friction Coefficient is a fundamental parameter that determines the amount of friction between the brake pads and the brake disc and has a direct

impact on the performance of the brake system. During braking, the force applied to the pads by the caliper pistons presses the pads firmly against the disc surface. As a result of this contact, the pads slide on the disc, creating a frictional force, which causes the vehicle's motion energy to be converted into heat, causing the vehicle to slow down or come to a complete stop. The coefficient of friction is a dimensionless quantity that expresses the ratio of the frictional force between two surfaces to the force (normal force) applied perpendicular to those surfaces. In other words, the higher the coefficient of friction of the lining in the brake system, the more frictional force and therefore more effective braking is achieved with the same compression force. The lining coefficient of friction is affected by many factors, notably the lining material. Different pad materials, such as organic, semi-metallic or ceramic, can exhibit varying frictional properties over different temperature ranges and under different road conditions. The coefficient of friction can also change as the pad and disc temperature increases during braking; some pads perform better at higher temperatures, while others may suffer a loss of braking performance known as “fading” when overheated. In addition, the ingress of water or moisture between the pad and the disc can significantly reduce the coefficient of friction and reduce braking effectiveness. The applied pressure can also slightly affect the coefficient of friction. In simulation, this parameter is represented by the friction variable and usually takes a value between 0.1 and 0.8. While a high pad friction coefficient means shorter stopping distances and more powerful braking, it can also lead to more heat generation during braking and faster wear of the pads. Therefore, the choice of pad friction coefficient is crucial to strike a balance between braking performance, durability and safety.

Tire/Road Friction Coefficient is a fundamental parameter that determines the amount of friction between the vehicle's tires and the road surface and plays a critical role in all handling dynamics such as braking, acceleration and cornering. During braking, how much of the force transmitted to the wheels by the braking system can be transferred to the road surface and how short a distance the vehicle can stop depends largely on this coefficient. The coefficient of friction is defined as the ratio of the frictional force between the tire and the road to the perpendicular force (normal force) exerted by the tire on the ground and is a dimensionless quantity. The value of this coefficient varies greatly with the condition of the road and the characteristics of the tire. On a dry, clean asphalt road, this value is usually around 0.8-1.0, whereas on wet, icy or muddy surfaces the coefficient of friction drops significantly. For example, it can be as low as 0.5 on wet roads and 0.2 on icy roads. The type of tire, tread depth, air pressure and the roughness of the road surface also affect this coefficient. The higher the coefficient of friction, the better the tires adhere to the road surface and the more force can be transferred during braking, which means shorter stopping distances and safer driving. In the simulation, this parameter is represented by the variable

tireFriction and can be set automatically according to the road condition. For example, the coefficient of friction is set to 0.8 when the road condition is dry, 0.5 when wet, 0.2 when icy and 0.3 when muddy. This value limits the maximum force that can be transmitted to the wheels during braking; if the braking force exceeds this limit, the wheels will lock and skidding will start. Therefore, the tire/road friction coefficient is one of the most important external parameters that determine how effectively the force that the braking system can generate can be transferred to the road.

Road Slope is an important parameter that expresses the angle of the road the vehicle is traveling on relative to the horizontal plane and directly affects braking performance. Road slope, usually expressed in degrees, represents an uphill road at positive values and a downhill road at negative values. Road gradient changes the components of the gravitational force acting on the vehicle during braking; on downhill descents, gravity pushes the vehicle forward, extending the braking distance and increasing the load on the braking system, so the driver has to apply more braking force to stop the vehicle. On uphill climbs, gravity helps to slow the vehicle down, shortening the braking distance and reducing the load on the braking system. In simulation, this parameter is represented by the slope variable and usually takes a value between -20° and $+20^{\circ}$.

Vehicle Mass is a fundamental physical parameter that expresses the total mass of a vehicle. In addition to the weight of the vehicle itself, this mass includes the weight of passengers, cargo and fuel. Vehicle mass directly affects all vehicle dynamics such as braking performance, acceleration, fuel consumption and handling. In the context of braking, vehicle mass is particularly important because the kinetic energy (energy of motion) of a vehicle is directly proportional to its mass. To stop a vehicle, all of this kinetic energy must be converted into heat by the braking system. So the greater the vehicle mass, the more kinetic energy needs to be dissipated to stop a vehicle traveling at the same speed. This requires more braking force and often leads to longer stopping distances. Heavy vehicles generate more heat and wear on brake discs and pads during braking. In the simulation, this parameter is represented in kilograms by the mass variable. When the user selects the vehicle type (car, truck, motorcycle), this mass value is automatically set to the typical mass of the respective vehicle, but can also be changed manually by the user.

Initial Speed is a fundamental dynamic parameter that expresses the speed of the vehicle at the moment when braking starts. This speed is one of the two main factors that determine the kinetic energy (energy of motion) of the vehicle (the other is the vehicle mass). To stop a vehicle, the braking system must convert all of this kinetic energy into heat. The effect of speed on kinetic energy is quadratic. So, when the speed doubles, the kinetic energy quadruples. This has a dramatic

effect on braking distance. The higher the vehicle speed, the greater the kinetic energy that needs to be stopped. This means that a much longer distance is needed to stop the vehicle with the same braking force. For example, if a vehicle traveling at 50 km/h has a stopping distance of X , the same vehicle traveling at 100 km/h may have a stopping distance of about $4X$. When braking at high speeds, much more heat and wear is generated on the brake discs and pads. This can strain the thermal capacity of the brake system and increase the risk of fading (reduced brake performance). In the simulation, this parameter is represented in kilometers per hour (km/h) by the variable speed. By varying this value, the user can observe the effects of different initial speeds on braking distance, deceleration and the braking system.

ABS (Anti-lock Braking System) is an electronic system that prevents the wheels from locking during braking, developed to improve safety in modern cars and commercial vehicles. In moments of sudden or hard braking, especially on slippery, wet or icy surfaces, the wheels can easily lock and the vehicle starts to skid, the driver loses steering control and the stopping distance of the vehicle increases. Thanks to sensors that continuously monitor the rotational speed of each wheel, ABS detects when one or more wheels begin to lock up and, via the central control unit, reduces and increases the brake pressure applied to the relevant wheel within milliseconds, allowing the wheel to turn. This process is repeated dozens of times per second, even when the brake pedal is kept depressed, allowing the wheels to continue turning in a way that ensures maximum grip with the road surface. The main advantage of ABS is that it maintains the vehicle's steering ability during braking, meaning that even when braking, the driver can steer the vehicle with steering movements and avoid obstacles. In addition, ABS maximizes the contact of the tires with the road and prevents uneven tire wear and the formation of "flat spots". In the simulation, this parameter is represented as a checkbox with the abs variable and, when active, the braking force cannot exceed the maximum allowed by the tire/road friction coefficient, so that wheel locking and skidding are not simulated. The user can observe the differences in braking performance and controllability of the vehicle when ABS is active or inactive. As a result, ABS not only provides a shorter and safer stopping distance in emergency situations, but also helps prevent accidents by maintaining the controllability of the vehicle, making it an indispensable active safety system in today's vehicles.

Vehicle Type is a parameter in the simulation that allows the user to choose between different categories of vehicles, such as cars, trucks or motorcycles, and to automatically set the typical physical parameters for the selected vehicle. This feature makes the simulation both user-friendly and realistic, as each vehicle type differs significantly from each other in terms of basic parameters such as mass,

brake disc diameter, caliper piston diameter and sometimes tire/road friction coefficient. For example, when a car is selected, the system usually assigns a medium mass (e.g. 1200 kg) and standard disc and caliper diameters; when a truck is selected, a much higher mass (e.g. 2500 kg) and larger brake components are automatically determined. When a motorcycle is selected, a low mass (e.g. 200 kg) and smaller brake components are used. When the user changes the vehicle type, the simulation updates the relevant parameters in the background, allowing a comparative analysis of the braking performance and dynamics of different vehicles. This automation allows users to quickly and practically observe the braking characteristics of different vehicle types without having to manually enter each parameter. It is also possible to understand why a truck stops longer than a car in real life or why a motorcycle may have a shorter stopping distance, based on mass and brake system dimensions.

Road Condition is an important parameter in the simulation that represents the physical properties and frictional conditions of the road surface and has a direct impact on braking performance. By selecting different road conditions such as dry, wet, icy or muddy, the user can experience the braking effects of various road conditions that can be encountered in real life. The selected road condition automatically changes the coefficient of friction (tireFriction) between the tire and the road in the background of the simulation; for example, this coefficient is high when a dry road is selected, whereas the coefficient of friction is significantly reduced on wet, icy or muddy surfaces. This change directly affects the maximum force that can be transferred to the wheels during braking and therefore the stopping distance of the vehicle. On a dry surface the tires have better grip on the road and can stop in a shorter distance, whereas on an icy or muddy surface the grip is reduced and the stopping distance is longer. Furthermore, the road condition can increase not only the braking distance, but also the risk of the vehicle skidding and skidding; especially on icy and muddy surfaces, the wheels are more likely to lock and the vehicle is more likely to lose control. The Road Condition parameter in the simulation realistically reflects the impact of different road surfaces and weather conditions on braking dynamics, giving the user the opportunity to better understand the principles of safe driving and braking and to compare different scenarios. In this way, drivers can make a more informed assessment of their braking behavior and safe following distances under varying road conditions.

4. RESULTS and DISCUSSIONS

This study is a comprehensive simulation and training tool that explains the physical and mechanical principles of the brake system to the user in a visual, interactive and educational way. The user can set many parameters such as vehicle type (car, truck, motorcycle), road condition (dry, wet, icy, muddy), pedal

force, pedal ratio, brake booster ratio, system efficiency, master cylinder and caliper diameter, disc diameter, pad friction coefficient, tire/road friction coefficient, road slope, vehicle mass, initial speed and ABS. In line with these parameters, values such as force and pressure transformations at all stages of the braking system starting from pedal force, brake torque, wheel/road limit, ABS activation status, brake force, deceleration and stopping distance are calculated with realistic formulas. In addition, brake temperature increases according to kinetic energy and pedal force, and brake fade in overheating is simulated. All these calculations are supported by animations and visuals, and presented to the user with dynamic elements such as pedal and system diagram, vehicle animation and temperature display. Chart.js visualizes the time-dependent change of values such as speed, deceleration, brake force and temperature. The user can easily compare results such as stopping distance and deceleration by adding different scenarios to the comparison table. In addition, a quiz section with short multiple-choice questions tests basic knowledge of the braking system, and correct/incorrect answers are enhanced with sound effects to increase interactivity. All these features, combined with a modern and mobile-friendly interface, allow engineering students and enthusiasts alike to experience and learn how the braking system works, what parameters affect what, and the importance of safety systems such as ABS.

Figure 2-4 shows a screenshot of the simulation. It is given in parts because it does not fit the full page. The simulations' source codes are available on GitHub <https://aliakyuz12.github.io/Brakesystem/> and can be used for educational purposes.

The first module of the platform (Figure 2) allows users to select the vehicle type (car, truck, motorcycle) and road condition (dry, wet, icy, muddy). In addition, a wide set of brake system parameters can be customized, including pedal force, mechanical ratio, booster ratio, master cylinder diameter, caliper dimensions, pad friction, tire–road adhesion coefficient, vehicle mass, gradient, and initial velocity. Users can also activate or deactivate ABS, enabling a versatile and realistic configuration of braking scenarios.

The second section (Figure 3) provides a detailed overview of system outputs. Brake temperature is continuously tracked, reflecting the thermal behavior of the braking process. The simulation instantly computes and displays key results such as pedal and booster output forces, hydraulic pressure, caliper and pad forces, brake torque, deceleration, tire/road friction limits, ABS status, and stopping distance. A comparison history module allows users to record multiple trials under different conditions and compare results side by side, making it possible to analyze the influence of various parameters.

In the final component (Figure 4), an integrated knowledge quiz evaluates user understanding of braking concepts, providing interactive reinforcement of theoretical knowledge. At the same time, dynamic graphs display braking behavior over time, including speed reduction, deceleration profile, brake force development, and temperature increase. These visualizations, combined with animations and interface feedback, transform abstract physical processes into concrete, engaging learning experiences, making the tool highly effective for educational purposes.

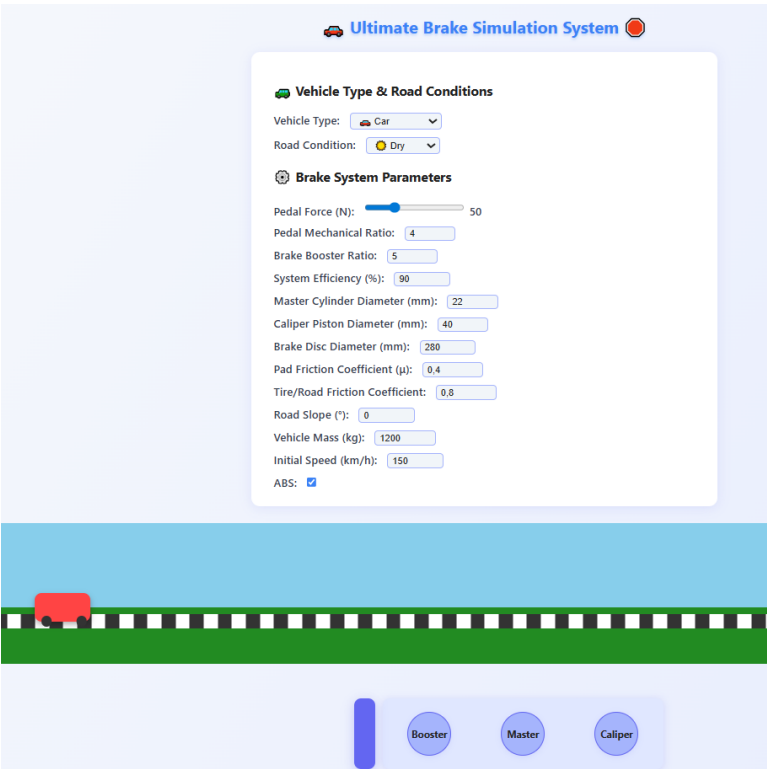


Figure 2. Vehicle Type, Road Control and Brake System Parameters

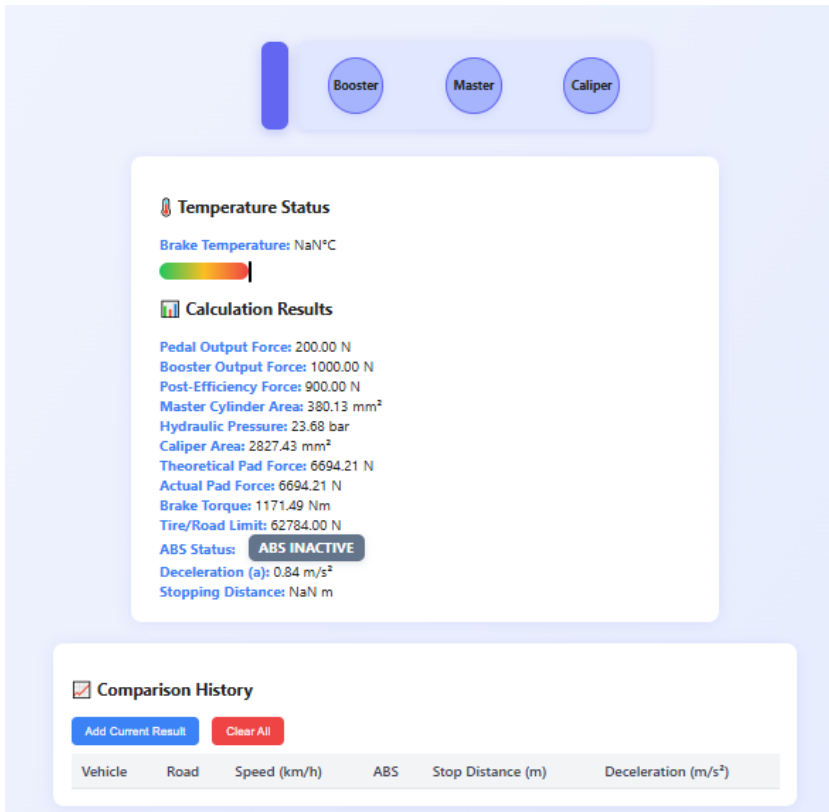


Figure 3. Temperature Status, Calculation Results, and Comparison History

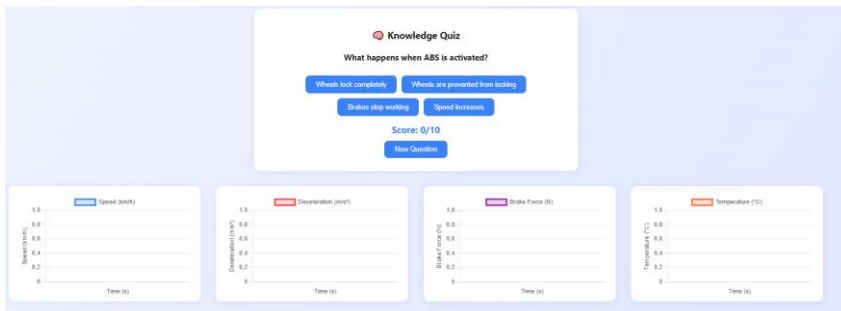


Figure 4. Knowledge Quiz and Real-Time Graphical Visualization

In the simulation, the pedal force received from the user is first multiplied by the pedal mechanical ratio and converted into pedal output force. This force is multiplied again by the brake booster ratio to obtain the booster output force and multiplied by the system efficiency to find the final force reaching the master cylinder. Using the diameter of the master cylinder, the master cylinder area is calculated using the circular area formula and the hydraulic pressure is obtained

by dividing the force applied to this area. In the same way, the caliper area is calculated from the caliper diameter and multiplied by the hydraulic pressure to calculate the theoretical force transmitted to the pads. Using the tire-road friction coefficient together with the vehicle mass, gravitational acceleration and road gradient, the maximum force the wheel can transmit is determined, i.e. the wheel/road limit.

If the pad force exceeds this limit, the force is reduced to this limit if ABS is engaged; if ABS is off, the force is reduced to seventy percent of the limit to simulate slip. The brake torque is calculated by multiplying the pad force by half the disc diameter. The deceleration value is obtained by dividing the brake force by the vehicle mass, and the road gradient is also taken into account. Finally, using the vehicle's initial speed and deceleration value, the stopping distance is calculated using the classical physics formula. In this whole process, the brake temperature increases with the kinetic energy released during braking and the applied pedal force; if the temperature rises too high, a reduction in pad force (fade) is simulated. Thus, all physical and mechanical processes of the braking system are modeled sequentially. During the simulation, instantaneous velocity, deceleration, brake force, and temperature values are continuously updated for animation and graphics. Table 2 shows all the physics formulas used in the simulation.

Table 2. Formulas used in the simulation

$F_{pedal,out} = F_{pedal} \cdot R_{pedal} \cdot R_{servo}$	(1)
$F_{\varepsilon} = F_{pedal,out} \cdot \varepsilon / 100$	(2)
$A_{caliper,master} = \pi \cdot (d_{caliper,master} / 2)^2$	(3)
$P = F_{\varepsilon} / A_{master}$	(4)
$F_{caliper} = P \cdot A_{caliper}$	(5)
$\tau_{brake} = F_{caliper} \cdot \mu_{pad} \cdot r_{disc}$	(6)
$\tau_{max,dis} = \mu_{tire} \cdot m \cdot g \cdot r_{disc}$	(7)
<i>If $\tau_{brake} > \tau_{max,dis}$ and ABS active $\rightarrow \tau_{brake,ABS} > \tau_{max,dis}$</i>	(8)
$F_{brake,wheel} = \tau_{brake} / r_{wheel}$	(9)
$a = \frac{F_{brake,wheel}}{m} \pm g \sin \theta$	(10)
$d = v_0^2 / 2a$	(11)
$t = v_0 / a$	(12)

$$T_{brake} = T_{initial} + \frac{\frac{1}{2}mv_0^2}{C_{brake}} \quad (13)$$

$$v(t) = v_0 - at \quad (14)$$

F_{pedal}: Force exerted by the driver, R_{pedal}: Pedal mechanical ratio,

R_{servo}: Hydrovac (servo) ratio (Brake booster ratio), F_{pedal,out} : Pedal output force,

ε: System efficiency %, Fε: Force after system efficiency,

A_{caliper, master}: Master cylinders and capillary areas, d_{caliper, master}: Diameter for circular pistons

P: Hydraulic pressure produced by the master cylinder, F_{caliper}: Caliper force

τ_{brake}: Brake torque (torque applied by the pads to the disc), μ_{pad}: Pad friction coefficient

r_{disc}: Disc radius, τ_{max, dis}: Maximum road torque (Wheel/road limit)

μ_{tire}: Tire/road friction coefficient, m: Vehicle mass, g=9.81 m/s²,

F_{brake, wheel}: Brake force transferred to the wheel, r_{wheel}: Wheel Radius, a: Acceleration

Θ: Road slope, d: Stopping distance, v₀: Initial velocity, t: Stop time

T_{brake}: Brake temperature, C_{brake}: Thermal capacity of the brake system (J/OC)

C_{brake} = m_{disc}.c_{disc} (8 kg. 900 J/OC), v(t): Velocity

The input parameters used in the simulation - their meaning and the values and value ranges they can take are summarized in Table 3.

Table 3. The input parameters used in the simulation and their meanings (Halderman & Mitchell, 2004; Breuer & Bill, 2003)

Parameter	Variable Name and Icon	Unit	Default Value	Min-Max Value	Description
Pedal Force	pedalForce (F _{pedal})	N	50	0-200	The force applied by the driver to the pedal
Pedal Mechanical Ratio	pedalRatio (R _{pedal})	-	4	1-10	Force magnification ratio of the pedal mechanism
Brake Booster Ratio	servoRatio (R _{servo})	-	5	1-10	Additional force ratio obtained with hydrovac (servo)
System Efficiency	Efficiency (ε)	%	90	70-100	Total efficiency of the braking system (including losses)
Master Cylinder Diameter	masterDiameter (d _{master})	mm	22	10-40	Master cylinder diameter
Caliper Piston Diameter	caliperDiameter (d _{caliper})	mm	40	20-60	Caliper Piston Diameter
Brake Disc Diameter	diskDiameter ((d _{disc})	mm	280	200-400	Brake Disc Diameter
Pad Friction Coefficient	Friction (μ _{pad})	-	0.4	0.1-0.8	Coefficient of friction between pad and disk
Tire/Road	tireFriction	-	0.8	0.1-1.2	Coefficient of tire-road

Friction Coefficient					friction
Road Slope	slope	Degree	0	-20-20	Slope of the road (in degrees, positive: uphill, negative: downhill)
Vehicle Mass	mass	Kg	1200	500-3000	Vehicle mass
Initial velocity	Speed	km/h	50	0-200	Velocity at the start of braking
ABS	Abs	- (true/false)	True	True/False	ABS active/passive status
Vehicle Type	vehicleType	-	-	-	Vehicle type (car, truck, motorcycle). Other parameters such as mass, disc diameter, caliper diameter are automatically adjusted according to the selection.
Condition	roadCondition	-	-	-	Road condition (dry, wet, icy, muddy). The tire/road friction coefficient is automatically adjusted according to the selection.

The simulation also includes a comparison table to compare different scenarios (no impact on the calculation) and animations and sound effects that only provide audio/visual feedback and have no impact on the physical calculations. A simple quiz system for students was also included (Table 4).

Table 4. Quiz questions in the simulation

What happens when ABS is activated?
Wheels lock completely - Wheels are prevented from locking - Brakes stop working - Speed increases
Which road condition provides the best braking performance?
Wet – Icy – Dry - Muddy
What is brake fade?
Brakes getting stronger - Brakes losing effectiveness due to heat - ABS activation - Pedal getting harder
Higher tire friction coefficient means:
Longer stopping distance - Shorter stopping distance - No change in stopping - ABS won't work
What does the brake booster do?
Reduces pedal force needed - Increases brake temperature - Activates ABS - Changes tire friction

5. CONCLUSION

This paper introduces a comprehensive simulation application that presents the basic principles and performance of vehicle braking systems to the user in both a visual and interactive way. By varying braking parameters for different vehicle types and road conditions, the user can instantly observe how critical values such as force, pressure, torque, deceleration, stopping distance and temperature are affected at all stages of the braking system. In addition, features such as animations, graphics, comparison tables and quizzes support the learning process and make the application fun and instructive.

This type of simulation is of great importance in education because students and users can develop a much more permanent and deeper understanding of complex physical and mechanical processes by changing parameters and directly observing the results, rather than learning theoretically. Especially in the field of engineering education and automotive technologies, such interactive tools contribute greatly to concretizing abstract concepts and understanding the real-life effects of safety systems (e.g. ABS). In conclusion, this simulation tool provides an opportunity for engineering students, automotive enthusiasts and trainers to learn the working logic of braking systems, the importance of safety technologies and the effects of different parameters on the system by experiencing them.

This study has some limitations. First, the simulation model developed addresses the dynamics of the brake system, which are quite complex in reality, through simplified mathematical relationships and assumptions. For example, temperature calculations were performed using a simple estimation model based solely on kinetic energy, excluding factors such as heat transfer, cooling, material properties, and repeated braking. Similarly, the ABS algorithm has been idealized; the complex sensor data, control logic, and nuances dependent on road conditions found in real systems are not represented in this model. Vehicle types (car, truck, motorcycle) are defined only by basic parameters such as mass, disc diameter, and caliper size; suspension, weight transfer, tire deformation, and aerodynamic effects are not included in the model. Furthermore, road conditions are represented by simple friction coefficient values, ignoring parameters such as surface roughness, temperature, or tire wear. Finally, since the study was created for visualization and educational purposes, it has not been validated with the high-accuracy experimental data required for engineering design; rather, the system focuses on conveying braking principles to users in an understandable manner.

REFERENCES

- Halderman, J. D., & Mitchell, C. D. (2004). *Automotive brake systems*. Pearson/Prentice Hall.
- Miller, J. I., Henderson, L. M., & Cebon, D. (2013). Designing and testing an advanced pneumatic braking system for heavy vehicles. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 227(8), 1715–1729.
- Pi, D., Cheng, Q., Xie, B., Wang, H., & Wang, X. (2019). A novel pneumatic brake pressure control algorithm for regenerative braking system of electric commercial trucks. *IEEE Access*, 7, 83372–83383.
- Lutanto, A., Fajri, A., Nugroho, K. C., & Falah, F. (2024). Advancements in automotive braking technology for enhanced safety: A review. *Multidisciplinary Innovations and Research in Applied Engineering*, 1(1), 6–20.
- Ghazaly, N. M., El-Sharkawy, M., & Ahmed, I. (2014). A review of automotive brake squeal mechanisms. *Journal of Mechanical Design and Vibration*, 1(1), 5–9.
- Hamsini, S., & Kathiresh, M. (2021). *Automotive safety systems*. In *Automotive Embedded Systems: Key Technologies, Innovations, and Applications* (pp. 1–18). Cham: Springer International Publishing.
- Fischer, A., & Kolagari, R. T. (2025). *Automotive Security Engineering*. In *Software Engineering 2025—Companion Proceedings* (pp. 10–18420). Gesellschaft für Informatik, Bonn.
- Kinkaid, N. M., O'Reilly, O. M., & Papadopoulos, P. (2003). Automotive disc brake squeal. *Journal of Sound and Vibration*, 267(1), 105–166.
- Kumar, G. R., Thriveni, S., Reddy, M. R., & Gowd, G. H. (2014). Design analysis & optimization of an automotive disc brake. *International Journal of Advanced Engineering Research and Science*, 1(3), 24–29.
- Ouyang, H., Nack, W., Yuan, Y., & Chen, F. (2005). Numerical analysis of automotive disc brake squeal: A review. *International Journal of Vehicle Noise and Vibration*, 1(3–4), 207–231.
- Dhir, D. K. (2018). Thermo-mechanical performance of automotive disc brakes. *Materials Today: Proceedings*, 5(1), 1864–1871.
- Kisilowski, J., & Lozia, Z. (1986). Modelling and simulating the braking process of automotive vehicle on uneven surface. *Vehicle System Dynamics*, 15(sup1), 250–263. <https://doi.org/10.1080/00423118608969140>

- Gerdes, J. C., & Hedrick, J. K. (1999). Brake system modeling for simulation and control. *Journal of Dynamic Systems, Measurement, and Control*, 121(3), 496–503. <https://doi.org/10.1115/1.2802501>
- Day, T. D., & Roberts, S. G. (2002). A simulation model for vehicle braking systems fitted with ABS. *SAE Transactions*, 111, 821–839.
- Li, W., Wang, X., Leng, X., & Wang, M. (2008, September). Modeling and simulation of automobile braking system based on kinetic energy conversion. In *2008 IEEE Vehicle Power and Propulsion Conference* (pp. 1–3). IEEE. <https://doi.org/10.1109/VPPC.2008.4677642>
- Magargle, R., Johnson, L., Mandloi, P., Davoudabadi, P., Kesarkar, O., Krishnaswamy, S., ... & Pitchaikani, A. (2017). A simulation-based digital twin for model-driven health monitoring and predictive maintenance of an automotive braking system. In *Modelica 2017 Conference Proceedings* (Vol. 132, pp. 35–44). <https://doi.org/10.3384/ecp1713235>
- Kalikate, S. M., Patil, S. R., & Sawant, S. M. (2018). Simulation-based estimation of an automotive magnetorheological brake system performance. *Journal of Advanced Research*, 14, 43–51. <https://doi.org/10.1016/j.jare.2018.05.009>
- Viselga, G., Ugnenko, E. B., Uzhviieva, E. N., Tymchenko, O. N., & Sorochuk, N. I. (2021). Simulation modeling of the automobile braking system performance. *IOP Conference Series: Materials Science and Engineering*, 1021(1), 012051. <https://doi.org/10.1088/1757-899X/1021/1/012051>
- Quan, V. H., Ngoc, N. A., Quynh, L. V., Khoa, N. X., & Tien, N. M. (2023). Modelling and simulation of automatic brake system of vehicle. *ARPJ Journal of Engineering and Applied Sciences*, 18(7).
- Kacprzycki, Ł., Brukalski, M., Abramowski, M. M., Sar, H., & Fundowicz, P. (2025). Simulation of an anti-lock braking system in a MATLAB/Simulink environment for various road adhesion properties. *Engineering Transactions*.
- Gowda, D. V., Ramachandra, A. C., Thippeswamy, M. N., Pandurangappa, C., & Naidu, P. R. (2025). Automotive braking system simulations V diagram approach. [Conference paper].
- Breuer, B., & Bill, K. H. (2003). Brake manual: Fundamentals, components, systems, vehicle dynamics; Bremsenhandbuch. Grundlagen, Komponenten, Systeme, Fahrndynamik. SAE International.

CHAPTER 6

A Digital Platform for Examining the Biomedical Foundations and Clinical Accuracy of Blood Gas Analysis

**Ali Özhan Akyüz¹ & Kazım Kumaş² &
Mustafa Ayan³ & Durmuş Temiz⁴**

¹ Prof. Dr., Burdur Mehmet Akif Ersoy University, Bucak Emin Gulmez Vocational School of Technical Sciences, 0000-0001-9265-7293

² Öğr. Gör., Burdur Mehmet Akif Ersoy University, Bucak Emin Gulmez Vocational School of Technical Sciences, 0000-0002-2348-4664

³ Öğr. Gör., Burdur Mehmet Akif Ersoy University, Bucak Emin Gulmez Vocational School of Technical Sciences, 0000-0001-6780-0625

⁴ Öğr. Gör. Dr., Burdur Mehmet Akif Ersoy University, Bucak Emin Gulmez Vocational School of Technical Sciences, 0000-0002-7350-7502

1. INTRODUCTION

Blood gas analysis is widely used in modern medicine as a rapid and reliable laboratory test for evaluating critically ill patients. This analysis can be performed on arterial, venous, or capillary blood samples; however, arterial blood gas (ABG) measurements generally provide the most accurate information because arterial blood directly reflects the oxygenated blood delivered to the tissues. Venous or capillary samples are preferred, especially in situations where invasive procedures are risky or when the patient's pain tolerance is low (Toffaletti & Rackley, 2021). The main advantage of blood gas analysis is its ability to provide rapid and comprehensive information about the patient's acid-base balance, respiratory function, oxygenation status, and metabolic compensation capacity. The major parameters measured include pH, $p\text{CO}_2$, $p\text{O}_2$, bicarbonate (HCO_3^-), base excess, and oxygen saturation (SaO_2). These parameters are used across a wide clinical spectrum, from monitoring patients on mechanical ventilation support in intensive care units to diagnosis and treatment in critical conditions like trauma, cardiac arrest, and shock in emergency departments. Furthermore, they play a vital role during anesthesia management in operating rooms and in the monitoring of chronic respiratory and cardiac diseases (COPD, asthma, interstitial lung diseases, heart failure) (Hennessey & Japp, 2015; Papoff, Moretti, & Nenna, 2024; Rabi, Kowal, & Ambalavanan, 2016). Blood gas analyses are critically important not only for diagnosis but also for evaluating treatment effectiveness and determining patient prognosis. Early detection of conditions such as hypoxemia, hypercapnia, metabolic acidosis, or alkalosis enables the timely implementation of appropriate therapeutic interventions. For this reason, blood gas measurements are considered one of the gold standard methods in intensive care and acute clinical practices (Toffaletti & Rackley, 2021). Historically, blood gas measurements have developed in parallel with advancements in the fields of biochemistry and biomedical engineering. This process, which began with manual titration methods in the early 20th century, progressed to its current level with the development of glass electrodes and pH meters, the refinement of CO_2 and O_2 measurements using Severinghaus and Clark electrodes, and subsequently the emergence of modern blood gas analyzers with integrated analyzers and microprocessor-based automation. These advancements have increased the clinical applicability of the test, shortened the measurement time, and improved patient comfort by reducing the sample volume (Severinghaus, 1986; Severinghaus, 2014).

2. HISTORY AND TECHNOLOGICAL DEVELOPMENT

The clinical applications of blood gas analysis have developed in parallel with technological advancements in the fields of biochemistry and biomedical engineering. In the early 20th century, manual titration methods were used to

assess the acid-base status of blood. These techniques involved measuring the capacity of the carbonic acid-bicarbonate buffer system by adding strong acid or base. However, these methods were both time-consuming and operator-dependent, making them impractical for clinical use (Severinghaus, 1986; Severinghaus, 2014).

A revolutionary advancement in pH measurement was achieved with the development of the glass electrode by Fritz Haber and Zygmunt Klemensiewicz in 1909. The glass electrode enabled the precise measurement of pH in biological fluids by generating a potential sensitive to the hydrogen ion activity in the solution. Arnold Beckman's introduction of the commercial pH meter in the 1930s led to the widespread adoption of pH measurement in clinical laboratories (Christian, 2020; Katz, 2025).

In carbon dioxide measurement, the Severinghaus electrode, developed by John W. Severinghaus and A. Freeman Bradley in 1958, made it possible to accurately and reliably measure pCO₂ levels in blood samples. This electrode operated through a combination of a glass pH electrode, a CO₂-permeable membrane, and a buffer solution. Carbon dioxide diffuses across the membrane, causing a change in pH, which is measured by the electrode to calculate the pCO₂ value.

In the field of oxygen measurement, the polarographic oxygen electrode (Clark electrode), developed by Leland C. Clark in 1953, enabled the direct amperometric determination of the blood's pO₂ level. This method relies on measuring the current generated by the reduction of oxygen at a platinum cathode, allowing for the continuous monitoring of blood oxygenation.

In the 1960s, the integration of pH, pCO₂, and pO₂ measurements into a single device made blood gas analysis both rapid and user-friendly. The Radiometer ABL1, introduced in 1969, pioneered this integrated approach, revolutionizing clinical use. In subsequent years, microprocessor-based devices offered advantages such as automatic calibration and error detection, reduced sample volume, and shorter measurement times. Furthermore, the measurement of electrolytes (Na⁺, K⁺, Cl⁻) and metabolites (lactate, glucose) was integrated (Clark & Clark, 1987; Severinghaus & Astrup, 1986).

Portable and cartridge-based systems (e.g., i-STAT, epoc), developed in the 2000s, made blood gas analysis usable in emergency departments, ambulances, and field conditions (Point-of-Care Testing, POCT). These systems reduced maintenance requirements, enabling calibration and reagent changes without the need for user intervention. Today's modern blood gas devices can also measure parameters like hematocrit, hemoglobin, co-oximetry (COHb, MetHb), and electrolytes (Abbasi et al., 2019; Nayak et al., 2017).

In the future, blood gas measurements are expected to be integrated into wearable devices or continuous monitoring systems, thanks to miniaturized sensors and biosensor technologies. These advancements will ensure that blood gas analysis can be reliably performed not only in clinical laboratories but also at the bedside and in mobile settings.

3. FUNDAMENTAL PARAMETERS AND MEASUREMENT PRINCIPLES

Blood gas analyses include both directly measured and derived parameters, which enable rapid and reliable clinical decision-making. The fundamental parameters provide critical information about the acid-base status, respiratory function, oxygenation, and metabolic compensation. This section explains the measurement principle, the sensors used, the calculation methods, the units, potential sources of error, and the clinical significance for each parameter.

1. pH (Hydrogen Ion Activity)

The measurement of pH is fundamentally based on the potential difference established between a glass membrane electrode and an Ag/AgCl reference electrode. The glass membrane is specifically designed to be sensitive to H⁺ ions in the solution, generating a potential that is then processed by electronic circuits and converted into a pH value using a calibration curve. This process is governed by the Nernst Equation, which relates the measured potential (E) to the pH of the solution: $E = E_0 - (2.303RT/F) \cdot \text{pH}$. The pH value is dimensionless, and clinical analyzers typically achieve an accuracy ranging from ± 0.01 to ± 0.03 pH units. Potential sources of error are critical and include electrode contamination, aging, degradation of the reference solution, temperature effects, and sample contamination. Clinically, pH is the most direct indicator of the patient's acid-base status, with a narrow normal range of 7.35–7.45. (Hahn, 1998; Kellum, 2000; Severinghaus & Astrup, 1986; Zavorsky et al., 2007).

2. pCO₂ (Partial Pressure of Carbon Dioxide)

The measurement of pCO₂ relies on the Severinghaus electrode, which is an elegant combination of a pH electrode sealed beneath a CO₂-permeable membrane. As CO₂ from the sample diffuses across this membrane, it reacts with the internal buffer solution, generating H⁺ ions and causing a measurable pH change. The device translates this pH shift back into the pCO₂ concentration, typically reported in mmHg or kPa, with an accuracy of ± 1 – 2 mmHg. Potential measurement errors often stem from membrane blockage, the presence of air bubbles in the sample, or temperature fluctuations. Clinically, pCO₂ is a direct indicator of a patient's ventilatory status: an elevated value signifies respiratory hypoventilation (inadequate breathing), while a decreased value indicates

hyperventilation (Severinghaus & Astrup, 1986; Severinghaus, Astrup, & Murray, 1998).

3. pO₂ (Partial Pressure of Oxygen)

pO₂ is measured using the Clark electrode, an amperometric sensor featuring an oxygen-permeable membrane and a platinum cathode/silver anode assembly. Oxygen molecules diffusing through the membrane are electrochemically reduced at the platinum cathode, producing an electric current that is directly proportional to the pO₂ level in the sample. This parameter is crucial for diagnosing hypoxemia, as it reflects the amount of oxygen physically dissolved in the blood. The accuracy usually ranges between ± 2 and ± 5 mmHg. Common sources of error include membrane damage, contamination of the electrode surface, and the presence of air bubbles, all of which can interfere with accurate oxygen reduction (Buerk, 2004; Clark & Clark, 1987; Suzuki, Hirakawa, Watanabe, & Kikuchi, 2001).

4. Bicarbonate (HCO₃⁻) — Derived Parameter

Bicarbonate is not measured directly but is calculated as a derived parameter that represents the core metabolic component of the acid-base system. Its concentration is computationally determined from the measured pH and pCO₂ values using the fundamental Henderson–Hasselbalch equation. Clinically, HCO₃⁻ provides essential diagnostic information: a low concentration is indicative of metabolic acidosis, suggesting a primary problem in buffering capacity, while a high concentration points toward metabolic alkalosis (Bijapur, Kudligi, & Asma, 2019; Reyes-Lazalde, Reyes-Monreal, Pérez-Bonilla, & Reyes-Luna, 2013).

5. Base Excess (BE / Standard Base Excess, SBE)

Base Excess is a quantitative metabolic parameter defined as the amount of strong acid or alkali required to restore a blood sample's pH to a neutral 7.40 at standard conditions (37 °C and pCO₂=40 mmHg). This value is calculated using complex algorithms, such as the Siggaard–Andersen nomograms implemented in the analyzer's software. BE provides a clearer indication of the non-respiratory, metabolic burden and is often used alongside the anion gap to help precisely pinpoint the etiology of a metabolic acidosis or alkalosis (Berend & Duits, 2019; Samuel, 2024).

6. Oxygen Saturation (SaO₂) and Co-Oximetry

Oxygen saturation (SaO₂) may be estimated based on the measured pO₂ using standard formulae like the Hill equation or the Severinghaus empirical equation. For a complete and accurate picture of oxygen transport, however, co-oximetry is used. This method utilizes optical measurement at multiple wavelengths to

precisely quantify the different functional forms of hemoglobin, including oxyhemoglobin (OxyHb), deoxyhemoglobin (DeoxyHb), carboxyhemoglobin (COHb), and methemoglobin (MetHb). Co-oximetry is vital because it accurately represents the total oxygen-carrying capacity of the blood and is indispensable in critical situations like carbon monoxide poisoning or methemoglobinemia (Ayres, 2012; Sinex, 1999; Zijlstra, 2005).

7. Electrolytes, Lactate, Hemoglobin, and Other Parameters

Modern blood gas analyzers have evolved into comprehensive critical care tools by integrating the measurement of several other key parameters. Electrolytes (Na^+ , K^+ , Cl^- , Ca^{2+}) are measured using ion-selective electrodes (ISE), and it is important to consider whether direct or indirect ISE methods are used. Lactate and Glucose levels are determined via enzymatic reactions where the resulting H_2O_2 (hydrogen peroxide) is measured amperometrically. Finally, parameters related to blood volume and oxygen capacity, such as Hemoglobin, Hematocrit, COHb, and MetHb, are often determined photometrically or via integrated co-oximetry (Kapoor, Srivastava, & Singh, 2014; Krasowski & Pelletier, 2024; Marija, Bernhard, & Beatrice, 2021; Toffaletti, 2001).

8. Calibration, Temperature Correction, and References

For sustained accuracy, the sensitive sensors for pH, pO_2 , pCO_2 , and the ion-selective electrodes must undergo mandatory regular calibration against known standards. Furthermore, temperature profoundly affects gas solubility and electrode response. Consequently, analyzers either apply a reference correction to standardize all readings to a normal body temperature of 37°C or, in advanced units, correct the values specifically to the patient's actual measured temperature, ensuring the physiological relevance of the results (Espadas-Torre, Telting-Diaz, & Meyerhoff, 1995; van der Weerd, Bierl, & Matysik, 2016).

4. BLOOD GAS SAMPLE TYPES AND CLINICAL USE

The types of samples used in blood gas analysis are selected based on clinical requirements, the patient's condition, and invasiveness. There are three main types of blood samples: arterial, venous, and capillary. Each has different advantages, limitations, and areas of clinical application.

1. Arterial Blood Gas (ABG)

Arterial Blood Gas (ABG) analysis is regarded as the gold standard method due to its superior accuracy, particularly in assessing oxygenation status (pO_2 , SaO_2), as arterial blood directly reflects the oxygen delivered to the tissues. It is routinely used in intensive care and emergency departments. However, ABG requires arterial puncture, making it an invasive procedure associated with risks such as pain, hematoma, infection, or arterial damage, which constitute its main

limitations. Clinically, ABG is essential for monitoring ventilation and oxygenation during anesthesia and serves as a diagnostic guide for differentiating between respiratory and metabolic disorders (Saberian et al., 2023; Taherian, Ahmadlou, Nejad, & Manian, 2025).

2. Venous Blood Gas (VBG)

Venous Blood Gas (VBG) samples offer the distinct advantage of being easier and safer to obtain, resulting in less pain and lower invasiveness compared to ABG. VBG provides reliable results for metabolic parameters such as pH and HCO_3^- . Conversely, its primary limitation is its unsuitability for accurate oxygenation assessment, as pO_2 and SaO_2 values derived from venous blood are not as reliable as arterial measurements. VBG is widely utilized for monitoring metabolic disorders, including diabetic ketoacidosis and lactic acidosis, and is often preferred for serial monitoring to minimize invasive procedures (Ekström et al., 2019; Turhan, 2025).

3. Capillary Blood Gas (CBG)

Capillary Blood Gas (CBG) analysis provides a minimally invasive option, typically drawn from the fingertip, earlobe, or heel. It is particularly valuable as an alternative to arterial sampling in pediatric patients. Similar to VBG, the main disadvantages are that oxygenation parameters are not as reliable as those obtained from ABG, and the results are highly susceptible to errors arising from incorrect sampling techniques. CBG is the first choice in neonates and children, and it may be used as an alternative to arterial samples in adults in specific, limited clinical circumstances (Ak, Kayis, Bayir, Ogun, & Kara, 2014; Eizadi-Mood, Alfred, Yaraghi, Huynh, & Moghadam, 2009; Richter et al., 2014).

The selection of blood gas sample types fundamentally depends on balancing measurement accuracy with patient comfort and clinical necessity. While ABG remains the gold standard for comprehensively evaluating oxygenation and respiratory parameters, venous and capillary samples offer crucial clinical advantages by providing reliable metabolic parameter monitoring while minimizing invasive intervention. The optimal choice must be optimized based on the patient's age, clinical severity, and specific monitoring requirements (Lee, 2025; Marik, 2014).

5. BLOOD GAS SIMULATOR-TECHNICAL AND CLINICAL EXPLANATION

Blood gas analysis is currently vital for evaluating critically ill patients. This simulator has been developed for educational purposes, offering users an interactive environment to visualize blood gas parameters and experience various clinical scenarios. The simulator is built using HTML, CSS, and JavaScript, with

the Chart.js library integrated for graphing. This structure allows both calculations and visual presentations to be provided on a single platform.

The core function of the simulation is to generate blood gas parameters based on the scenario selected by the user and to present these parameters both numerically and graphically. Users can choose the scenario based on the clinical situation (Normal, COPD, Shock, Hyperventilation, Vomiting, or Random), adjust the temperature at which the blood sample was taken, simulate conditions like sample delay and air bubbles, and add pH electrode drift in mV. These parameters allow for the digital modeling of both physiological and technical variations observed in real-world devices.

Calculations and Functions

HCO₃⁻ Calculation

In the simulation, bicarbonate concentration is calculated using the Henderson–Hasselbalch equation. The calculation is performed by the function `calcHCO3` (pH, pCO₂, tempC) and includes temperature correction. HCO₃⁻ is derived from the pCO₂ and pH values using the temperature-dependent solubility coefficient, α .

O₂ Saturation Calculation

Oxygen saturation is approximately calculated for the user from pO₂ values using the Hill equation within the function `calcO2Sat`(pO₂). This function is based on the hemoglobin-oxygen binding curve and provides the user with an estimated SaO₂ value.

Nernst Curve and pH Electrode Drift

The function `nernstSlope_mVPerpH`(tempC) is used to simulate how temperature changes affect pH measurement. Additionally, the effect of pH electrode drift is calculated with `applyPHDrift` (pH_true, drift_mV, slope_mVPerpH), and the measured pH value is updated with the added drift.

Sample Delay and Air Bubble Effect

The function simulates the delay in the blood sample reaching the laboratory by applying `DelayEffects` (pO₂, pCO₂, pH, delayMin). This function calculates the drop in pO₂, the increase in pCO₂, and the drop in pH for every minute of delay. The air bubble condition is modeled by the function `applyAirBubble` (pO₂, pCO₂, pH), which demonstrates to the user the resulting increase in pO₂, decrease in pCO₂, and increase in pH.

Interpretation Function

The measured results are clinically interpreted by the function interpreter (pH, pCO₂, HCO₃⁻). The function compares the values against normal ranges and classifies the disease as normal, primary respiratory/metabolic acidosis, alkalosis, or mixed disorders.

Graphical Visualization

- The function `updateCharts` (values, ranges, measuredPH, measuredPotential, slopeMV) draws both the bar charts and the Nernst curve.
- The main bar chart compares the pH, pCO₂, HCO₃⁻, pO₂, and SaO₂ values against their normal reference ranges.
- The Nernst curve plots the electrode potential versus pH and highlights the points for the true pH and the measured pH with the applied drift.

User Interface and Interaction

The interface, created with HTML/CSS, features sliders and dropdown menus that allow the user to modify the temperature, sample delay, air bubble status, and drift parameters. The simulation is executed by clicking the "Run Simulation" button, which updates the numerical values and the graphs.

A complete summary of all the fundamental JavaScript functions and their roles, which form the heart of the Blood Gas Simulator (HTML/JavaScript code), is provided in Table 1.

Table 1. Simulation Core Functions and Input/Output Parameters

Code / Function	Purpose	Parameters Used	Result / Output
scenarios	Initial values for clinical scenarios	pH, pCO ₂ , pO ₂	Generates initial values suitable for the scenario
calcHCO3(pH, pCO2, tempC)	HCO ₃ ⁻ calculation (Henderson-Hasselbalch)	pH, pCO ₂ , temperature	Bicarbonate concentration (mmol/L)
calcO2Sat(pO2)	O ₂ saturation calculation (Hill equation)	pO ₂	SaO ₂ (%)
nernstSlope_mVPerpH(tempC)	Nernst slope calculation	Temperature	mV/pH slope
applyPHDrift(pH_true, drift_mV, slope_mVPerpH)	Simulates electrode drift	pH_true, drift_mV, slope_mVPerpH	pH value with added drift
applyDelayEffects(pO2, pCO2, pH, delayMin)	Effect of sample delay	pO ₂ , pCO ₂ , pH, delay time	pO ₂ ↓, pCO ₂ ↑, pH↓ değerleri

applyAirBubble(pO2, pCO2, pH)	Air bubble effect	pO ₂ , pCO ₂ , pH	pO ₂ ↑, pCO ₂ ↓, pH↑ değerleri
interpret(pH, pCO2, HCO3)	Clinical interpretation	pH, pCO ₂ , HCO ₃ ⁻	Normal / acidosis/alkalosis / mixed disorder
updateCharts(values, ranges, measuredPH, measuredPotential, slopeMV)	Chart update	Measured values, reference ranges, slope, measuredPH	Bar chart and Nernst curve visualization
runSimulation()	Runs the entire simulation	User input parameters	Updates numerical values, draws charts, and generates an interpretation

The visual output of the simulation across three distinct clinical scenarios is shown in Figures 1, 2, and 3.

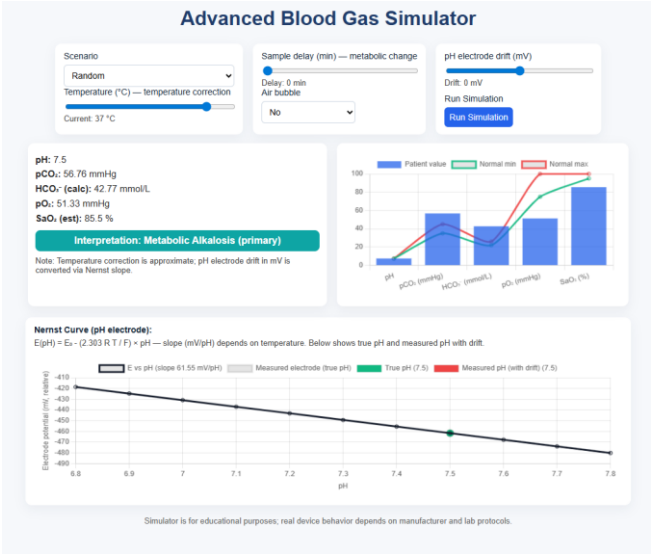


Figure 1. Screenshot of the Blood Gas Simulator showing the result for a Randomly Generated Scenario.

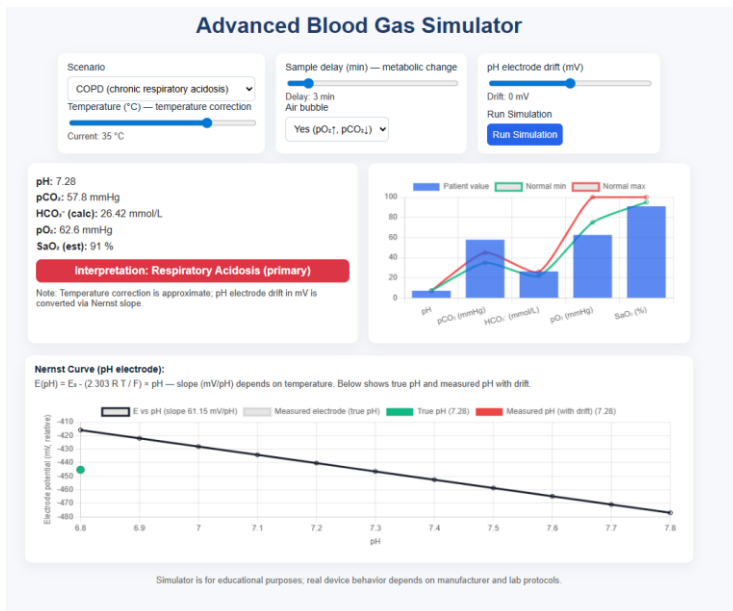


Figure 2. Simulation output demonstrating Chronic Respiratory Acidosis (COPD).

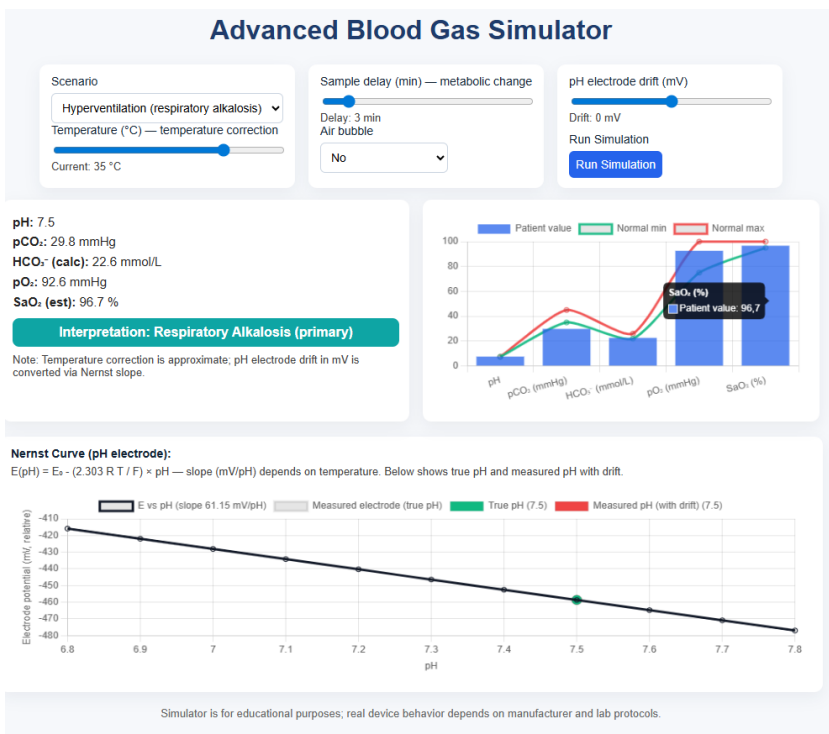


Figure 3. Simulation results for a Hyperventilation scenario (Respiratory Alkalosis).

Table 2 presents an example of clinical output and interpretation for all core scenarios within the Blood Gas Simulator, illustrating the different acid-base disturbances modeled by the platform.

Table 2. Clinical Output and Interpretation of Simulated Blood Gas Scenarios

Scenario	pH (measured)	pCO ₂ (mmHg)	HCO ₃ ⁻ (mmol/L)	pO ₂ (mmHg)	SaO ₂ (%)	Interpretation	Explanation
Normal	7.40	40	24	90	97	Normal	pH, pCO ₂ , and HCO ₃ ⁻ They are within the normal range; metabolic and respiratory balance is achieved.
COPD	7.28	60	26	60	86	Respiratory Acidosis (primary)	High pCO ₂ and low pH; chronic respiratory acidosis. HCO ₃ ⁻ is slightly elevated, indicating compensation.
Shock	7.25	30	16	70	91	Metabolic Acidosis (primary)	Low HCO ₃ ⁻ and pH; metabolic acidosis; pCO ₂ is low, indicating the start of compensation.
Hyperventilation	7.52	28	22	95	98	Respiratory Alkalosis (primary)	Low pCO ₂ High pH and respiratory alkalosis due to rapid breathing. HCO ₃ ⁻
Vomit	7.50	45	30	95	98	Metabolic Alkalosis (primary)	It's normal. Metabolic alkalosis following vomiting; pH is high, HCO ₃ ⁻

							is elevated; pCO ₂
							It is normal or slightly high.
							In random scenarios, pH, pCO ₂ , and HCO ₃ -
Random	7.33	52	20	82	94	Mixed / Compensated disorder	May show different combined disorders, with a high probability of a mixed disorder.

6. TECHNICAL AND CLINICAL ANALYSIS OF SCENARIOS

Normal Scenario

- Reference values are used in the simulation.
- pH is within the 7.35–7.45 range, pCO₂ 35–45 mmHg, HCO₃⁻ 22–26 mmol/L.
- Clinically, the patient is stable; there is no compensation or metabolic/respiratory disorder.

COPD (Chronic Respiratory Acidosis)

- pCO₂ is high and pH is low.
- HCO₃⁻ is slightly high, indicating chronic compensation by the kidneys.
- When combined with the simulation's delay and drift parameters, the values may deteriorate further.

Shock (Metabolic Acidosis)

- pH is low, HCO₃⁻ is low, and pCO₂ is normal or slightly decreased.
- This condition is a classic example of hypoperfusion and lactic acidosis.
- The simulation of delay or air bubbles creates significant effects on pO₂ and pH.

Hyperventilation (Respiratory Alkalosis)

- Rapid breathing leads to low $p\text{CO}_2$, and pH rises.
- HCO_3^- is normal; compensation is absent or slight.
- If drift is added, the measured pH may appear even higher.

Vomit (Metabolic Alkalosis)

- After vomiting, HCO_3^- rises, and pH rises.
- $p\text{CO}_2$ is normal; the lungs have not yet compensated.
- The simulation can visually estimate clinical electrolyte changes.

Random

- These scenarios are for educational purposes; they are ideal for testing mixed disorders and measurement errors.
- Clinical interpretation becomes more complex in combinations with drift, delay, and temperature changes.

CONCLUSION

Blood gas analysis (arterial blood gas) remains the gold standard method for the rapid and precise assessment of the acid-base balance, respiratory function, and metabolic status in patients facing critical illness. This crucial role has been solidified through extensive technological progress, advancing from early 20th-century manual titration techniques to the highly accurate, integrated electrode systems (such as the Clark and Severinghaus electrodes) used in modern biomedical engineering. The reliability of the test necessitates careful consideration of sample type, balancing the diagnostic advantage of arterial samples for assessing oxygenation with the reduced invasiveness offered by venous and capillary samples for monitoring metabolic changes.

The Blood Gas Simulator platform represents a crucial educational and research asset, utilizing web technologies (HTML, JavaScript) and the Chart.js library to offer a dynamic, visual learning experience. The primary strength and contribution of this tool lie in its sophisticated integration of realistic technical and physiological error modeling. Users can simulate critical pre-analytical variables, including the consequences of delayed sample arrival (leading to metabolic shifts like lactic acid buildup), the introduction of air bubbles, temperature inconsistencies, and instrumental faults such as electrode potential drift. By guiding users through the interpretation of parameters derived from core biophysical laws—like the Henderson–Hasselbalch equation for bicarbonate and the Nernst equation for electrode potential—the simulator effectively reinforces the diagnostic process for conditions such as primary respiratory acidosis, metabolic alkalosis, or mixed acid-base disturbances.

In conclusion, this platform successfully bridges the divide between abstract theoretical knowledge and concrete practical application. It provides students and clinical professionals with a robust, accessible, and interactive environment necessary to master the quantitative analysis of blood gas values. It also helps them thoroughly understand how crucial pre-analytical variables can significantly impact patient diagnosis and subsequent treatment decisions.

REFERENCES

- Toffaletti, J. G., & Rackley, C. R. (2021). Blood gases and critical care testing: Physiology, clinical interpretations, and laboratory applications. Academic Press.
- Papoff, P., Moretti, C., & Nenna, R. (2024). Blood gas assessment and oximetry. In ERS handbook of paediatric respiratory medicine (pp. 121–128). European Respiratory Society.
- Rabi, Y., Kowal, D., & Ambalavanan, N. (2016). Blood gases: Technical aspects and interpretation. In Assisted ventilation of the neonate: An evidence-based approach to newborn respiratory care (pp. 80–96).
- Hennessey, I. A., & Japp, A. G. (2015). Arterial blood gases made easy. Elsevier Health Sciences.
- Severinghaus, J. W. (2014). History of respiratory gas monitoring in anesthesia. In The wondrous story of anesthesia (pp. 745–759). Springer New York.
- Severinghaus, J. W. (1986). Historical development of oxygenation monitoring. In Pulse oximetry (pp. 1–18). Springer London.
- Katz, E. (2025). Evolution of electrochemistry: Key figures and their contributions. Newman Springs Publishing.
- Christian, G. D. (2020). Evolution of analytical sciences in the United States: A historical account. *Annual Review of Analytical Chemistry*, 13(1), 475–496.
- Severinghaus, J. W., & Bradley, A. F. (1958). Electrodes for blood pO₂ and pCO₂ determination. *Journal of Applied Physiology*, 13(3), 515–520.
- Severinghaus, J. W., & Astrup, P. B. (1986). History of blood gas analysis. V. Oxygen measurement. *Journal of Clinical Monitoring*, 2(3), 174–189.
- Clark, L. C., Jr., & Clark, E. W. (1987). A personalized history of the Clark oxygen electrode. *International Anesthesiology Clinics*, 25(3), 1–29.
- Nayak, S., Blumenfeld, N. R., Laksanasopin, T., & Sia, S. K. (2017). Point-of-care diagnostics: Recent developments in a connected age. *Analytical Chemistry*, 89(1), 102–123.
- Abbasi, U., Chowdhury, P., Subramaniam, S., Jain, P., Muthe, N., Sheikh, F., ... & Kumaran, V. (2019). A cartridge-based point-of-care device for complete blood count. *Scientific Reports*, 9(1), 18583.
- Kellum, J. A. (2000). Determinants of blood pH in health and disease. *Critical Care*, 4(1), 6.

- Zavorsky, G. S., Cao, J., Mayo, N. E., Gabbay, R., & Murias, J. M. (2007). Arterial versus capillary blood gases: A meta-analysis. *Respiratory Physiology & Neurobiology*, 155(3), 268–279.
- Hahn, C. E. W. (1998). Electrochemical analysis of clinical blood gases, gases and vapours. *The Analyst*, 123(6), 57R–86R.
- Severinghaus, J. W., Astrup, P., & Murray, J. F. (1998). Blood gas analysis and critical care medicine. *American Journal of Respiratory and Critical Care Medicine*, 157(4), S114–S122.
- Suzuki, H., Hirakawa, T., Watanabe, I., & Kikuchi, Y. (2001). Determination of blood pO₂ using a micromachined Clark-type oxygen electrode. *Analytica Chimica Acta*, 431(2), 249–259.
- Buerk, D. G. (2004). Measuring tissue PO₂ with microelectrodes. In *Methods in enzymology* (Vol. 381, pp. 665–690). Academic Press.
- Reyes-Lazalde, A., Reyes-Monreal, M., Pérez-Bonilla, M. E., & Reyes-Luna, R. (2013). Development of a software to determine disturbances in the acid-base balance in human blood. *Revista Mexicana de Ingeniería Biomédica*, 34(2), 175–191.
- Bijapur, M. B., Kudligi, N. A., & Asma, S. (2019). Central venous blood gas analysis: An alternative to arterial blood gas analysis for pH, PCO₂, bicarbonate, sodium, potassium, and chloride in intensive care unit patients. *Indian Journal of Critical Care Medicine*, 23(6), 258.
- Berend, K., & Duits, A. J. (2019). The role of the clinical laboratory in diagnosing acid–base disorders. *Critical Reviews in Clinical Laboratory Sciences*, 56(3), 147–169.
- Samuel, T. R. (2024). Scientific breakthroughs from Siggaard–Andersen nomogram to the latest developed graphical tool representation for arterial blood gas (ABG) interpretation. *Frontiers in Health Informatics*, 13(3).
- Sinex, J. E. (1999). Pulse oximetry: Principles and limitations. *The American Journal of Emergency Medicine*, 17(1), 59–66.
- Zijlstra, W. G. (2005). Clinical assessment of oxygen transport-related quantities. *Clinical Chemistry*, 51(2), 291–292.
- Ayres, D. A. (2012). Pulse oximetry and CO-oximetry. In *Advanced monitoring and procedures for small animal emergency and critical care* (pp. 274–285).
- Toffaletti, J. G. (2001). Special topics in diagnostic testing: Blood gases and electrolytes. *American Association for Clinical Chemistry*.

- Kapoor, D., Srivastava, M., & Singh, P. (2014). Point of care blood gases with electrolytes and lactates in adult emergencies. *International Journal of Critical Illness and Injury Science*, 4(3), 216–222.
- Marija, K., Bernhard, K. F., & Beatrice, L. K. (2021). Blood-gas vs. central-laboratory analyzers: Interchangeability and reference intervals for sodium, potassium, glucose, lactate and hemoglobin. *Heliyon*, 7(11).
- Krasowski, M. D., & Pelletier, D. (2024). Electrolytes, blood gases, renal function. In *Clinical pathology board review e-book* (p. 109).
- Espadas-Torre, C., Telting-Diaz, M., & Meyerhoff, M. (1995). Electrochemical sensors for the continuous monitoring of blood gases and electrolytes. *The Electrochemical Society Interface*, 4(1), 41–47.
- van der Weerd, B., Bierl, R., & Matysik, F. M. (2016). Trends in electrochemical sensing of blood gases. In *Trends in bioelectroanalysis* (pp. 263–280).
- Saberian, L., Sharif, M., Aarabi, M., Broumand, B., Shafiee, M. A., & Shafiee, M. (2023). Arterial versus venous blood gas analysis: Comparisons, appropriateness, and alternatives in different acid/base clinical settings—A systematic review. *Cureus*, 15(7).
- Taherian, M., Ahmadlou, M., Nejad, A. E., & Manian, M. (2025). Blood gas analyzers and methodology. In *Up-to-date approach to blood gas analysis: Clues to diagnosis and treatment* (p. 7).
- Ekström, M., Engblom, A., Ilic, A., Holthius, N., Nordström, P., & Vaara, I. (2019). Calculated arterial blood gas values from a venous sample and pulse oximetry: Clinical validation. *PLoS One*, 14(4), e0215413.
- Turhan, S. (2025). Blood gas monitoring in anesthesia applications. In *Up-to-date approach to blood gas analysis: Clues to diagnosis and treatment* (p. 31).
- Eizadi-Mood, N., Alfred, S., Yaraghi, A., Huynh, C., & Moghadam, A. S. (2009). Comparison of arterial and capillary blood gas values in poisoning department assessment. *Human & Experimental Toxicology*, 28(10), 665–670.
- Ak, A., Kayis, S. A., Bayir, A., Ogun, C. O., & Kara, H. (2014). The value of capillary blood gas measurement in the intensive care unit. *Journal of Chest Diseases and Critical Care Medicine*, 1(1), 24–30.
- Richter, S., Kerry, C., Hassan, N., Chari, A., Lunn, D., & Nickol, A. (2014). Capillary blood gas as a substitute for arterial blood gas: A meta-analysis. *British Journal of Hospital Medicine*, 75(3), 136–A7.

- Lee, G. R. (2025). Clinical applications of blood gas analysis: A comparative review of arterial and venous blood gas monitoring in critical care. *Acute and Critical Care*, 40(2), 153.
- Marik, P. E. (2014). Arterial blood gas analysis. In *Evidence-based critical care* (pp. 329–347). Springer International Publishing.

CHAPTER 7

Experimental PID Tuning for Magnetometer-Based Heading Control of an Autonomous Mobile Robot

Osman Ünal¹

¹ Assist. Prof. Dr., Sakarya University of Applied Sciences, Maritime Vocational School,
ORCID: 0000-0003-1101-6561

1 INTRODUCTION

Autonomous mobile robots have been widely used in industrial, military, and service-oriented applications due to their ability to operate without human intervention (Victorpaol et al., 2017; Zghair & Al-Araji, 2021). Among these systems, differential wheeled mobile robots stand out because of their simple mechanical structure, high maneuverability, and low production cost (Zhang et al., 2023). One of the most critical problems in such robots is accurate heading control, which directly affects trajectory tracking performance and overall navigation accuracy. In this study, magnetometer-based heading estimation provides an effective and low-cost solution for determining the robot's orientation in outdoor and indoor environments.

Proportional–Integral–Derivative (PID) controllers are among the most widely used control strategies for mobile robot heading control due to their simple structure, ease of implementation, low computational cost, and satisfactory performance in linear and nonlinear systems (Tyagunov et al., 2018, Fahmizal et al., 2019). However, the performance of a PID controller strongly depends on the proper selection of its three gain parameters (Ali et al., 2025): proportional (K_p), integral (K_i), and derivative (K_d). Inappropriate tuning of these parameters may lead to excessive overshoot, long settling time, oscillations, or steady-state errors. Therefore, determining the optimal PID gains is a crucial issue for achieving high-precision heading control.

In the literature, various classical and intelligent tuning methods have been proposed for PID parameter optimization, such as Ziegler–Nichols (Åström & Hägglund, 2004), Cohen–Coon (Joseph & Olaiya, 2017), relay-based tuning (Nunes et al., 2014), genetic algorithms (Meng & Song, 2007), particle swarm optimization (Solihin et al., 2011), fuzzy logic (Visioli, 2001), and neural network-based approaches (Hernández-Alvarado et al., 2016). Although these methods can provide automatic or near-optimal solutions, they often suffer from several drawbacks such as high computational complexity, requirement of accurate system modeling, sensitivity to noise, long convergence times, and difficulties in real-time implementation on low-cost embedded platforms. Moreover, intelligent optimization-based methods generally require excessive experimental data and significant processing power, which limits their applicability in practical mobile robot systems.

In contrast, experimental PID tuning remains a reliable and practical approach, especially for real-world robotic platforms where system dynamics are highly nonlinear and difficult to model accurately. By directly observing the system response under real operating conditions, experimental tuning allows the controller parameters to be adjusted according to actual performance metrics such as rise time, overshoot, steady-state error, and settling time. This approach

eliminates the dependency on mathematical models and reduces the risk of performance degradation caused by modeling uncertainties.

In this study, the optimal PID parameters for the magnetometer-based heading control of a differential wheeled autonomous mobile robot are determined through a systematic experimental procedure. After defining the minimum, maximum, and step size for each of the K_p , K_i , and K_d coefficients, a comprehensive experimental dataset is generated using three nested loops that cover all possible parameter combinations within the specified ranges. For each parameter set, the robot's ability to reach and maintain the desired heading angle is evaluated, and its performance is quantitatively analyzed. Based on these experimental results, the PID parameter combination that yields the best heading tracking performance is identified.

The main contribution of this work is the proposal of a structured and practical experimental tuning framework for magnetometer-based PID heading control, which can be directly applied to low-cost autonomous mobile robots, along with the complete Arduino implementation required to obtain this framework being provided at the end of the study. The findings of this study demonstrate that experimentally tuned PID controllers can provide robust, efficient, and reliable heading control without the need for complex optimization algorithms or precise system modeling.

2 EXPERIMENTAL STUDY

This study experimentally investigates the tuning of a PID controller for magnetometer-based heading control of an autonomous mobile robot. The experimental setup, hardware architecture, communication interfaces, and the systematic PID parameter search methodology are described in detail in this section. The autonomous mobile robot used in the experiments is illustrated in Figure 1.

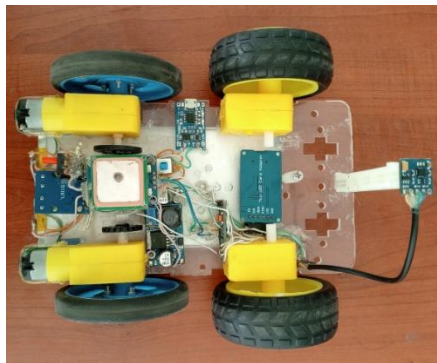


Figure 1. Autonomous mobile robot used for experimental validation.

2.1 Hardware Configuration

The control and data acquisition system of the autonomous mobile robot is based on an Arduino Pro Mini microcontroller, which provides sufficient computational capability, low power consumption, and compact size suitable for embedded robotic applications. Heading information is obtained using a QMC5883L three-axis digital magnetometer, which measures the Earth's magnetic field and enables absolute heading estimation. The magnetometer is interfaced with the microcontroller via the I²C communication protocol using the SDA and SCL lines connected to the A4 and A5 pins of the Arduino Pro Mini, respectively.

Motor actuation is performed using an L9110 dual-channel DC motor driver. Two DC motors equipped with 1/48 reduction gearboxes are used to drive the left and right wheels of the mobile robot. The PWM control signals for the motors are generated by the Arduino Pro Mini and applied to the driver through digital pin D3 for the left wheel and pin D9 for the right wheel, enabling differential drive motion. The gear-reduction mechanism increases torque while maintaining stable low-speed performance, which is essential for precise heading correction.

For experimental data logging, a microSD card module is integrated into the system. The module communicates with the Arduino Pro Mini via the SPI protocol using pins D10–D13. All experimental results, including the PID parameter combinations and corresponding accumulated heading errors, are stored on the microSD card for off-line evaluation and performance comparison. This logging capability ensures reliable and non-volatile storage of experimental results.

2.2 Experimental PID Tuning Strategy

The PID tuning process is conducted through a systematic parameter sweep covering predefined ranges of proportional, integral, and derivative gains. The minimum, maximum, and step values of the controller gains are defined as follows:

- $K_p \in [2.4, 21.6]$ with a step size of 4.8, resulting in 5 distinct values,
- $K_i \in [0, 27]$ with a step size of 5.4, resulting in 6 distinct values,
- $K_d \in [0, 0.1]$ with a step size of 0.05, resulting in 3 distinct values.

Accordingly, a total of $5 \times 6 \times 3 = 90$ unique PID gain combinations are tested. For each combination, the robot performs an independent experimental run using nested loops implemented in the Arduino firmware.

Before each test, the motors are stopped to eliminate residual dynamics. The current heading of the robot is then measured using the magnetometer, and a new

reference heading is defined by adding 180° to the measured value. This ensures that each experiment evaluates the controller's ability to correct a large angular deviation. Angular normalization is applied to keep the setpoint within the range of 0°–360°.

Each PID parameter set is tested for a fixed duration of 5 seconds. During this interval, real-time heading measurements are continuously obtained from the magnetometer, and the angular error is computed as the shortest signed angular distance between the measured heading and the setpoint. The error is constrained to the interval $[-180^\circ, 180^\circ]$ to avoid discontinuities.

The accumulated absolute heading error is calculated over the entire 5-second interval using Equation (1):

$$J = \sum_{k=1}^N |e(k)| \quad (1)$$

where $e(k)$ is the heading error at the k -th sampling instant. This accumulated absolute error is used as the primary performance index for evaluating each PID combination.

2.3 Real-Time PID Control Implementation

The PID controller is implemented in discrete time with a nominal sampling period of approximately 30 ms. The control law is defined as in Equation (2).

$$u(t) = K_p e(t) + K_i \int_t^{t+1} e(t) dt + K_d \times \frac{de(t)}{dt} \quad (2)$$

To prevent integral windup, the integral term is limited within ± 50 . The control signal u is used to adjust the PWM duty cycles of the left and right motors in a differential manner. Specifically, the right motor speed is increased by $+u$, while the left motor speed is decreased by $-u$, ensuring proportional rotational correction toward the desired heading. The PWM signals are constrained within the valid range of 0–255.

For each PID combination, the total accumulated absolute error is recorded on the microSD card. Following the completion of all 90 experiments, the PID parameters yielding the minimum accumulated error are automatically selected as the optimal gains. These optimal gains are then permanently applied to the controller, and continuous real-time heading control is performed using the selected parameters.

2.4 Data Logging and Performance Evaluation

All experimental results are stored in a text file on the microSD card, including each tested K_p , K_i , and K_d combination and the corresponding accumulated

absolute error. After the full parameter sweep, the best-performing gain set and its error value are also recorded. During closed-loop operation with the optimal gains, instantaneous heading error and control signal values are continuously logged for further post-processing and dynamic response analysis. The complete Arduino code used for the automated PID tuning procedure, real-time control implementation, and data logging is provided in Appendix A. This code fully implements the nested-loop parameter sweep, setpoint generation, real-time PID computation, and SD-card-based data storage used throughout the experiments described in this study.

3 RESULTS AND DISCUSSION

This section presents and discusses the experimental results obtained from the systematic PID parameter search described in the previous sections. A total of 90 distinct (K_p , K_i , and K_d) combinations were evaluated based on the accumulated absolute heading error over a 5-second test interval for each case. A video link providing an illustrative sample of the experimental runs performed with the autonomous robot has been included in Appendix B. The complete set of experimental results is reported in Table 1.

Table 1. Results of 90 tested PID gain combinations.

Kp	Ki	Kd	Sum Err (°)	Kp	Ki	Kd	Sum Err (°)	Kp	Ki	Kd	Sum Err (°)
2.4	0	0	6859	7.2	21.6	0	9485	16.8	10.8	0	9120
2.4	0	0.05	9386	7.2	21.6	0.05	8822	16.8	10.8	0.05	7648
2.4	0	0.1	7459	7.2	21.6	0.1	10642	16.8	10.8	0.1	7846
2.4	5.4	0	9727	7.2	27	0	11119	16.8	16.2	0	8772
2.4	5.4	0.05	9687	7.2	27	0.05	12704	16.8	16.2	0.05	7950
2.4	5.4	0.1	16060	7.2	27	0.1	12158	16.8	16.2	0.1	7648
2.4	10.8	0	11057	12	0	0	5506	16.8	21.6	0	8144
2.4	10.8	0.05	11911	12	0	0.05	4569	16.8	21.6	0.05	7760
2.4	10.8	0.1	12490	12	0	0.1	5444	16.8	21.6	0.1	8556
2.4	16.2	0	18391	12	5.4	0	7749	16.8	27	0	8036
2.4	16.2	0.05	18437	12	5.4	0.05	7088	16.8	27	0.05	8981
2.4	16.2	0.1	21186	12	5.4	0.1	6933	16.8	27	0.1	7796
2.4	21.6	0	19048	12	10.8	0	7454	21.6	0	0	5825
2.4	21.6	0.05	21663	12	10.8	0.05	8967	21.6	0	0.05	5181
2.4	21.6	0.1	20922	12	10.8	0.1	7739	21.6	0	0.1	8418
2.4	27	0	20236	12	16.2	0	8601	21.6	5.4	0	7163
2.4	27	0.05	21155	12	16.2	0.05	7994	21.6	5.4	0.05	7108
2.4	27	0.1	20762	12	16.2	0.1	7667	21.6	5.4	0.1	6752
7.2	0	0	5883	12	21.6	0	7349	21.6	10.8	0	11782
7.2	0	0.05	5764	12	21.6	0.05	8397	21.6	10.8	0.05	10140

7.2	0	0.1	6067	12	21.6	0.1	11569	21.6	10.8	0.1	8068
7.2	5.4	0	9786	12	27	0	6970	21.6	16.2	0	7995
7.2	5.4	0.05	8078	12	27	0.05	8833	21.6	16.2	0.05	9047
7.2	5.4	0.1	9113	12	27	0.1	8023	21.6	16.2	0.1	8325
7.2	10.8	0	8072	16.8	0	0	5365	21.6	21.6	0	8899
7.2	10.8	0.05	8108	16.8	0	0.05	6475	21.6	21.6	0.05	7814
7.2	10.8	0.1	6847	16.8	0	0.1	5342	21.6	21.6	0.1	8911
7.2	16.2	0	8809	16.8	5.4	0	7344	21.6	27	0	8279
7.2	16.2	0.05	7829	16.8	5.4	0.05	6795	21.6	27	0.05	7858
7.2	16.2	0.1	8894	16.8	5.4	0.1	7365	21.6	27	0.1	8771

3.1 Analysis of the PID Parameter Dataset

Table 1 lists the tested proportional, integral, and derivative gain values along with their corresponding accumulated absolute heading errors. The results indicate a strong dependence of the heading control performance on the selected PID parameters. Among all tested combinations, the minimum accumulated absolute error was obtained for $K_p=12$, $K_i=0$, and $K_d=0.05$ with a total error of 4569; this result is highlighted in green in Table 1. This result indicates that a moderate proportional gain combined with a small derivative action and zero integral contribution provides the most effective heading correction for the considered robotic platform and experimental conditions. The absence of the integral term in the optimal parameter set suggests that steady-state error was already sufficiently minimized by the proportional–derivative structure, while the integral term tended to introduce excessive accumulation and potential instability. In contrast, the worst performance was observed for $K_p=2.4$, $K_i=21.6$, and $K_d=0.05$ which resulted in a maximum accumulated absolute error of 21663, as highlighted in red in Table 1. This configuration is characterized by an insufficient proportional gain combined with a very large integral gain. Such a gain structure leads to slow initial response and excessive integral windup, which severely degrades the heading correction performance. Overall, the results in Table 1 reveal a clear performance trend:

Low proportional gains fail to generate sufficient corrective torque, leading to large heading errors.

Excessive integral gains increase oscillatory behavior and destabilize the system.

The inclusion of a small derivative gain significantly improves transient performance by damping oscillations and reducing overshoot.

These observations confirm that stable and accurate heading control cannot be achieved through arbitrary PID parameter selection and that systematic experimental tuning is essential.

3.2 Time-Domain Error Behavior for Best and Worst Parameters

The time-domain evolution of the heading error for the best and worst PID parameter sets is illustrated in Figure 2. In this figure, the blue curve represents the error response corresponding to the optimal gains ($Kp=12$, $Ki=0$, and $Kd=0.05$), while the red curve shows the response obtained with the worst-performing gains ($Kp=2.4$, $Ki=21.6$, and $Kd=0.05$).

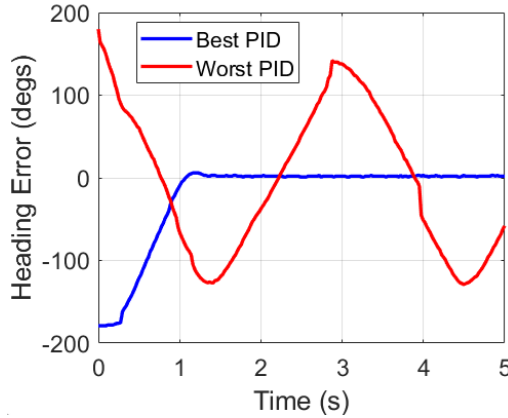


Figure 2. Heading error responses for optimal and worst PID parameters.

For the optimal parameter set, the error decreases steadily and consistently toward zero. The response exhibits a well-damped behavior with no significant oscillations, indicating that the robot is able to achieve the reference heading in a stable and robust manner. The smooth reduction in error confirms the effectiveness of the selected proportional-derivative structure in compensating for both the steady-state and transient components of the heading deviation.

In contrast, the worst-performing parameter set produces a highly oscillatory and unstable error response, as shown by the red curve in Figure 2. The error fails to converge to zero within the 5-second test interval and instead fluctuates around large values. This behavior can be directly attributed to the excessively high integral gain combined with a weak proportional action, which leads to integral windup and delayed corrective response. As a result, the robot is unable to align itself with the reference heading and exhibits persistent instability.

The stark contrast between the blue and red curves in Figure 2 clearly illustrates how improper PID tuning can lead to severe performance degradation in real-world robotic systems.

3.3 Discussion on Experimental PID Optimization

The experimental results clearly demonstrate that the performance of magnetometer-based heading control is strongly dependent on the PID gain

selection. The fact that the accumulated error varies from 4569 to 21663 across the tested parameter space shows that poor tuning can result in nearly five times larger tracking error compared to the optimal configuration. Moreover, the superiority of the experimentally optimized gains over randomly selected values highlights the importance of data-driven, systematic tuning approaches in embedded robotic control. Due to nonlinearities in motor dynamics, friction, backlash in gear mechanisms, sensor noise in the magnetometer, and unmodeled disturbances, analytical tuning rules alone are insufficient to guarantee optimal performance. The experimental search strategy adopted in this study effectively captures these real-world effects and yields controller gains that ensure both stability and high tracking accuracy.

3.4 Key Findings

The main findings derived from the presented results can be summarized as follows:

- The optimal PID gains were experimentally determined as $K_p=12$, $K_i=0$, and $K_d=0.05$, yielding the minimum accumulated absolute error.
- Excessive integral gain combined with low proportional gain leads to large tracking errors and unstable behavior.
- The time-domain responses in Figure 2 clearly demonstrate the difference between stable and unstable heading control.
- The obtained results strongly emphasize the critical importance of experimental PID parameter optimization for reliable autonomous mobile robot navigation.

4 CONCLUSION

This study presented an experimental investigation of PID tuning for magnetometer-based heading control of an autonomous mobile robot. A systematic and fully automated experimental framework was developed using an embedded microcontroller platform, a digital magnetometer sensor, differential DC motor actuation, and SD-card-based data logging. The proposed approach enabled the evaluation of 90 different PID parameter combinations under identical experimental conditions, ensuring a fair and repeatable performance comparison.

The experimental results demonstrated that the heading control performance is highly sensitive to the selection of PID gains. The accumulated absolute heading error varied significantly across the tested parameter space, confirming that improper gain selection can severely degrade tracking performance and

system stability. The optimal gain set ($Kp=12$, $Ki=0$, and $Kd=0.05$) achieved the minimum total error, yielding fast, smooth, and well-damped convergence to the reference heading. In contrast, the worst-performing parameter set resulted in large oscillations, instability, and failure to track the desired orientation.

The time-domain error responses further emphasized the impact of correct tuning. While the optimally tuned controller exhibited stable and smoothly error reduction, the poorly tuned controller showed persistent oscillations and divergence from the reference heading. These findings clearly indicate that integral action, when improperly selected, can introduce instability through windup effects, whereas a properly balanced proportional–derivative structure ensures both fast transient response and stable steady-state behavior.

Overall, this work experimentally verifies that data-driven PID optimization is essential for reliable magnetometer-based heading control in autonomous mobile robots. Unlike purely theoretical tuning rules, the proposed experimental approach effectively captures real-world nonlinearities such as motor dynamics, friction, sensor noise, and unmodeled disturbances. The presented methodology is simple, low-cost, and easily transferable to other embedded robotic platforms and navigation tasks.

Future work may focus on modeling the relationship between the PID gains and the corresponding accumulated absolute error in order to estimate intermediate parameter values that are not directly included in the experimental dataset but are expected to yield lower tracking errors through interpolation and optimization of the experimental response surface.

REFERENCES

- Ali, I. S., Areir, M. M., & Abulajras, E. E. (2025). Optimization of PID controller using PSO algorithm for a gimbal system. *Humanities & Natural Sciences Journal*, 6(4), 340-348.
- Åström, K. J., & Hägglund, T. (2004). Revisiting the Ziegler–Nichols step response method for PID control. *Journal of process control*, 14(6), 635-650.
- Fahmizal, F., Priyatmoko, A., Apriaskar, E., & Mayub, A. (2019, October). Heading control on differential drive wheeled mobile robot with odometry for tracking problem. In *2019 International Conference on Advanced Mechatronics, Intelligent Manufacture and Industrial Automation (ICAMIMIA)* (pp. 47-52). IEEE.
- Hernández-Alvarado, R., García-Valdovinos, L. G., Salgado-Jiménez, T., Gómez-Espinosa, A., & Fonseca-Navarro, F. (2016). Neural network-based self-tuning PID control for underwater vehicles. *sensors*, 16(9), 1429.
- Joseph, E. A., & Olaiya, O. O. (2017). Cohen-coon PID tuning method; A better option to Ziegler Nichols-PID tuning method. *ENginnering Research*, 2(11), 141-145.
- Meng, X., & Song, B. (2007, August). Fast genetic algorithms used for PID parameter optimization. In *2007 IEEE International Conference on Automation and Logistics* (pp. 2144-2148). IEEE.
- Nunes, D. C., Pinto, J. E., Fonseca, D. G., Maitelli, A. L., & Araújo, F. M. (2014, September). Relay based PID auto-tuning applied to a multivariable level control system. In *2014 11th International Conference on Informatics in Control, Automation and Robotics (ICINCO)* (Vol. 1, pp. 741-748). IEEE.
- Solihin, M. I., Tack, L. F., & Kean, M. L. (2011). Tuning of PID controller using particle swarm optimization (PSO). *International Journal on Advanced Science, Engineering and Information Technology*, 1(4), 458-461.
- Tyagunov, O. A., Teplov, M. A., & Slepynina, E. A. (2018, January). Adjustment of parameters of autonomous mobile robot speed and heading stabilization subsystem controllers on the basis of pareto-optimal solution construction process. In *2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EConRus)* (pp. 1011-1016). IEEE.
- Victerpaul, P., Saravanan, D., Janakiraman, S., & Pradeep, J. (2017). Path planning of autonomous mobile robots: A survey and comparison. *Journal of Advanced Research in Dynamical and Control Systems*, 9(12), 1535-1565.
- Visioli, A. (2001). Tuning of PID controllers with fuzzy logic. *IEE Proceedings-Control Theory and Applications*, 148(1), 1-8.

Zghair, N. A. K., & Al-Araji, A. S. (2021). A one decade survey of autonomous mobile robot systems. *International Journal of Electrical and Computer Engineering*, 11(6), 4891.

Zhang, P., Chen, Q., & Chen, Y. (2023). Prescribed-Time Fractional Order Control of DWR via Time-Varying Scaling Function. *International Journal of Rotating Machinery*, 2023(1), 6622657.

APPENDICES

Appendix A: Arduino code.

```
#include <QMC5883L.h>
#include <Wire.h>
#include <SPI.h>
#include <SD.h>
QMC5883L compass;
unsigned long now = 0, prev_time = 0;
float dt = 0.03;
float Kp = 10, Ki = 0, Kd = 0;
float integral = 0, derivative = 0, u = 0;
int Setpoint = 0, Input = 0, Error = 0, prev_error = 0;
int S_R = 0, S_L = 0;
const float Kp_min = 2.4;
const float Kp_max = 21.6;
const float Kp_step = 4.8;
const float Ki_min = 0.0;
const float Ki_max = 27.0;
const float Ki_step = 5.4;
const float Kd_min = 0.0;
const float Kd_max = 0.1;
const float Kd_step = 0.05;
const float I_LIMIT = 50.0;
void setup() {
```

```

Serial.begin(9600);
Wire.begin();
compass.init();
compass.setOversampling(512);
compass.setRange(2);
compass.setSamplingRate(50);
SD.begin(10);
pinMode(3, OUTPUT);
pinMode(9, OUTPUT);
analogWrite(3, 255);
analogWrite(9, 255);
if (SD.exists("datalog.txt")) SD.remove("datalog.txt");
}

int readHeading() {
    int h = compass.readHeading();
    if (h < 0) h += 360;
    if (h >= 360) h -= 360;
    return h;
}

void stopMotors() {
    analogWrite(3, 255);
    analogWrite(9, 255);
}

void writeSummaryToSD(float kp, float ki, float kd, float absErr) {
    File dataFile = SD.open("datalog.txt", FILE_WRITE);
    dataFile.print("COMBO ");
    dataFile.print(kp, 2); dataFile.print(" ");
    dataFile.print(ki, 2); dataFile.print(" ");
    dataFile.print(kd, 2); dataFile.print(" ");

```



```

dataFile.print("ABSERR ");
dataFile.println(absErr, 3);
dataFile.close();
}

void writeBestToSD(float kp, float ki, float kd, float absErr) {
  File dataFile = SD.open("datalog.txt", FILE_WRITE);
  dataFile.print("BEST Kp "); dataFile.print(kp, 4); dataFile.print(" ");
  dataFile.print("Ki "); dataFile.print(ki, 4); dataFile.print(" ");
  dataFile.print("Kd "); dataFile.print(kd, 4); dataFile.print(" ");
  dataFile.print("ABSERR "); dataFile.println(absErr, 4);
  dataFile.close();
}

void writeRunErrToSD(int err, float uval) {
  File dataFile = SD.open("datalog.txt", FILE_WRITE);
  dataFile.print("RUN_ERR ");
  dataFile.print(err);
  dataFile.print(" U ");
  dataFile.println(uval, 4);
  dataFile.close();
}

void applyPIDandDrive() {
  integral += Error * dt;
  if (integral > I_LIMIT) integral = I_LIMIT;
  if (integral < -I_LIMIT) integral = -I_LIMIT;
  derivative = (Error - prev_error) / dt;
  u = Kp * Error + Ki * integral + Kd * derivative;
  prev_error = Error;
  S_R = round(255 + u);
  S_L = round(255 - u);
}

```

```

S_R = constrain(S_R, 0, 255);
S_L = constrain(S_L, 0, 255);
analogWrite(9, S_R);
analogWrite(3, S_L);
}

void loop() {
    float bestKp = 0, bestKi = 0, bestKd = 0;
    float bestAbsErr = 1e12;
    for (int kp_i = 0; kp_i <= (int)((Kp_max - Kp_min)/Kp_step); kp_i++) {
        float testKp = Kp_min + kp_i * Kp_step;
        for (int ki_i = 0; ki_i <= (int)((Ki_max - Ki_min)/Ki_step); ki_i++) {
            float testKi = Ki_min + ki_i * Ki_step;
            for (int kd_i = 0; kd_i <= (int)((Kd_max - Kd_min)/Kd_step); kd_i++) {
                float testKd = Kd_min + kd_i * Kd_step;
                stopMotors();
                delay(500);
                int currentHeading = readHeading();
                Setpoint = currentHeading + 180;
                if (Setpoint >= 360) Setpoint -= 360;
                if (Setpoint < 0) Setpoint += 360;
                Kp = testKp;
                Ki = testKi;
                Kd = testKd;
                integral = 0;
                prev_error = 0;
                float sumAbsErr = 0.0;

                unsigned long tstart = millis();
                prev_time = millis();
            }
        }
    }
}

```

```

while (millis() - tstart < 5000) {
    now = millis();
    dt = (now - prev_time) / 1000.0;
    if (dt <= 0) dt = 0.03;
    prev_time = now;
    Input = readHeading();
    int diff = Setpoint - Input;
    if (diff > 180) diff -= 360;
    if (diff < -180) diff += 360;
    Error = diff;
    sumAbsErr += abs(Error);
    applyPIDandDrive();
    //writeRunErrToSD(Error, u);
}
stopMotors();
writeSummaryToSD(testKp, testKi, testKd, sumAbsErr);
if (sumAbsErr < bestAbsErr) {
    bestAbsErr = sumAbsErr;
    bestKp = testKp;
    bestKi = testKi;
    bestKd = testKd;
}
delay(500);
}
}
}
writeBestToSD(bestKp, bestKi, bestKd, bestAbsErr);

Kp = bestKp;

```

```

Ki = bestKi;
Kd = bestKd;
integral = 0;
prev_error = 0;

while (1) {
    now = millis();
    dt = (now - prev_time) / 1000.0;
    if (dt <= 0) dt = 0.03;
    prev_time = now;
    Input = readHeading();
    int diff = Setpoint - Input;
    if (diff > 180) diff -= 360;
    if (diff < -180) diff += 360;
    Error = diff;
    applyPIDandDrive();
    writeRunErrToSD(Error, u);
}
}

```

Appendix B: Video recordings of autonomous robot experiments.

<https://drive.google.com/file/d/1eTYxCydkjtjSSMoQyBL8wGXd4uwoA-HZs/view?usp=sharing>

CHAPTER 8

The Impact of Transformational Leadership on Sustainability and Business Performance Through ERP Systems and Supply Chain Integration

**İrem Candaş¹ & Sertan Çal² &
Batin Latif Aylak³ & Mehmet Ali Taş⁴**

¹ Turkish-German University, ORCID: 0009-0002-2780-5864

² Turkish-German University, ORCID: 0009-0004-3026-6587

³ Assoc. Prof. Dr., Turkish-German University, ORCID: 0000-0003-0067-1835

⁴ Re. Asst., Turkish-German University, ORCID: 0000-0003-3333-7972

1 INTRODUCTION

Nowadays, consumers pay attention not only to product or service quality, but also to an ethical and environmentally friendly supply chain. Generation Z in particular prefers brands with sustainable practices and thus influences the market. Investors are also increasingly valuing environmental, social, and governance (ESG) criteria. Sustainability is thus becoming a central component of long-term corporate strategies (Weston & Nnadi, 2023). Increasing pressure to be sustainable is forcing companies to redesign their processes. Supply chain management and ERP systems play a key role in this, for example in resource efficiency, CO₂ reduction, digitalization, and transparency. Sustainable supply chain management and data-driven decision support through ERP systems have a direct impact on sustainability performance. Therefore, SCM and ERP systems must become increasingly integrated, flexible, and sustainable (De Soete, 2016). The effective introduction of ERP systems and sustainable SCM practices depends heavily on good leadership, as these changes require both technical and organizational adjustments. New ways of working, digital processes, and employee adaptation can trigger resistance, which can be counteracted by a visionary, participatory leadership style. Transformational leadership motivates, inspires, and supports the success of such change processes (Pawar, 2016).

Transformational leadership is predominantly addressed in the literature in the context of topics such as corporate performance, employee motivation, innovation, and organizational commitment (Hu et al., 2013). However, how this leadership approach interacts with corporate sustainability strategies, particularly in sectors with intensive resource consumption and significant environmental impacts, remains insufficiently researched. This situation creates a significant theoretical gap regarding how a complex and long-term corporate goal like sustainability is shaped by leadership styles. The food sector, in particular, offers an extremely suitable basis for analyzing the impact of this leadership style due to its high energy and water consumption, complex supply chain structure, and intense social responsibility pressures (Maloni & Brown, 2006). In this context, this study examines the role of transformational leadership in achieving sustainability goals, improving performance, and integrating ERP and supply chain systems. First, the leadership characteristics of two different CEOs Paul Bulcke & Mark Schneider are assessed using the MLQ model to determine their levels of transformational leadership across four key dimensions. The observed leadership characteristics are then systematically compared with company-related outcomes (specifically sustainability, ERP and supply chain integration, and corporate performance) using the House of Quality (HoQ) methodology. The

combination of the MLQ and HoQ provides both qualitative and quantitative insights, providing an objective basis for evaluating the two leadership eras. The results are presented in the next section. Finally, the key findings are summarized, placed within a theoretical framework, and recommendations for future research are offered.

This study consists of five sections. Following an introduction, the second section introduces the concept of transformational leadership and provides a literature review. The third section describes the materials and methodology used, and the fourth section presents the results and discussion. The study concludes with a conclusion.

2 DEFINITION OF THE CONCEPT AND LITERATURE REVIEW

Leadership is the process by which a person, under certain conditions, influences and directs the activities of others to achieve personal or group goals (Ibrahim & Daniel, 2019). Leaders not only possess authority but are also individuals who can build trust, develop a vision, and motivate others to act in accordance with that vision (Fairholm, 1994). Among the most prominent and extensively discussed modern leadership approaches in the literature are charismatic leadership, transactional leadership, transformational leadership, and visionary leadership. Transformational leadership is one of the most comprehensive and effective models among modern leadership approaches. The fundamental difference from other approaches lies in the fact that the leader does not focus solely on results but strives for long-term change and organizational commitment by transforming the beliefs, values, and potential of employees (Aniebonam et al., 2023).

The term transformational leadership was first introduced in 1978 by political scientist James MacGregor Burns (1978). This concept was further developed by Bass (1985). According to Bass (1985), a transformational leader not only pursues the achievement of organizational goals but also places great emphasis on the development of their employees, encouraging them to discover their own potential (Bass, 1985). The four pillars of transformational leadership are often referred to in the literature as "Bass's Four I's (Figure 1)." These "Four I's" are crucial for understanding how leaders transform both individuals and organizations (Bass, 1985):



Figure 1: Bass' Four I's (Bass, 1985)

Idealized Influence (charisma): A transformational leader serves as a role model for their employees through their behavior. High ethical standards, a consistent personality, and decisiveness in difficult situations create trust and respect among the leader's environment.

Inspirational Motivation: In this dimension, the leader develops a strong and meaningful vision that aligns employees towards a common goal.

Intellectual Stimulation: A transformational leader encourages employees to question existing thought patterns and develop new perspectives. Mistakes are seen as learning opportunities, and innovative ideas are fostered.

Individualized Consideration: A leader who is aware of the diverse needs, interests, and potential of their employees prioritizes individual development. Through methods such as mentoring, coaching, and feedback, they support employees in recognizing their strengths.

The impact of this leadership approach can be observed not only in the private sector but also in areas such as public administration and technology, through concrete examples. A prime example of transformational leadership is Yvon Chouinard, founder of Patagonia (Chouinard & Stanley, 2013). He integrated personal environmental and ethical values into the company's strategy and motivated employees far beyond financial targets. His visionary leadership style made Patagonia both economically successful and a pioneer in sustainability. Patagonia consistently uses recycled and organic materials as standard and pursues comprehensive social responsibility – from fair working conditions to

family-friendly policies. While many companies are criticized for problematic supply chains, Patagonia firmly embedded ethical practices in its global processes. The transfer of all company shares to an environmental foundation in 2022 is considered a radical milestone. With this, Patagonia demonstrates that transformational leadership enables values-based, long-term thinking and can foster organizational change (Zucchella et al., 2023). The impact of this approach on sustainability and corporate performance will be explored further in the following chapters using the example of Nestlé.

The relationship between transformational leadership and organizational performance has become a focus of numerous studies in past years (García-Morales et al., 2008; Samad, 2012; Mutahar et al., 2015; Orabi, 2016; Rahmatullah et al., 2022; Hilton et al., 2023), particularly in the context of digitalization (Suwanto et al., 2022), information systems (Faradina & Mabrur, 2023), and supply chain integration. A notable example in this regard is the empirical study by Setiabudi et al. (2021). This work analyzes the impact of transformational leadership behavior on organizational performance by modeling two key elements of organizational infrastructure: enterprise resource planning (ERP) systems and supply chain integration (SCI). The findings show that while the direct influence of transformational leadership on organizational performance is not statistically significant, it does occur indirectly through ERP systems and supply chain integration. In particular, ERP systems are positively influenced by transformational leadership behavior, and these systems strengthen both supply chain integration and directly enhance organizational performance. These results highlight the indirect contribution of leadership to performance through its ability to integrate technical systems and manage digital transformation processes (Setiabudi et al., 2021). Elkhani et al. (2014) examined the relationship between ERP and transformational leadership and proposed an approach called TAM. A study of Shao et al. (2012), which is a research in China and compatible with the same context, can be given as an example.

3 MATERIAL AND METHOD

The aim of this study is to illustrate the impact of transformational leadership on corporate sustainability, ERP systems, and supply chain integration not only on a theoretical level, but also using a concrete and comparative methodology. To make this influence more visible and measurable, the study compares two different CEOs who worked within the same company. As a first step, the Multifactor Leadership Questionnaire (MLQ) is applied to assess the leadership characteristics of both CEOs and systematically quantify the degree of their transformational leadership. Subsequently, the changes observed within the

company during their respective leadership periods in the areas of sustainability practices, ERP systems and supply chain integration, as well as company performance, are examined. Based on the findings, the influence of transformational leadership characteristics on these categories is analyzed in detail using the House of Quality (HoQ) methodology. By testing the hypotheses formulated within this holistic assessment, the aim is to obtain objective and comparative results regarding how transformational leadership influences the organizational structure.

In this context, the following subsections first explain the general characteristics of the company selected for analysis and the reasons for its selection. Next, the leadership periods and strategic approaches of the two CEOs being evaluated are presented. Following this, the criteria for determining the categories of sustainability, ERP and supply chain integration, and corporate performance—in which the impact of transformational leadership is to be analyzed—are explained. The Multifactor Leadership Questionnaire (MLQ) is then used to assess the transformational leadership characteristics of both CEOs; this analysis forms the basis for evaluating and comparatively classifying their leadership approaches. After the MLQ analysis, all categories are consolidated and systematically evaluated using the House of Quality (HoQ) methodology. Finally, the data collection process and the evaluation system are defined to illustrate the analysis process step by step. In this way, the methodological framework of the work is presented to the reader in a transparent and comprehensive manner.

3.1 Introduction of the company

In this study, the food sector was chosen as the focus industry. Nestlé was selected for analysis as one of the leading companies in the sector. Its broad product portfolio and global network make Nestlé a suitable example for examining supply chain integration and ERP systems. Furthermore, Nestlé has long pursued an active corporate strategy in the area of sustainability. Information available on Nestlé, which allows for an in-depth analysis of the company, can be accessed through publicly available sources (Nestlé, 2024a). This makes it possible to base the assessments made within the study not only on qualitative observations but also on data-driven and objective analyses.

3.2 Paul Bulcke & Mark Schneider: Their terms of office and strategic approaches

This study comparatively examines two Nestlé leaders, Paul Bulcke and Mark Schneider, within the context of transformational leadership. Both leaders played

a crucial role in Nestlé's global strategic direction but these two CEOs represent different profiles in terms of background, leadership style, and the business areas they prioritized (Takahashi, 2021).

Paul Bulcke, a Belgian entrepreneur, joined Nestlé in 1979 and spent the majority of his career within the company. Under Bulcke's leadership, Nestlé focused primarily on restructuring its brand portfolio, improving operational efficiency, and maintaining organizational stability. As an executive from within the company, he was intimately familiar with its traditional structure and implemented measures to reinforce existing structures rather than initiating fundamental changes. While topics such as sustainability and digital transformation were present, they received only limited strategic attention during this period. Bulcke currently serves as Honorary Chairman of Nestlé (Nestlé, 2025a).

Mark Schneider is the first CEO in Nestlé's history to be appointed from outside the company (Nestlé, 2025b). During his tenure at Nestlé, he pursued the goal of transforming the corporate culture into a more agile and innovation-oriented structure; in doing so, he placed sustainability-oriented measures such as the further development of ERP systems, the digital integration of the supply chain, the transparency of ESG performance, and commitments to achieving net-zero emissions at the center of his strategic agenda. Schneider is the CEO of the company (Nestlé, 2025b).

3.3 MLQ analysis for Paul Bulcke

The Multifactor Leadership Questionnaire (MLQ) is one of the most widely used and scientifically validated psychometric instruments for measuring leadership styles. The MLQ examines the transformational leadership approach using four main dimensions: idealized influence, inspirational motivation, intellectual stimulation, and individual consideration.

3.3.1 Idealized influence (II)

Under Schneider's leadership, Nestlé has demonstrated that it is not only a company focused on economic performance, but also positions itself as an institution that takes responsibility for society, the environment, and future generations. The company's long-term strategic commitments were clearly expressed in the "Creating Shared Value Report and meeting our commitments 2020" report: "...our Net Zero Roadmap, we are aiming to halve emissions by 2030 and be net zero by 2050." (Nestlé, 2020, p. 6) The same document points out that Nestlé has committed to systematically disseminating regenerative

agriculture practices to achieve environmental improvements in the food supply chain (Nestlé, 2020, p. 7). This approach is presented not only as a performance target, but also as a form of ethical responsibility towards future generations.

During Schneider's tenure, the term "nutrition" was also closely linked to the company's purpose, and Nestlé's vision of "working for a healthier, longer life" was reiterated in various documents (Nestlé, 2020, pp. 2–4). This consistency in corporate discourse is a factor that reinforces the ethical and principled stance of the leadership. At the same time, Nestlé's global product donations during the COVID-19 pandemic, its collaborations with healthcare systems, and its community support projects demonstrate a leadership profile that remains true to the company's values, even in times of crisis. Such examples show that Schneider's leadership is not only strategic but also values-driven.

3.3.2 Inspirational motivation (IM)

Schneider frequently emphasizes social goals such as reaching low-income communities, reducing health inequalities, and implementing inclusive product strategies have been formulated at a strategic level. This vision was clearly articulated in the presentation at the CECF CEO Investor Forum 2019 (Nestlé, 2019a), which stressed the need to make improved nutrition accessible and affordable. In this spirit, under Schneider's leadership, Nestlé has transformed its corporate motivation into a socially oriented one – with goals such as carbon neutrality, combating climate change, implementing regenerative agricultural practices, and developing programs aimed at younger generations. Furthermore, the phrase “Good Food, Good Life” became a central element of the company’s identity during this period, and the promotion of a healthier and longer life became the focus of Nestlé’s operational strategy (Nestlé, 2025c).

3.3.3 Intellectual stimulation (IS)

Nestlé's Creating Shared Value Report and meeting our commitments 2018 (Nestlé, 2018) highlighted the company's science-based innovation strategy. Investments in personalized nutrition, in particular, demonstrate the innovative focus of the Schneider era. During this period, health-, data-, and nutritional-oriented product development processes gained strategic priority; the company's R&D policy focused not only on technological advancement but also on addressing societal challenges. Furthermore, Nestlé positioned digitalization not merely as an operational infrastructure feature but as a cornerstone of its corporate transformation. Another key element of Nestlé's innovation strategy was its interdisciplinary approach. Collaboration across departments was fostered

in new product development, market analysis, and sustainability projects; data-driven decision-making became an integral part of the corporate culture.

3.3.4 Individual consideration (IC)

The Creating Shared Value Report and meeting our commitments 2018 (Nestlé, 2018) outlined the scope of Nestlé's global initiatives to support the economic participation of young people. Training and internship opportunities to improve the employability of young talent were implemented through regional programs in Africa, Latin America, and Europe. In the same document, Nestlé presents the goal of creating an inclusive and diversity-conscious workplace as a strategic corporate policy. Programs to promote female leadership, community-based educational initiatives, and HR measures to represent diverse cultural profiles are identified as components of this commitment. However, all these measures are predominantly structured at the organizational level. The reports do not systematically document concrete implementations such as individual mentoring programs, coaching, or personal development plans.

3.4. MLQ analysis for Mark Schneider

3.4.1 Idealisierter Einfluss (II)

Nestlé's corporate vision during this period was clearly and decisively formulated as follows: "Enhancing quality of life and contributing to a healthier future" (Nestlé, 2019b). This statement was systematically reiterated as a corporate objective in both strategic documents and annual reports. Under Bulcke's leadership, the company's strategic direction was linked not only to short-term performance but also to stakeholder trust and long-term sustainability. At the UN Global Compact event held that same year, messages regarding ethical business responsibility and accountability to society were also conveyed. It was emphasized that companies should not only be profit-oriented but should also act in the service of society.

3.4.2 Inspirational motivation (IM)

Youth programs designed for developing countries, initiatives to promote local female entrepreneurs, and food safety strategies demonstrate that this vision has been implemented at the operational level. Nestlé's 2013-2014 (Nestlé, 2013; Nestlé, 2014a) annual and Creating Shared Value reports emphasized that promoting healthy living is not only a business necessity but also an ethical responsibility.

3.4.3 Intellectual stimulation (IS)

Founded in 2011, Nestlé Health Science (NHS, 2025) exemplifies the company's commitment to personalized nutrition and health solutions. The 2014 report highlighted NHS several times, emphasizing its role in driving product development and contributing to public health through innovative approaches in the health sector (Nestlé, 2014b). Nestlé Health Science is a core corporate entity that translates its scientific research capabilities into health-based innovation. Furthermore, during Bulcke's tenure, external sources of innovation were incorporated as part of digital transformation initiatives, and open environments for innovative projects were created. Although platforms like "Henri@Nestlé" are not explicitly mentioned in the reports, the company's overall innovation policy was based on an open innovation approach that supported internal R&D. Nevertheless, systematic insights into whether innovation was integrated into the overall corporate culture remain limited. Innovation initiatives were predominantly designed at the management level, while concrete examples of employee participation in critical thinking were less visible.

3.4.4 Individual consideration (IC)

During Paul Bulcke's tenure as CEO, Nestlé made strong corporate commitments, particularly regarding the integration of young people into the labor market and their contribution to local development. In this context, the "Nestlé Needs YOUth" initiative was launched to promote the education, employment, and professional development of young people in Europe and emerging markets. Training, technical support, and infrastructure investments for small-scale producers were implemented as part of Nestlé's rural development strategy. These measures aim to increase the long-term productivity and economic participation of individuals. However, these programs were largely organized at the community level, and concrete examples of direct employee mentoring, individual leadership development, or personal coaching systems were not systematically documented. For instance, publicly available reports contain only limited information on internal mentoring programs.

3.5 Comparison between Paul Bulcke and Mark Schneider

Both executives received full marks in the dimensions of Idealized Influence (II) and Inspiring Motivation (IM) thus fulfilling the core criteria of transformational leadership to a high degree. During Paul Bulcke's tenure, it was clearly documented that the company firmly adhered to its ethical values and integrated Nestlé's mission—"Enhancing quality of life and contributing to a healthier future"—into its corporate identity. Schneider, on the other hand,

viewed ethical leadership within the framework of a broader sustainability approach, underpinning it with stakeholder-based trust and environmental responsibility. In the context of "Inspiring Motivation (IM)," both executives developed inspiring statements and strategies. However, Mark Schneider, with his visionary approach—for example, through climate goals such as "Net Zero by 2050" and programs like "Nutrition for All," which aim for global inclusion—elevated corporate motivation to a more comprehensive level. Therefore, Schneider demonstrates a more advanced profile in terms of convincingly communicating a vision not only strategically but also on an emotional level.

Regarding the dimension of Individual Consideration (IC) both executives received comparatively lower scores. While Bulcke implemented initiatives to promote youth employment, no systematic implementation of individual measures such as mentoring or internal coaching was observed. Under Schneider, diversity, equality, and youth initiatives were more firmly anchored at the institutional level; however, the documentation of individual development plans remains limited. For this reason, the scores were set at 2 and 3 points, respectively (Table 1).

This comparison shows that while both executives exhibit characteristics of transformational leadership, this leadership approach was implemented more strategically, in a more integrated manner, and more systematically under Mark Schneider. In particular, the degree of institutionalization in the IS and IC dimensions demonstrates that Schneider's leadership style aligns more closely with modern approaches to transformational leadership.

Table 1: Points for MLQ analysis (Authors’ own elaboration)

MLQ	Paul Bulcke	Mark Schneider
Idealized Influence	4	4
Inspirational Motivation	4	4
Intellectual Stimulation	3	4
Individualized Consideration	2	3
Average	3.25 / 4	3.75 / 4

3.6 HoQ Analyse

This study uses the HoQ method to structure the relationships between the dimensions of transformational leadership (MLQ-5X model) and strategic business outputs (business performance, sustainability, ERP systems, supply chain integration) and to illustrate priorities, impact, and structural relationships

through comparative analysis. In this way, the influence of leadership behavior on strategic objectives can be quantified and systematically analyzed (Assefa et al., 2025). The HoQ structure forms a relationship matrix between customer requirements, the so-called WHATs, and the technical or strategic solutions, the HOWs, that aim to fulfill these requirements. These relationships are rated on a scale of 0-1-3-9 and multiplied by the respective importance of the WHATs, thereby determining prioritized design criteria (weightings) (Chan & Wu, 2002). The top-level structure in the upper part of the matrix illustrates the technical and strategic interactions between the HOWs.

3.6.1 WHATs

The WHATs correspond to the four main dimensions of transformational leadership as described in the relevant literature and defined by the MLQ-5X scale:

Idealized Influence (II): Ethical behavior, building trust, role model function

Inspirational Motivation (IM): Communicating an inspiring vision and motivation

Intellectual Stimulation (IS): Fostering innovation and critical thinking individualized

Consideration (IC): Sensitivity to individual development and coaching

These dimensions were evaluated using MLQ analysis based on content from the leadership periods of both CEOs on a scale of 0 to 4.

3.6.2 HOWs

The HOWs are the technical, strategic, or organizational outputs aimed at fulfilling the requirements represented by the WHATs. The HOWs were analyzed both in terms of their connections to the WHATs (rated 0-1-3-9) and their mutual correlations. For this study, 22 HOWs were defined and grouped into four main categories: business performance, sustainability, ERP systems, and supply chain integration.

Company performance: This category was developed to measure the financial and operational results that can be directly influenced by leadership. Subcategories:

- Revenue growth: Demonstrates leadership in marketing and strategic direction. Assessed based on the annual net revenue growth rate,

- Profit margin: This is the result of the company's financial efficiency, expense management, and strategic planning. The development of gross and net profit margins is analyzed.
- Market share growth: Reflects the competitiveness and influence of the leadership in the market. The company's market position is assessed on an annual basis.
- Innovation and product development: One of the areas most strongly influenced by transformational leadership. The number of new products, R&D investments, and patent data form the basis for evaluation.

Operational efficiency: This is related to process optimization, productivity, and cost reduction. The evaluation is based on process data and reports.

- Crisis management and resilience: Assesses how leadership has steered the organization through external shocks such as the pandemic. Performance data from crisis periods are taken into account.

Sustainability: Sustainability stands not only for current performance, but also for the leadership's sense of responsibility towards the future. Subcategories:

- Environmental sustainability: Aspects such as carbon footprint, energy management, and reduction of environmental impacts are assessed. The company's environmental policy and its implementation results are analyzed and evaluated.
- Social sustainability: Employee rights, occupational safety, diversity, and social engagement are the focus. The company's CSR projects and HR practices are analyzed.
- Economic sustainability: Resilience to crises, long-term growth, and risk management strategies are assessed. Financial sustainability indicators form the basis of the evaluation.
- Sustainability reporting: The extent of reporting according to standards such as GRI and SDGs is examined. Assessment focuses on transparency, use of metrics, and consistency.
- Corporate sustainability strategy: This assesses the extent to which sustainability is integrated into the company's strategy. Responsibilities at the top management level, established goals, and strategic plans are taken into account.

ERP systems: This category was developed to assess the impact of transformational leadership on digital transformation and process integration.

- **System quality:** The technical stability, reliability, and performance of the ERP system are evaluated. The evaluation is based on the depth of integration of platforms such as SAP into the company's processes.
- **Response time:** The data processing and response speed of the ERP system is analyzed. Particular focus is placed on operational agility and system efficiency.
- **Information quality:** The accuracy, consistency, and decision relevance of the data generated by the ERP system are evaluated. Reporting standards and data quality are included in the evaluation.
- **User-related effectiveness:** The acceptance and use of the ERP system by employees, user-friendliness, and the availability of training programs are evaluated.
- **Impact on business processes:** The analysis examines the extent to which the ERP system contributes to increased efficiency, process improvement, and faster decision-making. Specific case studies are used for evaluation.

Supply chain integration (SCI): This category shows the extent to which leadership promotes organizational collaboration, process coordination, and transparency in the supply chain. Subcategories:

- **Demand forecasting capability:** This assesses how accurately and promptly the company forecasts future demand. The evaluation is based on forecasting systems and success rates.
- **Management of supply contracts:** The flexibility, sustainability and crisis resistance of contracts with suppliers are analyzed.
- **Supplier cooperation:** The depth of strategic partnerships with suppliers, the joint level of innovation and the mutual transfer of knowledge are evaluated.
- **Information exchange:** Real-time data transmission, the integration of digital systems, and the quality of communication between partners are evaluated.

- Logistics integration: The integration of logistics networks into ERP systems, as well as traceability and digital coordination of delivery processes, are analyzed.
- Internal problem-solving capability: The assessment evaluates the company's ability to resolve internal supply chain problems quickly and using data. Crisis management and decision-making mechanisms are of particular importance.

3.6.3 Relationship Matrix

One of the central components of the House of Quality (HoQ) analysis is the numerical evaluation of the relationship between customer needs (WHATs) and technical or strategic outputs (HOWs). This evaluation serves as analytical decision support and is defined in the QFD literature by the standardized scale of 0-1-3-9 (Hauser & Clausing, 1988; Akao, 1990) (Table 2).

Table 2: Relationship symbols for HoQ analysis (Authors' own elaboration)

Points	Symbol	Explanation
0		No relationship
1		Weak relationship
3		Moderate relationship
9		Strong relationship

3.6.4 Weighting:

In House-of-Quality (HoQ) analysis, the calculation of weighted scores is the central method for numerically determining the priority of each technical or strategic solution (HOW) in Eq (1):

$$\text{Weight}(\text{TR}_i) = V(\text{TR}_{i1}) \times \text{Imp}(\text{CA}_1) + V(\text{TR}_{i2}) \times \text{Imp}(\text{CA}_2) + \dots + V(\text{TR}_{in}) \times \text{Imp}(\text{CA}_n) \quad (1)$$

Here, $V(\text{TR}_{in})$ represents the correlation value between the technical requirement TR_i and the customer request CA_n . The function $\text{Imp}(\text{CA}_n)$ describes the importance or priority of the respective customer request CA_n . The weighting of a technical requirement TR_i is thus calculated as the weighted sum of all relationships between TR_i and the individual customer requirements, where each relationship is multiplied by the importance of the corresponding customer request. As a result of this calculation, a total numerical value is determined for each HOW. This value represents the strength of the connection between the respective output and the leadership behavior, as well as its relative priority.

3.6.5 Correlation matrix (Roof Matrix)

The roof structure (roof matrix) serves to analyze the reciprocal relationships and interactions between the HOW elements, which are defined as technical or strategic outputs. The roof shows whether a particular HOW has a positive, negative, or neutral relationship with other HOWs (Mazur, 1993). In the literature, this method is often described as a structured analysis based on qualitative observations (Hauser & Clausing, 1988; Mazur, 1993). The main objective of the overarching matrix is to visually and systematically represent technical dependencies or strategic synergies between the HOWs (How-To Works). This particularly enables the identification of potential positive or negative impacts of one output on another when a specific outcome is being pursued (Chan & Wu, 2002).

Synergistic relationships (e.g., ++ or +) help identify strategies that strengthen multiple goals simultaneously. Conflict relationships (e.g., – or —), on the other hand, highlight the risk that strategic decisions might bring improvements in one area while harming another.

The correlations are evaluated as follows (Table 3):

Table 3: Symbols for correlation

Points	Symbol		Explanation
++	Strong correlation	positive	The outputs develop together
+	Moderate relationship	positive	Progress is possible
	No relationship		The outputs are independent of each other
-	Weak correlation	negative	When one rises, the other is potentially inhibited
--	Strong correlation	negative	Conflict of objectives between the outputs

This structure plays a crucial role in understanding whether output goals can be improved collectively (Zairi & Youssef, 1995). Based on the HoQ components explained in detail above, this study quantified and systematically structured the relationships between the four central dimensions of transformational leadership (WHATs) and the 22 strategic outputs (HOWs) according to their importance and levels of interrelationship.

3.7 House of Quality Analysis

The resulting House of Quality Matrix represents a holistic and visual decision support tool that simultaneously illustrates both the impact of leadership behavior on organizational outcomes and the strategic relationships between these outputs. The figure below depicts the House of Quality Matrix, which is the main outcome of the study and consolidates all the structured assessments, correlations, and calculations from the analysis process under one framework. HoQ analysis for Paul and Mark is shown in Figure 2.

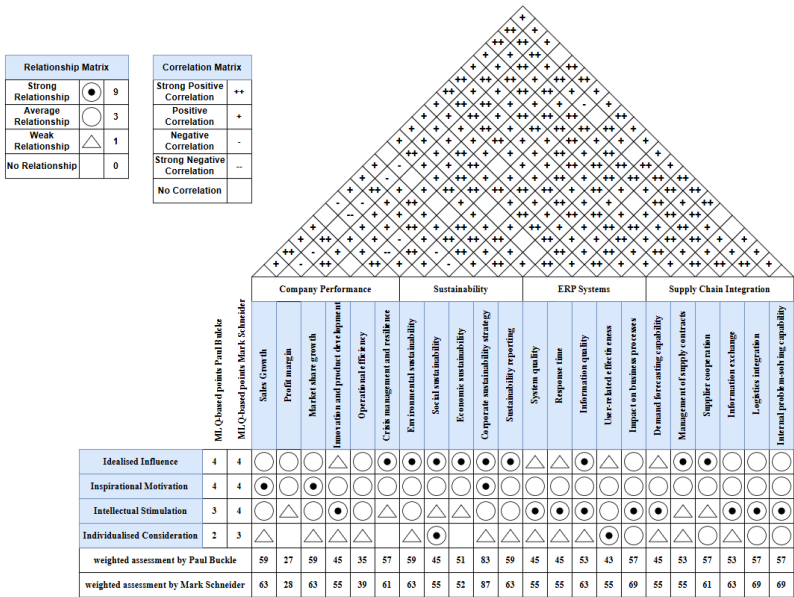


Figure 2: HoQ analysis for Paul and Mark

In accordance with the detailed methodological framework, the relationships between the two CEOs' evaluation scores in the dimensions of transformational leadership and the defined outcome categories were quantified using the HoQ matrix. The analysis table shown above quantitatively illustrates the influence of leadership behavior on strategic business outcomes and forms the empirical basis of this work. This analysis provides the structural foundation for the results explained in more detail in the next section.

In the literature review, the study by Setiabudi et al. (2021) stands out as one of the most significant empirical models, in which the impact of transformational leadership on business performance is indirectly tested via ERP systems and

supply chain integration (SCI). Based on this model from the literature, the following six hypotheses were formulated:

H1: Transformational leadership has a positive impact on ERP systems.

H2: Transformational leadership increases supply chain integration.

H3: Transformational leadership directly improves business performance.

H4: ERP systems have a positive impact on business performance.

H5: Integrating the supply chain has a positive impact on business performance.

H6: ERP systems strengthen the integration of the supply chain.

The original contribution of this work to the literature lies in the integration of the sustainability factor into the conceptual framework. Therefore, four new hypotheses were developed within the framework of this dissertation and integrated into the holistic analysis:

H7: Transformational leadership promotes the implementation of sustainability practices.

H8: Sustainability increases the strategic effectiveness of ERP systems.

H9: Sustainability contributes to the integration of the supply chain.

H10: Sustainability practices improve business performance.

This structure made it possible to assess not only the influence of transformational leadership on technical systems (ERP-SCI), but also its structural impact through sustainability.

4 RESULTS

4.1 Hypotheses

4.1.1. H1: Transformational leadership positively influences the acceptance of ERP systems

In the literature, the organizational acceptance of ERP systems is linked to leadership behavior; in particular, it has been argued that components of transformational leadership such as vision, openness to change and support of users are crucial for the success of ERP systems (Setiabudi et al., 2021) (Table 4).

Table 4: MLQ dimension for H1 (Authors' own elaboration)

MLQ-Dimension	Assigned ERP results
Intellectual stimulation	Information quality, influence on business processes
Individualized consideration	Effectiveness at user level
Idealized influence & inspiring motivation	System quality, response time

In this hypothesis, all four dimensions of transformational leadership (II, IM, IS, IC) can influence ERP results. The relationships mentioned above are particularly noteworthy (Table 5).

Table 5: H1 scores for Hows (Authors' own calculation)

ERP subcomponent (HOW)	Paul Bulcke	Mark Schneider
ERP1 – System quality	45	55
ERP2 – Response time	45	55
ERP3 – Information quality	53	63
ERP4 – Effectiveness at user level	43	55
ERP5 – Influence on business processes	57	69
Overall ERP score	243	297

Mark Schneider's weighted overall influence on the ERP results is 54 points higher compared to Paul Bulcke. H1 results from correlation matrix are shown in Table 6.

Table 6: H1 results from correlation matrix (Authors' own elaboration)

Results of correlation matrix	
ERP1 & ERP3	++
ERP3 & ERP5	++
ERP4 & ERP5	+
ERP2 & ERP5	+

ERP expenditures show strong systemic interrelationships: This correlation structure demonstrates that transformational leadership not only improves individual results but also enables the holistic integration of the ERP system.

H1 hypothesis evaluation: H1 is supported.

During Mark Schneider's tenure, a more balanced and comprehensive influence on ERP outcomes was observed, particularly due to high scores in Intellectual Stimulation (IS) and Individualized Consideration (IC). ERP systems received support not only technically, but also culturally and at the user level. In contrast, during Paul Bulcke's tenure, ERP applications developed primarily with a focus on operational efficiency, while user engagement and strategic transformation remained limited. The weighted HoQ scores and the roof correlation matrix clearly demonstrate the influence of transformational leadership on ERP systems. This influence became particularly evident at the institutional level during Mark Schneider's tenure and was comprehensively integrated into the organizational structure.

4.1.2 H2: Transformational leadership supports the integration of supply chains

Leadership literature argues that cooperative, visionary, and technologically savvy leaders make information exchange with suppliers, contract management, and logistical synchronization more effective (Setiabudi et al., 2021). SCI results are directly related to organizational integration. Therefore, the impact of the dimensions Intellectual Stimulation (IS) and Idealized Influence (II) on SCI is of particular importance. Table 7: MLQ dimension for H2 is displayed in Table 7.

Table 7: MLQ dimension for H2 (Authors’ own elaboration)

MLQ Dimension	Associated Results
Intellectual Stimulation	SCI1 Forecasting, SCI4 Info Exchange, SCI 6 Internal Problem Solving
Idealized Influence	SCI2 Supply Contract, SCI3 Partnership
Individualized Consideration	SCI6 Internal Problem Solving
Inspirational Motivation	SCI3 Partnership, SCI5 Logistics Integration

Mark Schneider's total weighted influence on SCI (Supply Chain Integration) is 50 points higher than Paul Bulcke's (Table 8).

Table 8: H2 scores for Hows (Authors' own calculation)

SCI subcomponent (HOW)	Paul Bulcke	Mark Schneider
SCI1 Forecasting	45	55
SCI2 Supply Contract	53	55
SCI3 Partnership	57	61
SCI4 Information Exchange	53	63
SCI5 Logistics Integration	57	69
SCI6 Internal Problem Solving	57	69
Total SCI score	322	372

This structure shows that integration is influenced not only technically, but also relationally and process-related through leadership. H2 results from correlation matrix is in Table 9.

Table 9: H2 results from correlation matrix (Authors' own elaboration)

Results of correlation matrix	
SCI1 & SCI4	++
SCI3 & SCI4	++
SCI5 & SCI6	+
SCI2 & SCI3	++

H2 hypothesis evaluation: H2 is supported.

Investments made during Mark Schneider's tenure—such as the digital supply chain, data-driven forecasting systems, and more transparent contract management—were directly institutionalized at the organizational level in the area of Supply Chain Integration (SCI) outcomes. Supplier relationships and system coordination were strengthened, particularly through the leadership behavior dimensions of Intellectual Stimulation (IS) and Idealized Influence (II). In contrast, a more traditional supply chain structure was maintained during Paul Bulcke's tenure; aspects such as collaboration and digitalization remained limited. The weighted scores and correlations within the SCI category demonstrate that transformational leadership—especially during Mark Schneider's tenure—structurally supported supply chain integration.

4.1.3 H3: Transformational leadership directly improves company performance

The literature emphasizes positive correlations between the dimensions of transformational leadership – such as vision, innovation and motivation – and organizational efficiency, market success and crisis resilience (Bass & Avolio, 1996).

The subcategories of organizational performance can be directly influenced by management strategies. The following is a summary of the MLQ-HOW mappings that are directly related to this hypothesis (Table 10):

Table 10: MLQ dimension for H3 (Authors’ own elaboration)

MLQ Dimension	Assigned Results
Intellectual Stimulation	FP4 Innovation Capability and Product Development, FP5 Operational Efficiency
Idealized Influence	FP1 Revenue Growth, FP3 Market Share Growth
Individualized Consideration	FP6 Crisis Management and Organizational Resilience
Inspiring motivation	FP1 Revenue growth, FP4 Innovation capability and product development

Mark Schneider's overall influence on the company's performance is 27 points higher than Paul Bulcke's (Table 11):

Table 11: H3 score for Hows (Authors’ own calculation)

Subcomponent (HOW)	Paul Bulcke	Mark Schneider
FP1 Revenue growth	59	63
FP2 Profit margin	27	28
FP3 Market share increase	59	63
FP4 Innovation capability and product development	45	55
FP5 Operational efficiency	35	39
FP6 Crisis resilience	57	61
Overall score	282	309

The correlation structure (Table 12) shows that strong internal relationships exist between the output categories directly influenced by transformational leadership. However, potential conflicts of objectives can arise for certain indicators – such as profit margin (FP2).

Table 12: H3 results from correlation matrix (Authors' own elaboration)

Results of correlation matrix	
FP1 & FP3	++
FP5 & FP6	++
FP3 & FP4	++
FP2 & FP4	-
FP1 & FP6	0

H3 hypothesis evaluation: H3 is supported.

Mark Schneider's high scores in the MLQ-5X dimensions—particularly in Intellectual Stimulation (IS) and Individualized Consideration (IC)—made a significant difference in performance areas such as FP4 (Innovation) and FP6 (Crisis Resilience). This difference was supported not only quantitatively but also by strategic transformation. In contrast, while strong results were achieved in categories such as revenue growth and market share during Paul Bulcke's tenure, innovation and user-level flexibility remained limited, restricting the company's efficiency and adaptability. Transformative leadership, especially during Mark Schneider's tenure, had a significant and direct impact on the company's key performance indicators. The weighted scores of the HoQ analysis and the correlations between the results clearly demonstrate this relationship. However, in some indicators, such as profit margin, the impact was limited, and indirect factors proved decisive.

4.1.4 H4: ERP systems improve business performance

This hypothesis posits that two different HOW groups interact with each other: ERP system subcategories (ERP1–ERP5) and company performance indicators (FP1–FP6). Therefore, both the cross-category correlation structure between these two groups and the correlation of the total weighting points should be considered. The correlation pair for H4 is shown in Table 13.

Table 13: Correlation pair for H4 (Authors' own elaboration)

Correlation pair	Value	Comment
ERP3 & FP4 (information quality – innovation)	++	Data-driven innovation
ERP5 & FP5 (process influence – efficiency)	++	Performance improvement through processes
ERP2 & FP6 (response time – crisis resilience)	++	Positive in terms of crisis response capability

A total of 30 cross-category correlations between ERP and FP were evaluated, and overall strong positive correlations were found. H4 results from correlation matrix is in Table 14.

Table 14: H4 results from correlation matrix (Authors' own elaboration)

Results of correlation matrix	
FP1 & ERP5	++
FP5 & ERP1, ERP2	++
FP2 & ERP3	+

These similarities show that ERP systems not only offer technical benefits, but also enable strategic performance improvements.

H4 hypothesis evaluation: H4 is supported.

ERP systems were further developed, particularly during Mark Schneider's tenure, with a focus on digital transformation. This improved their alignment with key performance indicators and made the system's impact on processes measurable. During Paul Bulcke's tenure, ERP systems primarily offered functional support at the operational level, while their strategic contribution remained limited. Nevertheless, a positive trend can be observed in both leadership phases. Strong positive correlations were found between ERP systems and business performance; in particular, process influence, information quality, and system efficiency had a performance-enhancing effect. The correlation matrix and the weighted scores support this effect.

4.1.5 H5: Supply chain integration increases company performance

The hypothesis is tested via the interaction of two outcome categories: supply chain integration (SCI1–SCI6) and business performance indicators (FP1–FP6). The evaluation is based on cross-category correlations and relationship score assessments. Thirty-six pairs were analyzed between the SCI results and the company's performance indicators. Some of the most prominent examples are (Table 15):

Table 15: Correlation pair for H5 (Authors’ own elaboration)

Correlation pair	Value	Comment
SCI1 Forecasting & FP1 Sales growth	++	Accurate forecasts promote sales growth
SCI3 Partnership & FP3 Market share	++	Trust in suppliers strengthens competitiveness
SCI5 Logistics integration & FP5 Efficiency	++	Logistics integration supports efficiency
SCI6 Internal problem solving & FP6 Resilience	++	Ability to respond quickly in crisis situations

This structure shows that SCI results are strongly integrated with company performance. The supply chain results (SCI1–SCI6) are linked to numerous HOWs that influence business performance. These assessments, weighted by HoQ, demonstrate that the SCI category makes a significant contribution to performance. Table 16 shows H5 results of correlation matrix.

Table 16: H5 results of correlation matrix (Authors’ own elaboration)

Results of correlation matrix	
SCI1& FP1	++
SCI5 & FP5	++
SCI3 & FP3	+

H5 hypothesis evaluation: H5 is supported.

During Schneider's tenure, digital solutions, particularly in areas such as information exchange, forecasting systems, and logistics synchronization, came to the fore. This led to a stronger alignment between supply chain SCI results and corporate performance indicators. In contrast, during Paul Bulcke's tenure, SCI applications were more traditional, and the capacity for internal problem-solving within processes remained limited. The correlation data and weighted scores

demonstrate that SCI results have a direct impact on corporate performance. Strategic indicators such as revenue growth, market share, and crisis resilience were particularly positively influenced by this relationship.

4.1.6 H6: ERP systems promote supply chain integration

This hypothesis is tested using two technical outcome categories: ERP systems (ERP1–ERP5) and supply chain integration (SCI1–SCI6). The analysis is performed through cross-category correlations and a comparison of the weighted HoQ scores (Table 17).

Table 17: Correlation pair for H6 (Authors’ own elaboration)

Correlation pair	Value	Comment
ERP3 Information quality & SCI4 Information exchange	++	High information quality ensures a precise flow of information in the supply chain
ERP5 Process influence & SCI5 Logistics integration	++	Process influence has a direct impact on logistical synchronization
ERP1 System quality & SCI1 Forecasting	++	High system quality supports the development of accurate forecasting systems

A total of 30 relationship pairs were analyzed. Overall, a positive trend was observed for all SCI results. The following correlations show a high weighted contribution, indicating a direct relationship between ERP results and SCI. These correlations quantify the organizational influence of ERP systems on SCI (Table 18).

Table 18: H6 Results of correlation matrix (Authors’ own elaboration)

Results of correlation matrix	
ERP3& SCI4	++
ERP5 & SCI5	++
ERP2 & SCI6	+

H6 hypothesis evaluation: H6 is supported.

The redesign of the ERP systems under Mark Schneider, with its focus on digital transformation and data flow, had a significant impact on the SCI results. A direct effect of ERP integration was particularly noticeable in areas such as information exchange (SCI4), forecasting (SCI1), and logistics (SCI5). During

Paul Bulcke's tenure, while the ERP systems led to operational efficiency improvements, their integration into the SCI remained limited. The high correlation between ERP systems and SCI results demonstrates that integration has taken place at both the technical and strategic levels. The weighted HoQ scores also confirm this finding. Die Neugestaltung der ERP-Systeme unter Mark Schneider mit Fokus auf digitale Transformation und Datenfluss hatte einen deutlichen Einfluss auf die SCI-Ergebnisse.

4.1.7 H7: Transformational leadership fosters sustainable practices

In the literature, characteristics of transformational leaders such as ethics, long-term vision and responsibility towards stakeholders are described as supporting factors for the implementation of corporate sustainability strategies (Bass & Avolio, 1996). The sustainability results (SUS1–SUS5) show a high potential for interaction with the strategic steering power of transformational leadership. MLQ dimension for H7 is in Table 19.

Table 19: MLQ dimension for H7 (Authors’ own elaboration)

MLQ Dimension	Assigned Results
Intellectual Stimulation	SUS1 Environmental Sustainability, SUS3 Economic Sustainability
Idealized Influence	SUS1 Environmental Sustainability, SUS4 Corporate Sustainability Strategy, SUS5 ESG/GRI/SDG Reporting
Individualized Consideration	SUS2 Social Sustainability
Inspiring motivation	SUS2 Social sustainability, SUS4 Sustainability strategy in the company

Mark Schneider's overall influence is 23 points higher compared to Paul Bulcke (Table 20).

Table 20: H7 scores for Hows (Authors' own calculation)

Subcomponent (HOW)	Paul Bulcke	Mark Schneider
SUS1 Environmental sustainability	59	63
SUS2 Social sustainability	45	55
SUS3 Economic sustainability	51	52
SUS4 Sustainability strategy within the company	83	87
SUS5 ESG/GRI/SDG reporting	59	63
Overall score	297	320

These strong correlations within the sustainability results show that leadership influence can unfold synergistically in these areas (Table 21).

Table 21: H7 results from correlation matrix (Authors' own elaboration)

Results of correlation matrix	
SUS1& SUS4	++
SUS2 & SUS5	++
SUS4 & SUS5	++
SUS3 & SUS4	+

H7 hypothesis evaluation: H7 is supported.

During Mark Schneider's tenure, it became clear that the high scores in the MLQ dimensions aligned with sustainability results. Significant progress was made in ESG compliance, company-wide sustainability strategies, and social responsibility, and sustainability was integrated into the institutional decision-making process. This demonstrates that the leadership brought about a sustainable transformation not only rhetorically but also structurally. Both the numerically weighted scores and the overarching correlations show that transformational leadership supports sustainability strategies and that this influence increased significantly during Schneider's tenure.

4.1.8 H8: Sustainability strengthens the strategic effectiveness of ERP systems

The hypothesis focuses on the interaction of two groups of results: sustainability sub-areas (SUS1–SUS5) and ERP system results (ERP1–ERP5).

Correlations and the weighted HoQ scores form the central data basis. Correlations and weighted HoQ scores form the central data basis (Table 22).

Table 22: Correlation pair for H8 (Authors’ own elaboration)

Correlation pair	Value	Comment
US5 ESG reporting & ERP3 Information quality	++	A high-quality database for ESG reports depends on ERP systems
SUS1 Environment & ERP5 Process influence	+	Environmentally conscious processes can be optimized with ERP systems
SUS4 Corporate sustainability & ERP1 System Quality	+	A sustainability strategy requires a viable technical system basis
SUS2 Social Sustainability & ERP4 User Effectiveness	++	Trained users respond more sensitively to social sustainability goals

A total of 25 cross-category correlations were analyzed. These correlations are predominantly positive and significant. The impact of sustainability results on ERP results is particularly pronounced during Mark Schneider's term of office:

- SUS5 → ERP3: Need for high-quality data for ESG reporting
- SUS4 → ERP5: Strategic sustainability provides guidance for ERP processes
- SUS2 → ERP4: Social user impact was specifically promoted

Under Mark Schneider, sustainability goals were strategically aligned with ERP systems. During Paul Bulcke's tenure, however, this integration remained rather limited.

H8 hypothesis evaluation: H8 is supported.

Sustainability strategies – particularly ESG reporting, net-zero plans, and inclusion policies – significantly influenced the direction of ERP system results during Mark Schneider's tenure, especially regarding user-centricity, information accuracy, and process efficiency. This is clearly reflected in both the overarching correlations and the weighted scores. ESG requirements, which demand data transparency and high system quality in ERP, have further strengthened this relationship. It is evident that sustainability measures have strategically transformed ERP systems at both the technical and user levels. The correlations and result weights convincingly support this hypothesis.

4.1.9 H9: Sustainability supports the integration of supply chains

The hypothesis links two main categories: sustainability outcomes (SUS1–SUS5) and supply chain integration outcomes (SCI1–SCI6). The interpretations are based on correlations and the HoQ matrix. A total of 30 correlation pairs were evaluated. The strongest positive relationships are as follows (Table 23):

Table 23: Correlation pair for H9 (Authors’ own elaboration)

Correlation pair	Value	Comment
SUS1 Environment & SCI4 Information exchange	++	Carbon tracking and traceability are based on information exchange
SUS2 Social & SCI2 Contract management	++	Social responsibility is enshrined in contracts
SUS3 Economic & SCI5 Logistics integration	++	Cost- and sustainability-oriented logistics systems are closely linked
SUS4 Corporate Strategy & SCI3 Partnership	++	Strategic sustainability directly influences supplier selection
SUS5 ESG & SCI4 Information Exchange	++	ESG reporting requires transparency in the supply chain

The positive correlations extend across a wide range of SCI results. During Mark Schneider's tenure, the integration of sustainability goals with SCI results was systematically established. Notable correlations include:

SUS1 & SCI4: Environmental sustainability requires an efficient data flow

SUS4 & SCI3: Company-wide sustainability is directly linked to strategic partnerships

SUS5 & SCI6: ESG compliance depends on internal problem-solving mechanisms

During Bulcke's term in office, however, these connections remained more limited and were primarily at the operational level.

H9 hypothesis evaluation: H9 is supported.

The sustainability strategies pursued by Nestlé under Mark Schneider's leadership—such as net-zero targets, GRI-compliant reporting, and social contract mechanisms—were directly reflected in supply chain practices. In particular, issues like transparency, ethical sourcing, and environmental traceability significantly shaped the transformation in the area of SCI (Sustainable Corporate Governance). The HoQ (High Quality of Care) assessment also demonstrates that these sustainability-driven initiatives

structurally align with the SCI results. Correlations and HoQ data clearly showed that sustainability measures—especially through information sharing, partnership structures, and integration into logistics systems—exert a strong influence on the SCI area.

4.1.10 H10: Sustainability practices improve company performance

This hypothesis is based on the interaction between two categories: sustainability outcomes (SUS1–SUS5) and corporate performance indicators (FP1–FP6). The data are derived from cross-category correlation analyses as well as the interactions of weighted HoQ scores. Thirty relationships were evaluated. The strongest correlations are as follows (Table 24):

Table 24: Correlation pair for H10 (Authors’ own elaboration)

Correlation pair	Value	Comment
SUS2 Social sustainability & FP6 Crisis resilience	++	Social support systems increase resilience in crises
SUS3 Economic sustainability & FP2 Profit Margin	++	Economic sustainability is directly related to cost efficiency
SUS4 Corporate Strategy & FP5 Operational Efficiency	+	Sustainable strategies influence efficiency planning
SUS1 Environmental Sustainability & FP3 Market share	+	An environmentally friendly image can strengthen brand loyalty
SUS5 ESG reporting & FP4 Innovation	+	ESG metrics promote the development of new products

This structure shows that sustainability, in its various forms, exhibits positive relationships with FP results (corporate performance). During Schneider's tenure, sustainability results were strategically integrated with corporate performance. Key correlations include:

Table 25: H10 results from correlation matrix (Authors’ own elaboration)

Results of correlation matrix	
SUS2 & FP6	++
SUS3 & FP2	++
SUS5 & FP4	+

During Bulcke's term of office, however, these connections mostly remained limited to local social programs, and no systematic performance relationships could be established.

H10 hypothesis evaluation: H10 is supported.

Mark Schneider has established a direct strategic link between sustainability and corporate performance by integrating areas such as ESG compliance, carbon neutrality targets, and social inclusion into operational KPIs. This has strengthened both the organization's competitiveness and its resilience to crises. The weighted HoQ scores and correlation values quantify and underpin this effect. Sustainability measures have a direct and multidimensional positive impact on business performance. The correlations demonstrate this influence, particularly in the areas of crisis management, profitability, and innovation.

4.2 Overall assessment of the results

This study assessed the influence of transformational leadership on business outcomes using qualitative content analysis based on the MLQ-5X model. These findings were then quantified and systematically structured using the House of Quality (HoQ) methodology.

The results obtained by combining both levels of analysis comprehensively demonstrate the multifaceted impact of transformational leadership characteristics on business performance, sustainability, ERP systems, and supply chain integration.

First, it was found that transformational leadership exerts a strong and positive influence on ERP systems (H1), supply chain integration (H2), and direct business performance (H3). Particularly during Mark Schneider's tenure, high scores were achieved in the MLQ sub-dimensions (II, IM, IS, IC), resulting in more strategic and institutionalized effects on ERP and SCI outcomes. The weighted HoQ scores also quantitatively support this influence.

In a second step, strong positive correlations were also found between ERP systems and business performance (H4), supply chain integration and business performance (H5), and between ERP systems and supply chain integration (H6). These results confirm—in line with previous literature—the relationship between digital systems and organizational efficiency.

The sustainability hypotheses (H7–H10), which represent the original contribution of this study to the literature, demonstrated that transformational leadership systematically promotes sustainability measures (H7). These measures have a positive impact on both ERP systems (H8) and supply chain integration (H9) and ultimately contribute to business performance (H10). Particularly under the leadership of Mark Schneider, it became clear that sustainability strategies were integrated into institutional systems, and this integration is reflected in the HoQ analyses in the form of high weighted scores.

Overall, all hypotheses (H1–H10) were supported by the research findings; no hypothesis had to be rejected. Correlation matrices (Roof), weighted scores, and the combined analysis of MLQ and HoQ provide a sound theoretical and empirical model of the relationship between leadership behavior and technical or strategic outcomes.

In this context, this study, which examines the influence of transformational leadership on multidimensional strategic outcomes, offers not only conceptual but also practical implications. By quantifying the leadership factor within the context of organizational transformation processes, it proposes a model framework that can serve as a guide for decision-makers.

4.3. Validity of the hypotheses and general conclusions

All ten hypotheses developed in this study were supported by theoretical foundations from the literature and quantitative results from the HoQ analysis. In particular, the multifaceted impact of transformational leadership on both technical systems (ERP, supply chain integration) and organizational performance and sustainability was systematically quantified using MLQ scores and weighted HoQ results. This demonstrates that transformational leadership is not merely an abstract concept, but a concrete leadership approach that can demonstrably influence strategic systems within organizations. The validity of each hypothesis is systematically presented below (Table 26):

Table 26: Overall rating for all hypotheses (Authors’ own elaboration)

Hypothesis	Hypothesis content	Evaluation

H1	Transformational leadership has a positive influence on the acceptance of ERP systems.	Supported
H2	Transformational leadership supports the integration of supply chains.	Supported
H3	Transformational leadership directly increases corporate performance.	Supported
H4	ERP systems improve corporate performance.	Supported
H5	Supply chain integration increases corporate performance.	Supported
H6	ERP systems promote supply chain integration.	Supported
H7	Transformational leadership promotes sustainable practices.	Supported
H8	Sustainability strengthens the strategic effectiveness of ERP systems.	Supported
H9	Sustainability supports supply chain integration.	Supported
H10	Sustainability practices increase business performance.	Supported

This table demonstrates that the theoretical framework addressed in the study has gained validity at an analytical level. In particular, the use of the leadership scores determined within the MLQ analysis as "importance values" within the HoQ structure has made it possible to establish quantitative relationships between leadership behavior and strategic output categories.

The assessments conducted show that transformational leadership not only has an impact at the individual level – for example, in human resources or employee satisfaction – but also has a direct effect in structural and strategic areas such as ERP systems, supply chain integration, and sustainability. This leadership approach was implemented more institutionally and systematically, particularly during Mark Schneider's tenure; significant performance improvements were observed in ERP and SCI results.

Furthermore, this work makes an original contribution by demonstrating that incorporating the sustainability dimension strengthens the interaction between transformational leadership and company-wide digital systems in the environmental, social, and economic spheres. This proves that sustainability practices are not merely visions or values, but can also function as performance-enhancing factors in digital infrastructures such as ERP and SCI.

In summary, this set of hypotheses and the insights gained from them impressively demonstrate that transformational leadership has multidimensional strategic effects in today's companies. It becomes clear that leadership understanding must be considered integratively – in conjunction with technical systems, process integration, and corporate responsibility.

To further assess the significance and validity of the results obtained in this work, this section presents a systematic comparison with existing scientific studies. The aim is to reflect on the developed hypotheses in light of previous research literature and to determine the extent to which our empirical findings agree with or differ from earlier findings.

All ten hypotheses of this study are examined individually. For each hypothesis, at least one relevant scientific source is consulted, providing either confirmation or a contrasting assessment. The research methods used, the sectoral relevance, and the results of the comparative studies are presented for each hypothesis. This approach serves not only to provide a theoretical foundation for the study but also to strengthen its relevance to existing research.

This study attempted to statistically validate the developed hypotheses using classical inferential statistical methods such as the Mann-Whitney U test, correlation analyses (Spearman/Pearson), and regression models. The aim was to substantiate the observed relationships between transformational leadership and organizational performance indicators not only theoretically but also empirically.

However, it turned out that the available data – consisting of two comparison units (Paul Bulcke and Mark Schneider) – was insufficient for such statistical procedures. Inferential statistical tests generally require at least 5–10 observation points to ensure a meaningful distribution. Particularly with correlations, both variables must exhibit a certain degree of variance, which was not the case for us (e.g., with constant MLQ scores).

Therefore, a purely statistical analysis was omitted. Instead, a descriptive analysis of the observed differences was conducted, along with a structured comparison with existing study results. This approach is complemented by a qualitative assessment using the House of Quality (HoQ) method, which allows for the systematic identification of causal relationships between leadership dimensions and strategic outputs (Figure 3).

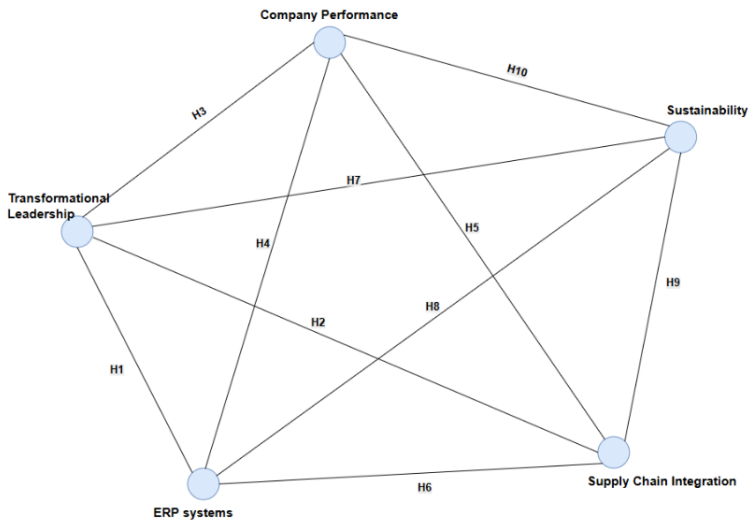


Figure 3: Research model

This research model illustrates the hypothesis-based relationship model of the five central variables in this study. Transformational leadership exerts a direct influence on business performance (H1), ERP systems (H2), supply chain integration (H3), and sustainability (H7). ERP systems, in turn, directly impact business performance (H4) and are reciprocally related to supply chain integration (H6), which also makes a direct contribution to business performance (H5). Sustainability influences both ERP systems (H8) and supply chain integration (H9) and ultimately contributes directly to improving business performance (H10). This network of directed receivers and labeled hypotheses clarifies the sequential logic and multiple pathways of influence of our research model.

The results of this study demonstrate the multifaceted impact of transformational leadership on business performance, sustainability, ERP systems, and supply chain integration. However, this research has certain limitations.

First, the chosen analytical method is largely based on qualitative content analysis and secondary assessment procedures such as the House of Quality (HoQ) method. No primary data, such as surveys, interviews, or direct observations, were collected, which means that organizational perceptions and the immediate effects of leadership behavior on employees could not be directly captured. This increases the dependence of the results on an interpretive framework and limits their empirical generalizability.

Secondly, the analysis is limited to a single company (Nestlé) and two leadership periods. Although Nestlé, as a globally operating company with comprehensive documentation, represents a suitable case study, the focus on only one industry limits the generalizability of the findings. Aspects such as sectoral diversity, different organizational structures, or cultural contexts could not be considered. For future research, it would therefore be desirable to comparatively analyze transformational leadership in other sectors—for example, in the technology sector, healthcare, or the public sector.

Third, in this study, the MLQ scores were not interpreted using classical quantitative methods (e.g., standardized surveys), but rather based on publicly available sources and qualitative content analyses. While this methodological approach offers an innovative way to systematically evaluate leadership behavior in historical contexts, it raises potential methodological questions regarding objectivity and measurement accuracy. Against this background, the following recommendations can be made for future studies:

Several case studies should be conducted to comparatively examine the effects of transformational leadership in different companies. Mixed-methods approaches are recommended, combining quantitative methods (e.g., MLQ questionnaires, structural equation modeling) with qualitative methods (e.g., interviews, case studies) to more comprehensively capture the multidimensional nature of leadership. Ethnographic studies and long-term observations focusing on the employee perspective would be highly beneficial for exploring the connections between sustainability, ERP systems, and supply chain integration. Furthermore, the impact of transformational leadership could be investigated in organizations with varying levels of maturity—for example, between SMEs and multinational corporations—in order to derive scalable strategic implications.

Another promising area of research arises from the changing world of work after the pandemic: topics such as remote leadership, digital leadership and the redefinition of transformational leadership approaches under changed conditions are becoming increasingly important.

In summary, this work provides a holistic framework for analyzing the impact of transformational leadership on organizational systems. However, given the dynamic and context-dependent nature of leadership, further studies with broader samples and interdisciplinary approaches are needed to fully grasp the phenomenon in its entirety.

5 CONCLUSION

While transformational leadership is often associated in the literature with aspects such as employee motivation or organizational commitment, this study takes a broader approach. It also examines the impact of this leadership style on sustainability practices, ERP systems, and supply chain integration. Using Nestlé as a case study, a comparative analysis was conducted to illustrate the multidimensional effects of leadership on corporate performance.

To assess these effects, the Multifactor Leadership Questionnaire (MLQ) model was used at the theoretical level. At the application level, the House of Quality (HoQ) method was employed to quantify the impact of transformational leadership on organizational outputs. The focus was on two distinct CEO periods at Nestlé: the tenures of Paul Bulcke and Mark Schneider. The analysis concentrated on four key areas of impact: corporate performance, sustainability, ERP systems, and supply chain integration.

While the MLQ served as a conceptual instrument for the qualitative analysis of leadership behavior, the HoQ method enabled a systematic measurement of the impact on strategic corporate outputs. This integrative approach allowed the multifaceted effects of transformational leadership to be captured and presented in a structured, comparative, and quantitative manner.

The analyses conducted show that transformational leadership has a significant impact on both technical systems (ERP, supply chain integration) and strategic outputs (corporate performance, sustainability). The MLQ analysis revealed particularly high leadership scores during Mark Schneider's tenure; modeling these scores in the HoQ matrix showed significant performance improvements in areas such as sustainability, digital integration, and organizational efficiency. All hypotheses (H1–H10) were supported by the study results. This confirms that transformational leadership not only influences employees but also significantly shapes an organization's digital and sustainable transformation capabilities. Significant differences were observed between Paul Bulcke's traditional, internally process-focused leadership style and Mark Schneider's transformation-oriented, strategic leadership.

This dissertation not only makes a theoretical contribution but also models the effects of transformational leadership on an analytical level, thus providing valuable insights for both academic research and practice. From a literary studies perspective, this study is one of the few works that examines the relationships between transformational leadership and ERP systems, supply chain integration, and sustainability within a holistic framework. Furthermore, it proposes a methodological alternative to the classic application of MLQ (Microsoft

Leadership Quality) by employing a content-analytical approach, thereby enabling a comparative evaluation of historical leadership periods.

From a practical perspective, this work represents a strategic instrument for decision-makers, visualizing the impact of transformational leadership not only on motivation, but also on digital transformation, process integration, and sustainable strategies.

References

- Akao, Y. (1990). *Quality function deployment: integrating customer requirements into product design*. Productivity press.
- Aniebonam, E. E., Chukwuba, K., Emeka, N., & Taylor, G. (2023). Transformational leadership and transactional leadership styles: systematic review of literature. *International Journal of Applied Research*, 9(1), 07-15.
- Assefa, Y., Moges, B. T., & Tilwani, S. A. (2025). The psychometric properties of teacher leadership scale: the study of adaptation and validation based on the MLQ-5X in higher education. *Journal of Applied Research in Higher Education*, 17(2), 719-737.
- Bass, B. M., (1985). *Leadership and performance beyond expectations*. New York: Free press.
- Bass, B. M., & Avolio, B. J. (1996). Multifactor leadership questionnaire. *Western Journal of Nursing Research*.
- Burns, J.M., 1978. *Leadership*, Harper & Row, New York.
- Chan, L. K., & Wu, M. L. (2002). Quality function deployment: A literature review. *European journal of operational research*, 143(3), 463-497.
- Chouinard, Y., & Stanley, V. (2013). *The responsible company: What we've learned from Patagonia's first 40 years*. Patagonia.
- De Soete, W. (2016). Towards a multidisciplinary approach on creating value: Sustainability through the supply chain and ERP systems. *Systems*, 4(1), 16.
- Elkhani, N., Soltani, S., & Nazir Ahmad, M. (2014). The effects of transformational leadership and ERP system self-efficacy on ERP system usage. *Journal of Enterprise Information Management*, 27(6), 759-785.
- Fairholm, G. W. (1994). *Leadership and the culture of trust*. Bloomsbury Publishing USA.
- Faradina, R., & Maburur, A. (2023). Role of Transformational Leadership in the Successful Implementation of Information Systems in the Government Sector. *Jurnal Pajak Dan Keuangan Negara (PKN)*, 5(1), 163-174.
- García-Morales, V. J., Lloréns-Montes, F. J., & Verdú-Jover, A. J. (2008). The effects of transformational leadership on organizational performance through knowledge and innovation. *British journal of management*, 19(4), 299-319.
- Hauser, J. R., & Clausing, D. (1988). The house of quality. *Harvard Business Review*, 66(3), 63-73.

- Hilton, S. K., Madilo, W., Awaah, F., & Arkorful, H. (2023). Dimensions of transformational leadership and organizational performance: the mediating effect of job satisfaction. *Management Research Review*, 46(1), 1-19.
- Hu, H., Gu, Q., & Chen, J. (2013). How and when does transformational leadership affect organizational creativity and innovation? Critical review and future directions. *Nankai Business Review International*, 4(2), 147-166.
- Ibrahim, A. U., & Daniel, C. O. (2019). Impact of leadership on organisational performance. *International Journal of Business, Management and Social Research*, 6(2), 367-374.
- Maloni, M. J., & Brown, M. E. (2006). Corporate social responsibility in the supply chain: an application in the food industry. *Journal of business ethics*, 68(1), 35-52.
- Mazur, G. H. (1993, June). QFD for service industries. In *The Fifth Symposium on QFD, Michigan*. Available at: https://mazur.net/pdfs/qfd_service.pdf (Accessed 20.05.2025).
- Mutahar, A. Y., Rasli, A. M., & Al-Ghazali, B. M. (2015). Relationship of transformational leadership, organizational learning and organizational performance. *International Journal of Economics and Financial Issues*, 5(1), 406-411.
- Nestlé (2013). *Annual Report 2013*. https://www.nestle.com/sites/default/files/asset-library/documents/library/documents/annual_reports/2013-annual-report-en.pdf (Accessed on: 20.05.2025)
- Nestlé (2014a). *Annual Report 2014*. <https://www.nestle.com/sites/default/files/2014-annual-report-en.pdf> (Accessed on: 20.05.2025)
- Nestlé (2014b). *Nestlé in society Creating Shared Value and meeting our commitments 2014*. Full report. <https://www.nestle.com/sites/default/files/asset-library/documents/creating-shared-value/nestle-csv-full-report-2014-en.pdf> (Accessed on: 22.05.2025)
- Nestlé (2018). *Creating Shared Value Report and meeting our commitments 2018*. Progress report. https://www.nestle.com/sites/default/files/asset-library/documents/library/documents/corporate_social_responsibility/creating-shared-value-report-2018-en.pdf (Accessed on: 23.05.2025)
- Nestlé (2019a). <https://www.nestle.com/media/mediaeventscalendar/allevnts/cccp-ceo-investor-forum-2019> (Accessed on: 23.05.2025)
- Nestlé (2019b). *Creating Shared Value and meeting our commitments . Nestlé Annual Review extract 2019*. <https://www.nestle.com/sites/default/files/2020-03/nestle-creating-shared-value-summary-report-2019-en.pdf> (Accessed on: 23.05.2025)

- Nestlé (2020). *Creating Shared Value and Sustainability Report 2020*. <https://www.nestle.com/sites/default/files/2021-03/creating-shared-value-report-2020-en.pdf> (Accessed on: 21.05.2025)
- Nestlé (2024a). *About Us* [online], Available at: <https://www.nestle.com/about/overview> (Accessed on: 19. 05. 2025).
- Nestlé (2025a). <https://www.nestle.com/investors/corporate-governance/honorary-chairman> (Accessed on: 20.05.2025)
- Nestlé (2025b). <http://nestle.com/investors/corporate-governance/former-executive-board-members/mark-schneider> (Accessed on: 20.05.2025)
- Nestlé (2025c). <https://www.nestle.com/> (Accessed on: 15.05.2025)
- Nestlé Health Science (NHS) (2025). <https://www.nestlehealthscience.com/> (Accessed on: 15.05.2025)
- Orabi, T. G. A. (2016). The impact of transformational leadership style on organizational performance: Evidence from Jordan. *International Journal of Human Resource Studies*, 6(2), 89-102.
- Pawar, A. (2016). Transformational leadership: inspirational, intellectual and motivational stimulation in business. *International Journal of Enhanced Research in Management & Computer Applications*, 5(5), 14-21.
- Rahmatullah, A., Ramdansyah, A. D., Kambara, R., & Haryadi, D. (2022). Improving organizational performance with organizational culture and transformational leadership through intervening organizational commitment variables. *Dinasti International Journal of Digital Business Management*, 3(2), 180-194.
- Samad, S. (2012). The influence of innovation and transformational leadership on organizational performance. *Procedia-Social and behavioral sciences*, 57, 486-493.
- Setiabudi, K. J., Siagian, H., & Tarigan, Z. J. H. (2021). *The effect of transformational leadership on firm performance through ERP systems and supply chain integration in the food and beverage industry* (Doctoral dissertation, Petra Christian University).
- Shao, Z., Feng, Y., & Liu, L. (2012). The mediating effect of organizational culture and knowledge sharing on transformational leadership and Enterprise Resource Planning systems success: An empirical study in China. *Computers in Human Behavior*, 28(6), 2400-2413.
- Suwanto, S., Sunarsi, D., & Achmad, W. (2022). Effect of transformational leadership, servant leadership, and digital transformation on MSMEs

performance and work innovation capabilities. *Central European Management Journal*, 30(4), 751-762.

Takahashi, H. (2021). *Everything Originated from Milk: Case Study of Nestlé*. World Scientific.

Weston, P., & Nnadi, M. (2023). Evaluation of strategic and financial variables of corporate sustainability and ESG policies on corporate finance performance. *Journal of Sustainable Finance & Investment*, 13(2), 1058-1074.

Zairi, M., & Youssef, M. (1995). A review of key publications on benchmarking part I. *Benchmarking for Quality Management & Technology*, 2(1), 65-72.

Zucchella, A., Hagen, B., & Serapio, M. G. (2023). *International entrepreneurship*. Edward Elgar Publishing.

CHAPTER 9

Scenario-Based Production and Capacity Planning Under Uncertainty: Pharmaceutical Industry

**Yağmur Erkan¹ & Mehmet Ali Taş² &
Batin Latif Aylak³**

¹ Turkish-German University, ORCID: 0009-0004-0751-8037

² Re. Asst., Turkish-German University, ORCID: 0000-0003-3333-7972

³ Assoc. Prof. Dr., Turkish-German University, ORCID: 0000-0003-0067-1835

1. Introduction

One of the greatest challenges facing businesses in the 21st century is adapting to a rapidly changing and increasingly uncertain global competitive environment. Economic volatility triggered by globalization and external shocks leads to unpredictable fluctuations in fundamental parameters such as demand and capacity, making uncertainty management inevitable for businesses (Klibi et al., 2010; Snyder & Shen, 2019). In the literature, these uncertainties are classified as environmental and internal factors, and it is stated that both types of uncertainty have significant effects on costs, productivity, and service levels.

These effects are felt particularly strongly in the pharmaceutical industry. Seasonal demand fluctuations, epidemics, and demographic trends are significant sources of uncertainty on the demand side, while technical failures, raw material shortages, and capacity constraints are significant sources of uncertainty on the supply side (Govindan et al., 2015; Queiroz et al., 2022). Furthermore, the sector has not only an economic responsibility but also an ethical responsibility in terms of the accessibility of vital products. The COVID-19 period has clearly demonstrated that cost-focused planning alone is not sufficient and that resilience has become a strategic necessity (Wang & Jie, 2020; Queiroz et al., 2022).

This structure enables a detailed analysis of system behavior in terms of cost-service balance under uncertainty and reveals how strategies change, particularly in stressful situations such as a pandemic. Thus, the model identifies which decision variables are sensitive to scenarios and which are more resilient, providing a practical, flexible, and applicable decision support tool for the pharmaceutical industry.

In this regard, the research questions of the study are as follows:

- How can capacity expansion, outsourcing, and investment decisions be optimized together under uncertainty?
- How do different scenarios affect total costs and service levels?
- Which decision variables are more resilient to uncertainty, and which are more sensitive?

The remainder of the study defines the problem and scope in the first section; reviews the relevant literature in the second section; presents the MILP model in detail in the third section; discusses scenario analyses in the fourth section; and summarizes the findings, provides evaluations, and makes recommendations for the future in the final section.

2. Literature Review

The pharmaceutical industry faces a high level of uncertainty in production and capacity planning due to its complex process structure and fragile supply chains. Studies in this field evaluate the impact of uncertainty on decision-making mechanisms from different perspectives. For instance, Gökalp (2017) demonstrated that stochastic optimization can reduce waiting times and costs in critical healthcare decisions. Nguyen et al. (2018) revealed that deterministic capacity models can perform similarly to stochastic models under certain conditions. Nasrabadi et al. (2020) increased service availability under uncertainty by combining queueing theory with robust optimization. Furthermore, Peltokorpi (2010) and Plantin & Johansson (2012) emphasize that improvements in planning and control structures increase efficiency.

Uncertainty in drug production is becoming even more pronounced due to demand volatility, capacity constraints, regulatory delays, and supply disruptions. For this reason, scenario-based and stochastic approaches have gained significant attention in the literature. Lindahl et al. (2024) show that the combined use of deterministic and stochastic methods in capacity expansion decisions improves resource allocation, while Vila-Parrish et al. (2012) show that two-stage models improve service quality for perishable drugs. Under pandemic conditions, Pathy & Rahimian (2023) analyzed how different risk attitudes affect production decisions. Blossey et al. (2022) argue that uncertainty in regulatory approval processes is decisive for strategic decisions such as capacity, location selection, and outsourcing. In addition, Rossetti et al. (2011), Uthayakumar and Priyan (2013), Marques et al. (2017), Tucker (2020), Kochakkashani et al. (2023), Vann Yaroson et al. (2024), and Bilal et al. (2024) emphasize the increasing importance of resilience, flexibility, and integrated planning in the pharmaceutical supply chain.

Within the framework of decision variables, studies in the literature can be divided into three main groups:

1. Operational models: These models focus on production, inventory, and backorder dynamics (Uthayakumar & Priyan, 2013).
2. Tactical models: This group address capacity management, flexibility, and service level targets (Gökalp, 2017; Nguyen et al., 2018).
3. Strategic models: These models examine long-term investment, capacity expansion, and outsourcing (Blossey et al., 2022; Lindahl et al., 2024).

Despite significant contributions in the existing literature, some gaps remain:

- Studies that comprehensively address all critical variables—such as production, inventory, backorders, investment, contract manufacturing, and service level—within the same model are quite limited.
- Demand and capacity uncertainty are often modeled independently of each other.
- Although service level is a critical performance indicator for the pharmaceutical industry, it is not explicitly defined as a constraint in many models.

This study addresses these gaps by bringing all key decision variables together in a single integrated MILP model; comprehensively modeling uncertainty by creating four scenarios that jointly represent demand and capacity; directly applying the service level constraint; and providing detailed scenario analyses. The following sections describe the model structure, data set, and solution method in detail.

3. Methodology

This section provides an overview of the methodological approach, scenario structure, dataset, and optimization model employed in the study. The mathematical model and its Pyomo implementation were developed with support from ChatGPT (OpenAI, version 4), while the scenarios and dataset were synthetically generated using Python. All modeling components were subsequently reviewed by the author to ensure scientific rigor and methodological coherence.

The study investigates production and capacity planning under uncertainty in the pharmaceutical industry through a scenario-based framework that integrates scenario planning, stochastic modeling, and MILP. In an environment characterized by high demand volatility and frequent capacity bottlenecks, deterministic approaches fall short. MILP, by contrast, enables the simultaneous optimization of service level, capacity expansion, investment, inventory, and outsourcing within a single coherent structure (Clautiaux & Ljubić, 2025). The model is constructed on the capacity–inventory framework of Dang and Elgersma (2020) and incorporates the scenario-based MILP structure proposed by Lindahl et al. (2023). This integration allows for joint optimization of investment, outsourcing, inventory, and unmet demand decisions, while embedding the service-level requirement directly into the planning process. The four scenarios are described both qualitatively and quantitatively in Table 1, and their associated probabilities are derived from sectoral assessments presented in Table 2. As no real company data was available, demand and capacity inputs were synthetically generated to reflect typical uncertainty patterns observed in the pharmaceutical sector.

Table 1: Definition of Scenarios (Authors' own elaboration)

Scenario Code	Scenario Name	Description
S1	Normal case	Reference scenario assuming average demand and average capacity levels.
S2	Pandemic-related double shock	Emergency scenario in which demand doubles and capacity is reduced by 25%.
S3	Seasonal summer slump	Scenario of the seasonal effect, in which demand falls by 50% in June, July, and August.
S4	Raw materials supply problem	Disruption of the supply chain with a 70% reduction in capacity for antibiotics alone.

Table 2: Scenario probabilities (Authors' own elaboration)

Scenario	Probability (π_i)
S1	0.35
S2	0.15
S3	0.30
S4	0.20

The model covers three product groups—antibiotics, painkillers, and vitamins—and a six-month planning period from May to October. The study defines a total of four scenarios, and demand and capacity values are synthetically generated.

For demand, the average value for each product (antibiotic: 1000, painkiller: 1200, vitamin: 500) was determined and sampled from a normal distribution with a standard deviation of 100. Capacity values were simulated using a normal distribution with a standard deviation of 90, based on the average monthly capacity for each product (1100, 1200, 1500). Table 3 shows the demand simulation results, while Table 4 shows the capacity simulation results.

Table 3: Demand data (Authors' own elaboration)

	Normal case			Pandemic-related double shock		
	Antibiotic s	Painkiller s	Vitami n	Antibiotic s	Painkiller s	Vitami n
May	986	1352	1477	1608	2439	3034
June	1077	1254	1453	1940	2256	3211
July	809	1144	1531	1647	2323	3122
August	859	1177	1358	2186	2338	3195
Septembe r	1011	1238	1471	1963	2161	3271
October	1185	1094	1378	2201	2271	3308
	Seasonal summer slump			Raw materials supply problem		
	Antibiotic s	Painkiller s	Vitami n	Antibiotic s	Painkiller s	Vitami n
May	1156	1282	1470	993	1197	1746
June	401	618	724	1030	1083	1575
July	475	616	776	909	1060	1719
August	548	584	677	943	1150	1507
Septembe r	1026	1177	1458	1047	1355	1468
October	920	1240	1517	877	1331	1518

Table 4: Capacity data (Authors' own elaboration)

	Normal case			Pandemic-related double shock		
	Antibiotics	Painkillers	Vitamin	Antibiotics	Painkillers	Vitamin
May	1145	1358	1279	844	855	1041
June	1242	1258	1258	815	842	934
July	1122	1145	1209	856	1004	914
August	1018	1432	1306	918	899	1005
September	1051	1196	1246	782	875	1048
October	1046	1299	1374	819	1008	1008
	Seasonal summer slump			Raw materials supply problem		

	Antibiotics	Painkillers	Vitamin	Antibiotics	Painkillers	Vitamin
May	1097	1064	1308	353	1127	1305
June	1108	1280	1433	313	1297	1403
July	1027	1382	1252	401	1426	1353
August	1109	1237	1265	241	1309	1160
September	1127	1300	1173	234	1217	1230
October	1069	1285	1470	403	1320	1155

The cost parameters used in the model were determined based on industry standards and are listed in Table 5. Further input parameters required for the model are listed in Table 6.

Table 5: Cost parameters (€) (Authors' own elaboration)

Parameters	Antibiotics	Painkillers	Vitamin
Production costs	15	11	8
Labor costs	18	14	11
Inventory costs	2	1.5	1
Backlog penalty	40	30	20
Fixed investment costs	10000	10000	10000

Table 6: Other input parameters (Authors' own elaboration)

Parameter	Value	Description
Initial inventory	300	The amount of inventory initially available for each product and each scenario.
Maximum capacity increase	Antibiotics: 500 Painkillers: 600 Vitamin: 700	Maximum capacity increase that can be achieved per product when investments are made.
Alpha	0.8	Target for service level; the aim is to cover at least 80% of total demand.
Wage production capacity		The limit of external production capacity for each scenario. Average 70% of capacity. (52.5% in the pandemic scenario). Normal distribution, standard deviation is 90).

The mathematical model used in this section of the paper is explained. The sets of the mathematical model are shown in Table 7. The decision variables of the mathematical model are also shown in Table 8. The parameters of the mathematical model are shown in Table 9, respectively.

Table 7: Sets (Authors' own elaboration)

Set	Significance
P	Products
T	Periods
S	Scenarios

Table 8: Decision variables (Authors' own elaboration)

Symbol	Significance
$x_{p,t,s}$	Number of internally produced units of product p per month t under scenario s
$x_{p,t,s}^{lohn}$	Number of units of product p produced externally (through contract manufacturing) in month t in scenario s
$i_{p,t,s}$	Inventory of product p at the end of month t under scenario s
$b_{p,t,s}$	Backorder of product p in month t under scenario s
$u_{p,t,s}$	Increased capacity for product p in month t and scenario s
$x_{p,s} \in \{0, 1\}$	Binary variable: 1 if investing in product p in scenario s ; otherwise 0

Table 9: Parameters (Authors' own elaboration)

Symbol	Significance
$d_{p,t,s}$	Demand for product p in month t under scenario s
c_p	Unit costs for internal production of product p
c_p^{lohn}	Unit costs for external production (contract manufacturing) of product p

h_p	Inventory costs per unit of product p
b_p	Penalty costs per unit for product p backlogs
f_p	Fixed costs for an investment in product p
k_p	Maximum capacity expansion per investment for product p
$z_{p,0}$	Initial stock of product p at the beginning of the planning period
$cap_{p,t,s}$	Available internal production capacity for product p in month t , scenario s
π_s	Probability of scenario s occurring
α	Target value for minimum service level
$L_{p,t,s}$	External production capacity limit (wages) for product p , month t , scenario s
μ_p^{cap}	Average internal production capacity for product p regardless of scenarios
θ^{cap}	The fixed standard deviation 90 for calculating the external production limit for all products

The objective function aims to minimize the total cost weighted by scenario probabilities (Equation (1)). Total cost consists of production, outsourcing, inventory holding, order delay penalties, and fixed investment costs incurred when investing.

$$\min \sum_{s \in S} \pi_s \cdot \left[\sum_{p \in P} \left(\sum_{t \in T} (c_p \cdot x_{p,t,s} + c_p^{lohn} \cdot x_{p,t,s}^{lohn} + h_p \cdot i_{p,t,s} + b_p \cdot b_{p,t,s}) + f_p \cdot y_{p,s} \right) \right] \quad (1)$$

This condition controls whether production, external production, and initial inventory cover total demand in each scenario and period for product p . Demand must be balanced in the form of inventory and backlog (Eq. (2)):

$$\begin{cases} z_{p,0} + x_{p,1,s} + x_{p,1,s}^{lohn} = d_{p,1,s} + i_{p,1,s} - b_{p,1,s} & \forall t = 1 \\ i_{p,t-1,s} + x_{p,t,s} + x_{p,t,s}^{lohn} = d_{p,t,s} + i_{p,t,s} - b_{p,t,s} & \forall t > 1 \end{cases} \quad (2)$$

In-house production may not exceed the sum of existing capacities and capacity expansions carried out in previous periods (Eq. (3)):

$$x_{p,t,s} \leq cap_{p,t,s} + \sum_{\tau=1}^t u_{p,\tau,s} \quad (3)$$

contract manufacturing capacity $L_{p,t,s}$ for each product p , month t , and scenario s is modeled as a random variable with normally distributed dispersion: $L_{p,t,s} \sim \mathcal{N}(\mu_p^{cap}, \gamma_s, \theta^{cap^2})$. The reduction factor γ_s is 0.7 for normal scenarios and 0.525 for the pandemic scenario. The standard deviation (σ) is 90. Negative realizations are set to zero (Eq. (4)):

$$\sum_{p \in P} \sum_{t \in T} x_{p,t,s}^{lohn} \leq L_{p,t,s} \quad (4)$$

For the relationship between investment and capacity increase, capacity can only be increased for invested products (Eq. (5)). When $y_{p,s} = 0$, $u_{p,t,s} = 0$.

$$u_{p,t,s} \leq k_p \cdot y_{p,s} \quad (5)$$

According to the maximum of two investment restrictions, each scenario established by Eq. (6) includes the condition that investments can be made in a maximum of two different products.

$$\sum_{p \in P} y_{p,s} \leq 2 \quad (6)$$

Eq. (7) In Scenario s , the total unmet demand (order backlog) must not exceed the specified service level threshold (α). is established for the Service Level Constraint.

$$\sum_{p \in P} \sum_{t \in T} b_{p,t,s} \leq (1 - \alpha) \cdot \sum_{p \in P} \sum_{t \in T} d_{p,t,s} \quad (7)$$

Subsequently, an investment ban constraint has been established in the raw material crisis scenario in Eq. (8). In the raw material crisis scenario, it is assumed that investment will be halted for all products because there are no raw materials in the system and capacities cannot be increased even if investments are made.

$$y_{p,s} = 0 \text{ when } s = \textit{Rohstoffversorgungsproblem} \bigwedge p = \textit{Antibiotikum} \tag{8}$$

After establishing and explaining the model, the implementation phase of the model with Pyomo is initiated. The numerical solution of the model was performed using open-source Python libraries. Mathematical modeling was carried out using the Pyomo library (Python Optimization Modeling Objects), and the model was solved with the CBC solver. The technical software and infrastructure used in modeling are listed in Table 10.

Table 10: The software information (Authors’ own elaboration)

Programming language	Phyton
Modeling library	Pyomo
Analyzer	CBC (Coin-or Branch and Cut)
Data processing	Pandas
Data format	Reading/writing data via Excel (.xlsx)
Simulation	NumPy

By developing and implementing the model in the Google Colab environment, a cloud-based and repeatable modeling process has been achieved. The Python code that generates demand and capacity data, defines the model, and solves it has not been included in the paper to avoid increasing its length and to ensure a reader-friendly experience (upon request, the relevant code can be sent to the researcher's contact address).

4. Results and discussion

In this section, the results obtained by the developed model are compared with the findings of similar studies in the literature; system behavior is evaluated based on key indicators such as total cost, service level, investment decisions, capacity utilization, and production distribution.

4.1 Overall results and comparison of scenarios

The model was run separately for four scenarios, and the performance results were compared based on the indicators presented in Table 11.

Table 11: Results of the model (Authors' own elaboration)

Indicator	S1	S2	S3	S4
Production	20.159	34.487	15.765	16.416
Contract production	795	6.091	0	5.192
Stock	184	255	379	249
Backorder	0	2.995	0	0
Capacity increase	0	7.800	0	0
Investments	0	2	0	0
Service level	100%	93,27%	100%	100%
Total cost	230.528 €	576.849	171.732	247.990

S1 – Normal Situation: The system operated without any disruptions and required only a limited level of external production (795 units). There were no delays, and the service level was maintained at 100%. The total cost was €230,528.

S2 – Pandemic: This scenario represents the situation where the highest pressure was felt on the system. Due to the increase in demand and the decrease in capacity, investments were made, external production was used intensively (6,091 units), and a total of 7,800 units of additional capacity was required. Despite all interventions, a delay of 2,995 units occurred, the service level dropped to 93%, and the total cost reached €576,849, the highest level among all scenarios.

S3 – Seasonal Demand Decline: This is the scenario where the model performed best. Due to low demand, there was no need for investment or external production; a 100% service level was achieved without any delays. The total cost was €171,732, the lowest among all scenarios.

S4 – Raw Material Supply Problem: The investment ban imposed on antibiotic products limited capacity expansion opportunities and led to an

increase in external production (5,192 units). However, the low total order volume allowed the service level to be maintained at 100%.Figure 1 compares the production strategies, inventory levels, and order backlogs of the four scenarios. The pandemic scenario, in particular, clearly demonstrates that capacity constraints play a critical role in both investment and external production decisions. The following Figure 2 shows the service level and the total cost ratio in each scenario.

Overall, the best performance was achieved in the scenario where seasonal demand declined. In contrast, the pandemic scenario produced the highest cost and lowest service level due to simultaneous high demand and low capacity constraints. This comparison clearly demonstrates the impact of the production strategies adopted by the system under different uncertainty conditions on performance.

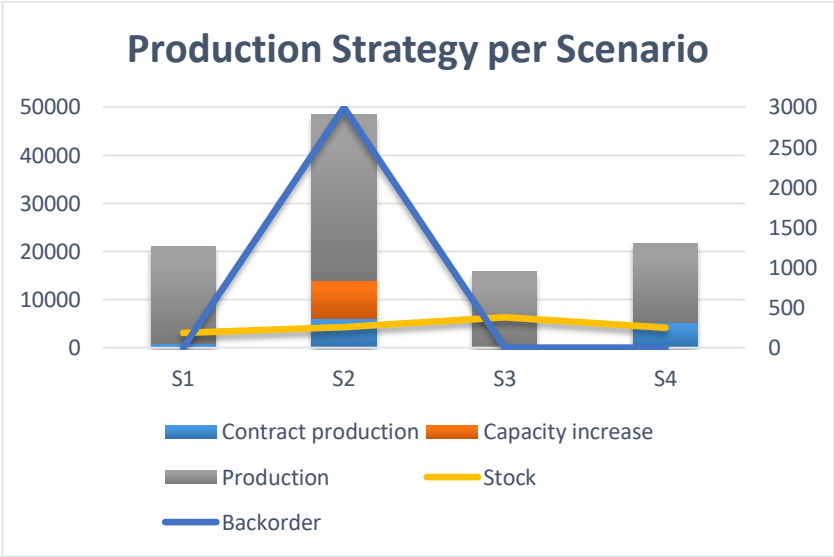


Figure 1: Production strategy per scenario (Authors’ own elaboration)

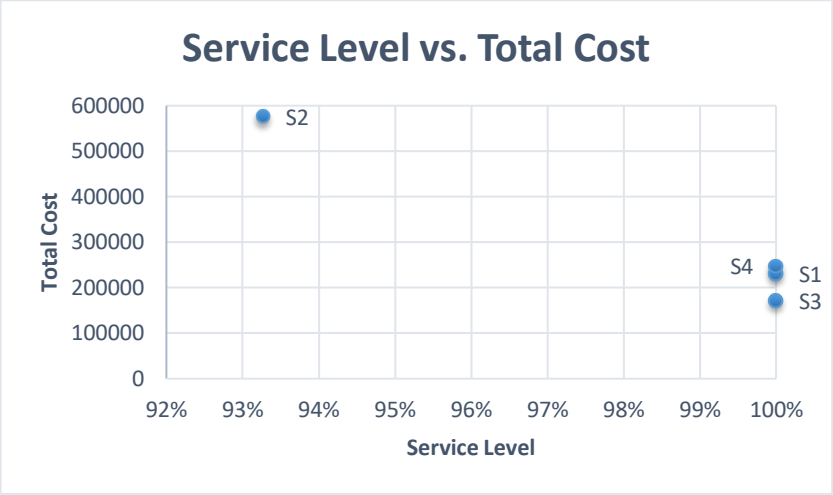


Figure 2: Service Level vs. Costs (Authors’ own elaboration)

4.2 Component-based cost analysis

This section analyzes and compares the effects of the components that influence the total costs based on the scenarios. The aim is to observe the variability of the cost structure depending on the conditions. The following table (Table 12) shows the cost components of each scenario. Additionally, the contribution of each scenario to the objective function is shown by weighting it according to the scenario probabilities.

Table 12: Cost components (Authors' own calculation)

Cost type	S1	S2	S3	S4	Weighted cost
Production cost	220.845	348.904	171.353	165.538	214.144,85
Wage cost	9.438	94.723	0	82.023	33.916,35
Storage cost	245	511	379	429	361,90
Backlog cost	0	112.711	0	0	16.906,65
Investment cost	0	20.000	0	0	3.000,00
Total cost	230.528	576.849	171.732	247.990	268.329,75
Weighted cost	80.684,8 0	86.527,35	51.519,6 0	49.598,0 0	268.329,75

The highest total costs were observed in the pandemic scenario (S1) at €576,849. The largest cost factor after production in this scenario is production costs (€348,904). In this scenario, the system was applied to all resources but was still insufficient. The second-highest total costs occurred in the raw material supply problem scenario (€247,990). Since investments in antibiotic production could not be made in this scenario, the dependence on labor production increased (€82,023). Under normal circumstances, neither penalties for delays nor investment costs are incurred. 96% of the total costs of €230,528 are production costs. Labor production is at a low level (€9,438). The lowest total costs are observed in the seasonal decline scenario (€171,732). Since demand has decreased significantly over a specific period, there is no need for additional resources; there is no external production, investment, or backorder. Figure 3 shows the influence of the cost components on the total costs on a scenario basis.

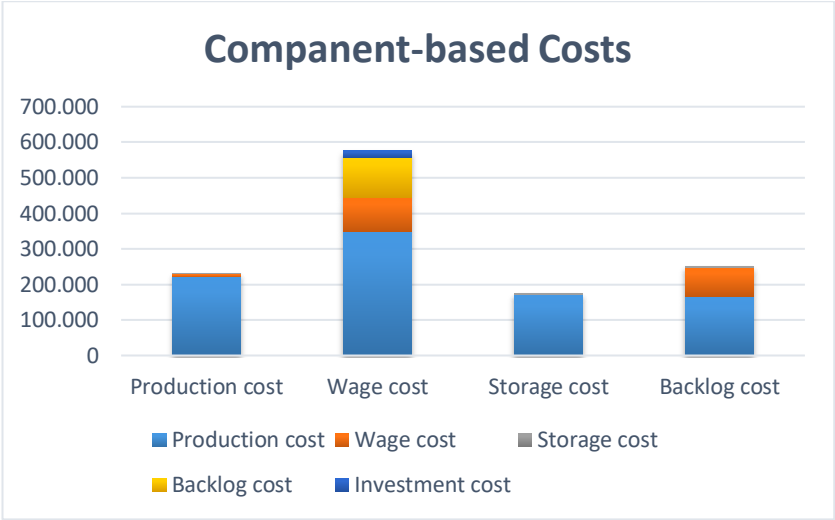


Figure 3: Component-based costs (Authors’ own elaboration)

4.3 Investment decisions and their effects

This section describes the scenarios and products under which investment decisions are made, as well as the reasons for and consequences of these decisions, and analyzes how the system reacts when the investment decision is completely reversed. According to the model results, investments in painkillers and vitamins were made only during the pandemic-induced double shock. Details of the invested products are listed in Table 13.

Table 13: Investment decisions (scenario- and product-based) (Authors’ own calculation)

Scenario	Antibiotic	Painkiller	Vitamin	Total investment
S1	0	0	0	0
S2	0	1	1	2
S3	0	0	0	0
S4	0	0	0	0

As a result of the investments, the relevant products were increased by the maximum capacity specified in the model specifications, and these investment decisions were implemented continuously for six months. Details can be found in Table 14.

Table 14: Capacity increase on a product basis (Authors' own calculation)

Scenario	Product	Monthly increase	Total increase (6 months)
S2	Painkillers	600	3600
S2	Vitamin	700	4200

In the pandemic scenario, demand for products far exceeds available capacity. Contract manufacturing capacity is limited (defined as 52.5%), again due to pandemic conditions. Penalties for backorders impose significantly higher costs on the system in the long run. Although the fixed investment costs (€10,000) are high, it offers a more economical long-term solution than external production or reordering.

4.3.1 Scenario “No Investments” (Pandemic)

In this subsection, an alternative scenario is created for the Pandemic Double Shock (S2) scenario only, by completely disabling the investment decision variable ($y_{p,s}$) defined in the model. This analysis is considered a sensitivity analysis to measure how the system would react without the investment option.

Accordingly, all investment variables in the model were fixed at $(y_{p,s}) = 0$, and the model was solved again. The model became unfeasible when it was forced to maintain the service level at 80% if the investment option was removed. Therefore, this analysis removes the service level constraint from the model. This demonstrates that the investment option is critical not only in terms of cost but also feasibility. The results of the model without investment are presented in Table 15.

Table 14: Summary of the results of the "No Investments" scenario (Authors' own calculation)

	Antibiotics	Painkillers	Vitamin	Total
Production	5.034	5.483	5.950	16.467
Contract Production	3.599	4.075	4.286	11.960
Backorder	2.612	3.930	8.605	15.147
Inventory	256	0	0	256
Service Level	77 %	71 %	55 %	66 %
Total cost				747.385 €

According to the results summarized in Table 15, total costs were €170,536 higher than in the original model with investments (€576,849 in the model with investments). The service level dropped significantly, reaching 66%. Contract manufacturing was used very intensively (11,960 units) because further capacity

expansion was not possible. For the product antibiotics, in addition to in-house production, significant outsourcing and a backlog were observed. Inventory was only maintained for this product. The overall results are as follows: The system attempts to meet demand through contract manufacturing and backlogs. However, this solution is very costly and reduces customer satisfaction. This analysis has numerically demonstrated that investment decisions play a crucial role in the model.

4.4 Analysis of the robustness of the decision variables

This section examines how model decisions differ across scenarios. To assess scenario-dependent variability, normalized variance in the range of 0–1 was calculated for each decision variable. Figure 4 shows this variability by product and period; the last row of the table (“Robust Score”) summarizes the overall stability level of each decision category.

The results show that inventory is the most stable variable across scenarios and acts more as a buffer. In contrast, investment is the decision with the highest variability, as it is triggered only in situations of high uncertainty, such as the pandemic scenario. The backlog of work is low and only emerges when other decision mechanisms are insufficient, reflecting the strong impact of service level constraints. Outsourced (contracted) production, on the other hand, exhibits medium-to-high variability, particularly increasing in scenarios where no investment is made.

Row Labels	Norm Prod.	Norm Wage prod.	Norm Stock	Norm Backord.	Norm Invest. (0/1)	Norm Cap. Inc.
Antibiotics	0,02	0,63	0,19	0,26	0,00	0,00
August	0,04	0,81	0,00	0,28	0,00	0,00
September	0,09	0,96	0,01	0,40	0,00	0,00
July	0,00	0,66	0,10	0,00	0,00	0,00
June	0,07	0,83	0,12	0,00	0,00	0,00
May	0,00	1,00	1,00	0,00	0,00	0,00
October	0,02	0,55	0,00	1,00	0,00	0,00
Painkillers	0,22	0,09	0,01	0,00	0,78	0,57
August	0,47	0,00	0,06	0,00	1,00	0,73
September	0,18	0,00	0,00	0,00	1,00	0,73
July	0,48	0,00	0,00	0,00	1,00	0,73
June	0,30	0,05	0,00	0,00	1,00	0,73
May	0,01	0,52	0,00	0,00	1,00	0,73
October	0,23	0,00	0,00	0,00	1,00	0,73
Vitamin	0,57	0,26	0,07	0,01	0,78	0,78
August	1,00	0,12	0,38	0,00	1,00	1,00
September	1,00	0,08	0,01	0,00	1,00	1,00
July	0,91	0,14	0,00	0,00	1,00	1,00
June	0,39	0,56	0,00	0,02	1,00	1,00
May	0,01	0,74	0,04	0,05	1,00	1,00
October	0,94	0,15	0,00	0,00	1,00	1,00
Robust Score	0,66	0,60	0,91	0,90	0,33	0,42

Figure 4: Normalized variance map of the decision variables (Authors’ own calculation)

The contribution of decision variables to total cost and stability across scenarios have been evaluated together. Figure 5 illustrates the relationship between these two dimensions. This relationship facilitates understanding both the cost impact of decisions and their stability under conditions of uncertainty. For example, a decision with high cost but low variance (e.g., in-house production) is a fundamental element that the system must implement in every scenario. In contrast, decisions with low cost but high variance (e.g., investment or external production) are flexibility tools for the system and vary depending on conditions. This analysis provides strategic guidance to decision-makers by clarifying which decisions are “fixed,” which are “flexible,” and which are “conditional” for long-term planning.

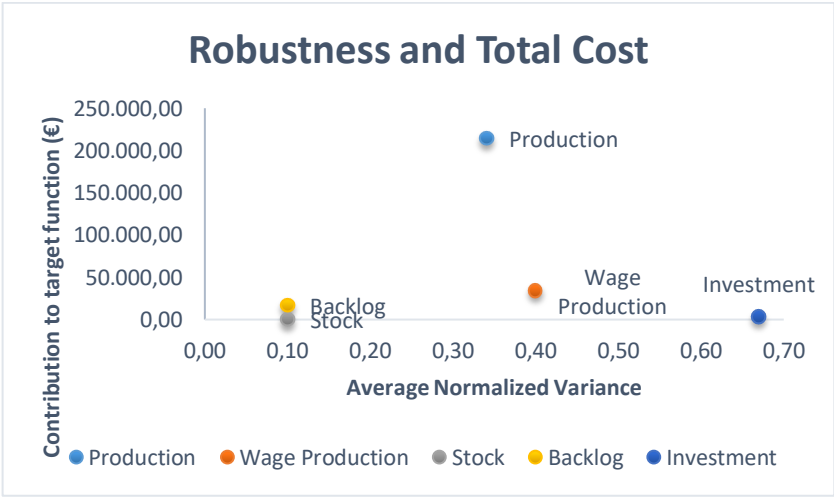


Figure 5: Variance and cost variance of total costs (Authors’ own Authors’ own elaboration)

4.5 Model validity: Sensitivity analysis

A two-way sensitivity analysis was conducted to examine how the model responds to changes in different parameters. This analysis focused specifically on two parameters: the delay penalty coefficient and fixed investment costs. The aim was to reveal the effects of changes in these parameters on production quantities, inventory levels, use of contracted production, capacity increases, delays, and total costs.

4.5.1 Penalty coefficients for arrears

When the delay penalty coefficient is reduced, the model's decision structure changes significantly. When the penalty coefficient is -100%, the system allows for a high amount of accumulated work (4,764 units) by keeping production and outsourcing to a minimum (total production 17,214; contracted production 942).

Although the service level drops to 95.48%, the total cost falls to €182,695, the lowest level throughout the analysis.

When the penalty factor is -50%, the model prefers to meet more demand; backlog falls to 1,302 units while contracted production rises to 1,392 units and inventory levels increase. As a result, the service level rises to 98.77% and total cost is €259,060.

When the penalty coefficient reaches values of -25% and above, the system's behavior stabilizes: the backlog stabilizes at 449 units, and the service level is maintained at approximately 99.57%. Increasing the penalty coefficient within the 0–100% range does not change the strategy; it only causes an increase in total costs. Sensitivity analysis of penalty coefficients for arrears is in Table 16.

Table 16: Sensitivity analysis of penalty coefficients for arrears (Authors' own calculation)

Change in penalty (%)	Total production	Total wage	Total inventory	Total backlog	Total capacity increase	Total investment	Total Cost	Service level
-100	17.214	943	196	4.764	1.170	0,3	182.696	95,48 %
-75	17.990	167	177	4.764	1.170	0,3	229.228	95,48 %
-50	20.227	1.392	216	1.302	1.170	0,3	259.060	98,77 %
-25	20.241	2.230	266	449	1.170	0,3	264.103	99,57 %
0	20.241	2.230	266	449	1.170	0,3	268.330	99,57 %
25	20.241	2.230	266	449	1.170	0,3	272.557	99,57 %
50	20.241	2.230	266	449	1.170	0,3	276.783	99,57 %
75	20.244	2.228	271	449	1.170	0,3	281.010	99,57 %
100	20.244	2.228	271	449	1.170	0,3	282.160	99,57 %

The fundamental reason for this trend is clear: as the penalty coefficient increases, the same amount of accumulated work leads to higher costs. The model prefers not to reduce the backlog below 449 units because this would require more production or higher inventory, and the cost of these options is higher than paying the penalty. Therefore, the system remains at the most economically rational equilibrium point.

Another important finding is that even when the penalty coefficient is zero, the service level is maintained at 99.57%. Even if the penalty cost is eliminated, the model continues to meet demand to a large extent. This shows that the model does not merely comply with constraints; it exhibits optimal behavior by considering the overall cost balance.

Overall, the sensitivity analysis reveals that even if the penalty coefficient increases, the model's decision equilibrium remains intact, the high service level is maintained, and unnecessary cost increases are avoided. These results confirm that the model is both theoretically consistent and practically reliable.

4.5.2 Sensitivity analysis of fixed investment costs (capacity increase)

In this sensitivity analysis, the model's capacity expansion decisions, production strategy, and total costs were evaluated by varying capital investment costs. The model was run for all four scenarios at each investment cost level, and the results were analyzed by weighting them with scenario probabilities. Thus, model performance was tested realistically under a broader uncertainty structure, not just based on individual scenarios. Sensitivity analysis of fixed investment costs is shown in Table 17.

Table 17: Sensitivity analysis of fixed investment costs							
Fixed investment costs	Total product ion	Total wage	Total inventory	Total backlog	Total capacity increase	Total investment	Total costs
0	20.847	1.624	75	449	6.510	1,7	263.294
2.500	20.567	1.905	266	449	2.010	0,5	265.604
5.000	20.241	2.230	266	449	1.170	0,3	266.830
10.000	20.241	2.230	266	449	1.170	0,3	268.330
15.000	20.244	2.228	271	449	1.170	0,3	269.830
20.000	20.241	2.230	266	449	1.170	0,3	271.330

50.000	20.241	2.230	266	449	1.170	0,3	280. 330
100.000	19.175	2.715	266	1.030	630	0,2	292. 826
1.000.000	19.175	2.715	266	1.030	630	0,2	427. 826

When the investment cost was €0, the system made a total investment of 1.7 units, directing most of this investment towards the pandemic scenario. This reduced the need for contracted production, maintained the service level, and resulted in a total cost of €263,294.

When the investment cost reached €2,500, the investment amount decreased to 0.5 units; the strategic approach remained unchanged, and the total cost increased only slightly to €265,604.

In the €5,000–50,000 range, the system invested 0.3 units in the pandemic scenario; investment was deemed unnecessary in other scenarios. The service level remained stable, and total costs rose in line with the increase in investment costs, ranging from €266,830 to €280,330.

When the investment cost reached €100,000, the investment made in the pandemic scenario fell to 1 unit, causing the capacity gap to increase; the contracted production volume rose, and the backlog reached 1,030 units. This brought the total cost to €292,826.

When the investment cost reached €1,000,000, only 1 unit was invested even under pandemic conditions; the capacity gap grew significantly, the service level declined, and the total cost reached €427,826, the highest value in the analysis.

These results show that the model shapes investment decisions not only based on average performance but also according to the level of pressure created by each scenario. Since the most severe conditions were observed in the pandemic scenario, investments were primarily directed toward this scenario. However, when investment costs exceeded critical thresholds, the system avoided investing; as a result, service levels declined, backlog increased, and total costs rose rapidly.

Overall, the model produces economically consistent strategies that rationally evaluate the cost-benefit balance under uncertainty, prioritize high-pressure scenarios, and align with real-world decision-making processes.

4.6 Discussion

This section demonstrates the alignment of the study with the literature and compares the findings with previous research. The developed scenario-based production and capacity planning model offers a flexible decision framework that

simultaneously evaluates both cost and service performance under uncertainty and exhibits an approach consistent with similar models in the literature.

As in the literature, this study clearly shows the decisive impact of investment decisions on system performance. Lindahl et al. (2024) emphasize that the type and timing of investment are directly related to strategic objectives. In this study, investment decisions were optimized in an integrated manner with contract manufacturing, and it was observed that the two mechanisms complemented each other, particularly in the pandemic scenario. Even with maximum investment capacity utilized, the service level declined to 96%; however, intensive contract manufacturing prevented an increase in backlog, thereby providing significant flexibility to the system.

Retrofit investments, as discussed by Lindahl et al. (2024), also follow a similar strategic logic. Although these investments are fast and relatively low-cost, capacity increases are limited; therefore, they are not as effective as contract manufacturing for large-scale capacity needs.

Vila-Parrish et al. (2012) define two basic strategies for meeting demand: (1) resolving all bottlenecks through outsourcing, (2) accelerating production based on internal inventory through outsourcing when necessary.

Their work shows that, despite the cost of outsourcing, it significantly increases the system's adaptability. In this study, contract manufacturing also played a role as a flexibility mechanism that supported investment or replaced it when investment was not possible. Similarly, it was observed that safety stocks for certain products balanced the need for external sourcing during periods of high demand. The variance analysis conducted in this study confirms that proactive stockpiling behavior emerged depending on the scenario, particularly for products such as vitamins.

Comparing this model with that of Tsang et al. (2007) further highlights the importance of investment flexibility. While Tsang's study focuses on clinical risks and NPV analyses, this study examines operational decision areas such as production quantities, service levels, and cost components. Therefore, depending on the level of pressure created by each scenario, investment emerges as an active decision tool in some cases and as an element replaced by mechanisms such as contract manufacturing or inventory in other cases. The common ground between the two approaches is that investment flexibility under uncertainty is a critical factor that enhances system performance.

Blossey et al. (2022) show that demand uncertainty and volatility in production approval dates directly affect capacity requirements and internal production–CMO balances. Although regulatory uncertainty was not directly modeled in this study, the capacity contraction that emerged in the raw material

supply problem scenario created a similar effect; the system was able to maintain service levels by shifting to contract manufacturing without investing.

The variance analyses also show parallels with Pathy and Rahimian (2023), revealing that inventory exhibited low volatility across scenarios and contributed to costs at a limited level. This result indicates that inventory functions as a complementary buffer mechanism rather than a strategic decision tool.

This finding is also consistent with Marques et al. (2017)'s observation that early decision-making is critical. In this model, all decisions are determined at the beginning of the planning period; investment, wage, inventory, and backlog strategies are optimized according to scenario probabilities. Thus, while resources are allocated efficiently at an early stage, unnecessary capacity losses and cost increases that may arise due to delayed responses are prevented. Both studies show that early, information-based, and proactive decisions significantly improve system performance.

Uthayakumar and Priyan (2013) developed a framework that integrates inventory, production, and distribution decisions in the producer–hospital supply chain, evaluating service level alongside cost minimization. However, while service level was treated as an indirect performance measure in that study, this study models service level directly with an 80% constraint. Furthermore, while Uthayakumar and Priyan's approach focuses more on logistics decision areas such as sales frequency, delivery volume, and distribution pattern, this study examines higher-level operational and strategic decisions such as production quantities, capacity expansion, and investment.

Rossetti et al. (2011) demonstrate that the structural transformation of the biopharmaceutical supply chain is shaped by pricing policies, intermediary roles, regulatory pressures, and external market conditions. Such external factors play a critical role in determining operational strategies such as service level, capacity utilization, and inventory. The scenario-based model developed in this study offers a similar perspective by systematically evaluating conditions such as pandemics, capacity bottlenecks, and supply uncertainties. Thus, the power relations and regulatory effects emphasized by Rossetti et al. (2011) are indirectly represented in this study through scenarios, demonstrating how they affect strategic outcomes such as production flexibility, capacity investment, and cost–service balance.

The scenario-based approach used in this study systematically addresses the visibility and flexibility requirements emphasized by Rossetti et al. (2011) and provides a robust decision framework for future supply chain designs. Thus, transformation trends, which have been discussed more at a theoretical level in the literature, can be tested in a practical and quantitative context through this model.

Tucker (2020) linked drug shortages to low inventory levels and fragile supply chains, showing that investing in flexible capacity structures is not always preferred due to cost pressures. A similar finding is observed in this thesis: high-cost investments are a rational option only in certain scenarios, while contract manufacturing emerges as a more suitable alternative under other conditions. The common point of both studies is that they point to the critical determinacy of downtime on system performance. However, while Tucker (2020) highlights policy and regulatory effects, this study focuses on multi-product production structures and capacity relationships at the firm level. This comparison clearly demonstrates the necessity of holistic strategies that address both economic and operational dimensions.

Kochakkashani et al. (2023) developed a MILP model aimed at optimizing service levels and costs for hospital supply chains under COVID-19 conditions and simplified the solution through product clusters. Similarly, in this study, production, inventory, backlog, and contract manufacturing decisions for different product groups are optimized within scenario frameworks. Both models aim to reduce costs while maintaining high service levels and to evaluate system behavior through scenario-based analysis. However, while Kochakkashani et al. (2023) presented scenarios based on regional characteristics, the scenarios used in this thesis were created based on industry-wide uncertainty themes; thus, they represented a broader uncertainty structure.

In summary, both studies show that while controlling costs in healthcare and pharmaceutical supply processes, scenario-based treatments in a high service range provide significant benefits to decision-makers. When evaluated alongside the micro-level production and capacity planning decisions in this thesis, the introduction by Kochakkashani et al. (2023) establishes a multifaceted foundation for integrated and flexible planning within the framework conditions.

Vann Yaroson et al. (2024) examined the role of resilience strategies in reducing drug shortages in the UK pharmaceutical supply chain using a mixed-methods approach. Qualitative and quantitative findings show that strengthening communication mechanisms, increasing the use of alternative sources, implementing flexible planning practices, and transitioning to collaborative supply chain structures significantly improve both short-term response capacity and long-term system resilience.

The scenario-based production planning model developed in this thesis also numerically represents similar resilience strategies. Through variables such as capacity expansion, contract manufacturing, initial inventory, and investment decisions, the model optimizes system performance under uncertainty and makes the impact of resilience-enhancing operational strategies measurable. While Vann Yaroson et al. (2024) analyze behavioral tendencies more at the application level, this study evaluates the quantitative effects of the same strategies directly through optimization results.

Therefore, although the two studies use different methodological approaches, they point to the same conclusion: Flexible and structural strategies that strengthen resilience are critical in terms of both maintaining service levels and controlling costs. This thesis presents an analytical decision support perspective that complements behavioral findings by revealing the numerical counterparts of these strategies.

Vann Yaroson et al. (2024) address supply chain resilience from social and organizational perspectives, while this study examines the same topic from a numerical optimization perspective. Bilal et al. (2024) found that proactive strategies such as regular inventory controls, strengthening supplier relationships, and process simplification are the most effective methods for reducing uncertainty; in contrast, reactive practices such as stockpiling and emergency procurement offer short-term solutions and may cause instability in the long term. These findings largely align with the results of the scenario-based model: The model shows that proactive strategies such as capacity expansion and contract manufacturing maintain service levels under demand and capacity uncertainty and reduce the risk of stockouts and backorders.

While Bilal et al. (2024) highlight structural problems such as data quality issues, bureaucratic barriers, and supplier reliability based on field data, this study presents quantitative counterparts of ideal strategic behaviors through numerical optimization. This indicates that model outputs need to be supported by administrative and organizational improvements to be effective in practice.

The supply chain competencies defined by Wang and Jie (2020)—transparency, agility, and flexibility—are consistent with the behavior of this model. The model demonstrates transparency by clearly evaluating decisions according to scenarios, agility by quickly adapting to contract manufacturing in situations where no investment is made, and flexibility by varying different strategies according to products and scenarios.

Overall, this study quantitatively validates findings in the literature and offers a multidimensional decision support system applicable in highly uncertain environments. The model serves as a powerful tool not only for cost minimization but also for simultaneously optimizing strategic flexibility, service level security, and resource efficiency.

5. Conclusion

This study develops a scenario-based MILP model for production and capacity planning under uncertainty. The model is based on the pharmaceutical industry, where demand and capacity fluctuations play a critical role, and covers a six-month planning horizon for three products. Decision variables such as production, outsourcing, inventory, backorders, capacity expansion, and investment are jointly optimized to enable a holistic assessment of the system.

Using five different scenarios and their assigned probability of occurrence, the model aims to minimize total cost while maintaining an 80% service level.

The findings show that the system develops more dynamic and scenario-specific strategies, especially under conditions of high uncertainty, such as in a pandemic scenario. In some cases, despite high investment costs, the preference for investment decisions demonstrates that the model exhibits rational behavior in terms of long-term cost reduction and more effective demand fulfillment.

Sensitivity analyses statistically validated the robustness of the model; it was clearly observed which decisions the system showed flexibility in and which decisions it maintained consistently when parameters such as investment costs or delay penalty coefficients were changed. In particular, it was found that when investment costs exceeded a certain threshold, there was a decline in service level and total costs increased rapidly.

Overall, the model generates decisions based not only on average performance but also on the unique pressure and uncertainty levels of each scenario, effectively prioritizing resources accordingly. In this regard, the study provides a robust decision support structure that offers strategic flexibility under uncertainty and is adaptable to realistic conditions.

Acknowledgement

The authors declare that they have no conflict of interest.

References

- Bilal, A. I., Bititci, U. S., & Fenta, T. G. (2024). Effective supply chain strategies in addressing demand and supply uncertainty: A case study of Ethiopian pharmaceutical supply services. *Pharmacy*, 12(5), 132.
- Blossey, G., Hahn, G. J., & Koberstein, A. (2022). Planning pharmaceutical manufacturing networks in the light of uncertain production approval times. *International Journal of Production Economics*, 244, 108343.
- Clautiaux, F., & Ljubić, I. (2025). Last fifty years of integer linear programming: A focus on recent practical advances. *European Journal of Operational Research*, 324(3), 707-731.
- Dang, H. T., & Elgersma, B. A. (2020). Capacity and Inventory Optimization for Pharmaceutical Industry. Signature. https://ctl.mit.edu/sites/ctl.mit.edu/files/theses/scm2020-elgersma-capacity-and-optimization-for-pharmaceutical-industry-capstone_0.pdf
- Govindan, K., Soleimani, H., & Kannan, D. (2015). Reverse logistics and closed-loop supply chain: A comprehensive review to explore the future. *European journal of operational research*, 240(3), 603-626.
- Gökalp, E. (2017) *Modelling and solving healthcare decision making problems under uncertainty*. phd. University of Warwick. Available at: <http://webcat.warwick.ac.uk/record=b3170066~S1> (Accessed: 3 June 2025).
- Klibi, W., Martel, A., & Guitouni, A. (2010). The design of robust value-creating supply chain networks: a critical review. *European journal of operational research*, 203(2), 283-293.
- Kochakkashani, F., Kayvanfar, V., & Haji, A. (2023). Supply chain planning of vaccine and pharmaceutical clusters under uncertainty: The case of COVID-19. *Socio-economic planning sciences*, 87, 101602.
- Lindahl, S. B., Babi, D. K., Gernaey, K. V., & Sin, G. (2023). Integrated capacity and production planning in the pharmaceutical supply chain: Framework and models. *Computers & Chemical Engineering*, 171, 108163.
- Lindahl, S. B., Babi, D. K., & Sin, G. (2024). Pharmaceutical capacity expansion under uncertainty: Framework and models. *Computers & Chemical Engineering*, 189, 108808.
- Marques, C. M., Moniz, S., de Sousa, J. P., & Barbosa-Póvoa, A. P. (2017). A simulation-optimization approach to integrate process design and planning decisions under technical and market uncertainties: A case from the chemical-pharmaceutical industry. *Computers & Chemical Engineering*, 106, 796-813.

- Nasrabadi, A. M., Najafi, M., & Zolfagharinia, H. (2020). Considering short-term and long-term uncertainties in location and capacity planning of public healthcare facilities. *European Journal of Operational Research*, 281(1), 152-173.
- Nguyen, T. B. T., Sivakumar, A. I., & Graves, S. C. (2018). Capacity planning with demand uncertainty for outpatient clinics. *European Journal of Operational Research*, 267(1), 338-348.
- Pathy, S. R., & Rahimian, H. (2023). A resilient inventory management of pharmaceutical supply chains under demand disruption. *Computers & Industrial Engineering*, 180, 109243.
- Peltokorpi, A. (2010) *Improving efficiency in surgical services a production planning and control approach*. Aalto-yliopiston teknillinen korkeakoulu. <https://aaltodoc.aalto.fi/handle/123456789/4838> (Accessed: 3 June 2025).
- Plantin, A., & Johansson, M. I. (2012). Implementing production planning processes in health care—a case study of a surgery clinic. In *Proceedings of PLAN Research and Application Conference, Lund, August 2012*. <https://research.chalmers.se/en/publication/160230> (Accessed: 3 June 2025).
- Queiroz, M. M., Ivanov, D., Dolgui, A., & Fosso Wamba, S. (2022). Impacts of epidemic outbreaks on supply chains: mapping a research agenda amid the COVID-19 pandemic through a structured literature review. *Annals of operations research*, 319(1), 1159-1196.
- Rossetti, C. L., Handfield, R., & Dooley, K. J. (2011). Forces, trends, and decisions in pharmaceutical supply chain management. *International Journal of Physical Distribution & Logistics Management*, 41(6), 601-622.
- Snyder, L. V., & Shen, Z-J. M. (2019). *Fundamentals of supply chain theory*. John Wiley & Sons.
- Tsang, K. H., Samsatli, N. J., & Shah, N. (2007). Capacity investment planning for multiple vaccines under uncertainty: 1: capacity planning. *Food and bioproducts processing*, 85(2), 120-128.
- Tucker, E. (2020) *Modeling Pharmaceutical Supply Chains to Mitigate Drug Shortages*. Thesis. <http://deepblue.lib.umich.edu/handle/2027.42/155172> (Accessed: 3 June 2025).
- Uthayakumar, R., & Priyan, S. (2013). Pharmaceutical supply chain and inventory management strategies: Optimization for a pharmaceutical company and a hospital. *Operations Research for Health Care*, 2(3), 52-64.
- Vann Yaroson, E., Breen, L., Hou, J., & Sowter, J. (2024). Examining the impact of resilience strategies in mitigating medicine shortages in the United Kingdom's (UK) pharmaceutical supply chain (PSC). *Benchmarking: An International Journal*, 31(3), 683-706.

- Vila-Parrish, A. R., Ivy, J. S., King, R. E., & Abel, S. R. (2012). Patient-based pharmaceutical inventory management: a two-stage inventory and production model for perishable products with Markovian demand. *Health Systems*, 1(1), 69-83.
- Wang, M., & Jie, F. (2020). Managing supply chain uncertainty and risk in the pharmaceutical industry. *Health services management research*, 33(3), 156-164.

CHAPTER 10

Medical Device Security and Anomaly Detection with AI: Risks, Models, and Future Perspectives in the IOMT

Resul Kara¹

¹ Prof. Dr., Düzce University, Faculty of Engineering, Department of Computer Eng.
ORCID: 0000-0001-8902-6837

1. Introduction: The Internet of Medical Things (IoMT) Paradigm and the Security Imperative

The healthcare sector is currently undergoing a significant transformation driven by the Internet of Medical Things (IoMT). IoMT is defined as a specialized subset of the broader Internet of Things (IoT) dedicated specifically to healthcare applications (Adil et al., 2025). This ecosystem encompasses internet-connected medical devices, sensors, wearable technologies, and health informatics systems. Its primary objective is to fundamentally improve the quality, accessibility, and efficiency of healthcare services through remote patient monitoring (Adil et al., 2025), real-time and accurate diagnostics, and Artificial Intelligence (AI)-driven decision-making processes (Deb et al., 2025). The utilization of intelligent, interconnected devices holds the potential to optimize clinical workflows and significantly reduce long-term healthcare costs.

However, the immense opportunities presented by IoMT technology must be balanced against unprecedented security risks and privacy concerns. The highly sensitive nature of the data collected, processed, and transmitted by these devices specifically Electronic Protected Health Information (ePHI) (Seh et al., 2020) renders these systems highly attractive targets for cyber adversaries (BSI Group, 2018). Unlike traditional Information Technology (IT) systems, a security breach in the IoMT ecosystem results in consequences far beyond data loss or financial damage. Many medical devices, such as pacemakers, insulin pumps, or hospital ventilators, possess the capacity to perform direct physical actions on patients (Deb et al., 2025). This reality shifts cybersecurity from the realm of data protection directly into the domain of patient safety (Deb et al., 2025; Seh et al., 2020).

The fundamental dilemma facing the IoMT ecosystem is that its greatest *value* proposition interconnectivity and real-time data flow simultaneously constitutes its greatest *vulnerability*. In a traditional enterprise network, a Distributed Denial of Service (DDoS) attack signifies service disruption and financial loss. In IoMT, the same attack could prevent critical vital sign data from reaching a clinician or halt the operation of a life-sustaining device, leading directly to a *clinical catastrophe* or patient mortality (BSI Group, 2018). Consequently, IoMT security is not merely an IT issue but a critical *public health* concern.

This chapter argues that the risks inherent to the nature of the IoMT ecosystem cannot be mitigated by traditional defense mechanisms alone. A significant portion of vulnerabilities in healthcare systems stems from reliance on outdated technology that lacks appropriate encryption and authentication mechanisms (Adil et al., 2025). Traditional, perimeter-based security models are increasingly insufficient given the distributed, heterogeneous, and dynamic nature of IoMT (Olatunji et al., 2025).

A primary source of vulnerability is "constrained design" rather than simply "malicious software." Devices at the perception layer (sensors, implants) are often so constrained in terms of memory, energy, and processing power (Adil et al., 2025; Deb et al., 2025) that they cannot support standard cryptographic methods or heavy security software (Al-Amri et al., 2025). This renders "lightweight" and *intelligent* (AI-driven) security solutions not merely a preference, but an absolute *necessity*. This chapter will analyze the layered risks within the IoMT ecosystem, examine AI and Machine Learning (ML)-driven proactive Anomaly-based Intrusion Detection Systems (A-IDS) that are replacing signature-based defenses, and discuss future perspectives in this domain, including Edge AI, Federated Learning, Explainable AI (XAI), and Zero Trust Architecture (Olatunji et al., 2025).

2. Anatomy of the IoMT Ecosystem and Security Fundamentals

To fully comprehend the security challenges of the IoMT ecosystem, it is essential to first understand the architectural structure comprising these systems and the fundamental security principles designed to protect them.

IoMT has a four-layer architectural model. The generally accepted IoMT architecture in the literature consists of four main layers defining the journey of data from the patient to the healthcare professional (Adil et al., 2025; Ghayvat et al., 2024).

Perception/Sensor Layer: This is the physical foundation of the architecture. It includes components that collect physiological data directly from the patient or the patient's environment. These include implantable devices (e.g., pacemakers, deep brain stimulators) (Deb et al., 2025), wearable devices (e.g., smartwatches, ECG monitors, pulse oximeters) (Ghayvat et al., 2024), smart pill sensors, or bedside monitors in hospital rooms. The primary function of this layer is to collect data in raw format.

Gateway Layer: This layer serves as a bridge between the low-power devices in the perception layer and the cloud layer. Devices in the perception layer typically use short-range, low-energy communication technologies such as Bluetooth Low Energy (BLE), Wi-Fi, or Zigbee (Ghayvat et al., 2024). The gateway layer (typically a smartphone, a home hub, or a specialized hospital device) aggregates raw data from these devices and then securely transmits it to the next layer via the internet, 4G/5G, or a Local Area Network (LAN).

Cloud Layer: This is the high-capacity server infrastructure designed to store, manage, and process the massive amounts of data incoming from gateways. Complex tasks requiring high computational power are executed in this layer. Specifically, Big Data analytics and the training/execution of AI and ML models

(e.g., disease prediction, anomaly detection) occur within the cloud layer (Adil et al., 2025).

Application Layer: This is the final layer of the architecture where end-users (doctors, nurses, patients, and healthcare administrators) interact with the system (Adil et al., 2025). Data processed and interpreted in the cloud layer is presented in this layer as clinical decision support systems, dashboards for healthcare professionals, or mobile health management tools and alerts for patients.

2.1.Decentralized Approaches: Edge and Fog Computing

The cloud-centric nature of the traditional four-layer model creates challenges regarding latency and data privacy. In response, Edge and Fog computing paradigms have been developed. These approaches move computation and data analysis from the centralized cloud closer to the data source (the perception layer), specifically to gateways or local servers (the edge) (Sethi & Sarangi, 2017).

Instead of sending data to the cloud, it is analyzed directly on the device or at the local gateway. The rise of Edge Computing is not only a technical performance optimization but also a fundamental *security architecture* choice. Processing data locally naturally mitigates the risks of network attacks (e.g., Man-in-the-Middle) during transport to the cloud and risks associated with centralized storage (e.g., large-scale data breaches) (BSI Group, 2018). Edge computing reduces the attack surface, lowers latency, and decreases network bandwidth load.

Security in the IoMT ecosystem requires a multidimensional approach that extends beyond traditional cybersecurity principles.

The Traditional CIA Triad: The foundation of medical device security rests on the classic information security model, the CIA Triad:

Confidentiality: Preventing unauthorized access to data and devices (Seh et al., 2020).

Integrity: Preventing unauthorized modification or tampering of data (Seh et al., 2020).

Availability: Ensuring data and devices (e.g., a ventilator) are available and operational for authorized users when needed (Seh et al., 2020).

Beyond CIA: However, some researchers argue that the CIA triad is *insufficient* to fully address the complexity of medical device security (Thigpen, 2020). According to this view, factors such as accountability, trustworthiness, and non-repudiation must also be included in fundamental security analysis.

Patient Safety-Centric Approach: A critical critique in the literature is that cybersecurity focus often centers on protecting patient *records* (ePHI) (BSI Group, 2018; Thigpen, 2020), whereas the primary and more severe threat is to patient *health* and *life* (BSI Group, 2018). Poor or inadequate cybersecurity practices can directly and irreversibly impact patient health.

The evolution of security principles in IoMT demonstrates a paradigm shift from "Data Security" to "Patient Safety." In traditional IT, "Availability" means a website is functional; in IoMT, it means the *life-sustaining* function of a ventilator or pacemaker is operational. Similarly, "Integrity" is clinically far more dangerous if an attacker *alters* patient data (Thigpen, 2020); this could lead to an incorrect diagnosis or treatment plan. In this context, altered (compromised integrity) data is clinically more hazardous than stolen (compromised confidentiality) data.

Modern Frameworks (TIPPSS): Reflecting this patient-centric approach, next-generation standards like IEEE/UL 2933™-2024 adopt the TIPPSS framework. This framework defines security through six components: Trust, Identity, Privacy, Protection, Safety, and Security (IEEE, 2024). This approach formally recognizes "Safety" as an integral and central part of the cybersecurity equation.

3. Risks in the IoMT Ecosystem: A Layered Vulnerability Analysis

Each layer of the IoMT architecture harbors unique security vulnerabilities and attack vectors (Adil et al., 2025). A vulnerability in one layer can compromise the entire system and, consequently, the patient.

3.1. Perception Layer Risks (Device Level)

Due to the physical nature of devices and severe resource constraints, this layer is one of the most vulnerable links in the ecosystem.

Weak Security Configurations and Legacy Firmware: Many IoMT devices ship with default and weak security settings (e.g., easily guessable passwords) . More importantly, these devices often operate on outdated firmware containing unpatched security vulnerabilities.

Resource Constraints: Limits on memory, energy, and processing power make the implementation of traditional cryptographic security methods (e.g., heavy encryption protocols) or antivirus software technically impossible (Al-Amri et al., 2025).

Physical Attacks and Tampering: Devices located in a patient's home or public areas are particularly susceptible to physical access. Attackers can perform

hardware tampering attacks, implant fake devices, or gain unauthorized access via physical interfaces (Deb et al., 2025).

3.2. Gateway and Network Layer Risks (Communication)

This layer, where data is transmitted from devices to the cloud or gateways, is vulnerable to exploits in communication protocols and network configurations.

Distributed Denial of Service (DDoS): Attackers can target IoMT networks or healthcare servers via botnets (often comprised of other compromised IoT devices). These attacks can lock down networks with excessive traffic, causing service interruptions, critical delays in accessing health resources, and even total system collapse (American Hospital Association, 2018). The significant increase in DDoS attacks against healthcare systems during the COVID-19 pandemic demonstrated the reality and danger of this threat.

Man-in-the-Middle (MitM): Unencrypted or weakly encrypted (e.g., legacy protocols) communication channels allow attackers to intercept the connection between the device and the server to read, steal, or more dangerously, *alter* data during transmission (American Hospital Association, 2018).

Jamming and Wireless Attacks: Bluetooth or Wi-Fi connections extensively used by IoMT devices are vulnerable to jamming attacks. An attacker can physically block data transmission by flooding the wireless communication channel (Al-Amri et al., 2025). This leads to a loss of network availability and the failure to transmit critical data (e.g., an arrhythmia alert).

3.3. Cloud and Application Layer Risks (Data and Access)

Where data is collected, stored, and accessed by end-users, these layers are primary targets for data breaches and high-level logical attacks.

Ransomware: Specifically targeting hospitals, "Crypto Ransomware" attacks encrypt and lock patient records (ePHI) and hospital operations (e.g., appointment systems, lab results), halting services (BSI Group, 2018). The 2017 "WannaCry" attack, which hit the UK's National Health Service (NHS), illustrated the scale of this threat. New variants like "Ransomcloud" target cloud services and can rapidly encrypt massive amounts of cloud-stored data.

Data Breaches and ePHI Theft: Stolen health records can be valued at 10 times or more than credit card information on the black market (American Hospital Association, 2018; Thigpen, 2020). This high financial motivation makes health data a primary target for cybercriminals.

Software Vulnerabilities and Unauthorized Access: Weak authentication mechanisms in the application layer, social engineering attacks, or software-

based attacks like SQL injection and cloud malware injection allow unauthorized individuals to infiltrate systems and access data (Adil et al., 2025).

3.4.Direct Impacts on Patient Safety

The consequence of these attacks is not merely a data privacy breach or financial loss; the threat directly targets patient life (BSI Group, 2018).

Service Disruption and Denial of Care: A DDoS or ransomware attack can render care impossible by preventing doctors from accessing life-saving medical devices (e.g., ICU ventilators) or necessary patient records (BSI Group, 2018; Thigpen, 2020).

Compromised Data Integrity (Data Alteration): Perhaps the most insidious threat is the *alteration* of patient data (intentional or unintentional) (Thigpen, 2020). Changing a patient's blood type, allergy information, or ECG data can lead to completely incorrect diagnoses, wrong medication administration, or erroneous surgical interventions.

Device Manipulation: The most dangerous and direct scenario is the hijacking of device functionality (Adil et al., 2025). This can cause life-threatening device malfunctions or direct harm, such as an insulin pump delivering a lethal dose (Deb et al., 2025).

The IoMT ecosystem functions as a "cyber-physical kill chain," linking a vulnerability in the cyber world (e.g., a weak password at the Application layer) to a *lethal* outcome in the physical world (e.g., a device executing a wrong command at the Perception layer) (Deb et al., 2025). This is the most tangible example of Cyber-Physical System (CPS) risk in medicine. The threat is not data loss, but *physical harm*.

4. Beyond Traditional Defense: Why AI-Based Anomaly Detection?

The layered risk profile of the IoMT ecosystem reveals why traditional cybersecurity defenses fall short and why new, intelligent approaches are urgently needed.

4.1.Limitations of Traditional Intrusion Detection Systems (IDS)

Traditional network security relies heavily on signature-based Intrusion Detection Systems (IDS). These systems depend on massive databases of "known" attack patterns (signatures), working much like a virus fingerprint. They compare network traffic against this database and trigger an alarm upon finding a match.

This approach has two fundamental inadequacies in the context of IoMT:

Blindness to Zero-Day Attacks: The greatest disadvantage of signature-based IDSs is that they are inherently *reactive*. They can only recognize attacks that have been previously seen, analyzed, and added to the database. Consequently, they are completely blind to *new* or *zero-day* attacks for which no signature exists (Al-Amri et al., 2025).

Static Nature: These systems require constant signature updates and are easily bypassed by attackers using simple evasion techniques (e.g., slightly modifying attack code to invalidate the signature).

4.2.The Need for Anomaly-Based IDS (A-IDS)

In contrast to the blindness of signature-based systems, Anomaly-Based Intrusion Detection Systems (A-IDS) operate on a fundamentally different philosophy. Instead of keeping a list of "bad" things, an A-IDS learns the "normal" behavior profile of the system or network. It operates by detecting *deviations* (anomalies) from this normal behavior.

The critical advantage for IoMT is that this approach does not require an attack signature or prior knowledge of the attack. The system can understand that something is "not normal" without knowing exactly "what" the attack is. Thus, A-IDSs have the capability to detect both *known* and *unknown* (zero-day) attack patterns (Adil et al., 2025; Al-Amri et al., 2025).

4.3.IoMT-Specific Challenges and the Necessity of AI

The need for A-IDS, combined with two fundamental challenges specific to IoMT, places Artificial Intelligence (AI) at the center of the equation:

Resource Constraints: As noted in Section 3, traditional security mechanisms are typically complex and resource-intensive (Al-Amri et al., 2025). IoMT devices possess severe limitations in processing power, memory, and battery life (Adil et al., 2025; BSI Group, 2018). These devices simply *cannot run* standard security solutions. Furthermore, a device's data transmission frequency may constantly change based on a patient's sleep, activity, or illness status. Defining this dynamic "normal" profile manually or with simple rules is impossible.

The transition to AI-based A-IDS stems from a *dual imperative* in IoMT. First (Threat Aspect): Cyberattacks (especially zero-day) are evolving too fast and becoming too complex for traditional signature-based systems. Second (Device Aspect): IoMT devices are too weak and resource-constrained to run traditional (heavy) security software (Al-Amri et al., 2025; Seh et al., 2020).

AI and Machine Learning are the only existing paradigms capable of solving both constraints simultaneously. AI forms the engine of these A-IDSs. AI/ML models can learn normal patterns in complex IoMT traffic or sensor data and

detect zero-day attacks (Adil et al., 2025) or device malfunctions in real-time. Therefore, there is a critical need for *lightweight* A-IDS solutions specifically designed for IoMT networks (Al-Amri et al., 2025).

Finally, the concept of "anomaly" in IoMT has a dual nature. AI-based systems can be used to detect not only 1) Security Anomalies (e.g., abnormal network traffic caused by an attacker, DDoS) but also 2) Clinical Anomalies (e.g., abnormal physiological data indicating health deterioration, arrhythmia) (Binbusayyis et al., 2022). Ideally, a robust future A-IDS should be able to distinguish whether a device is under cyberattack or transmitting clinically concerning data.

5. AI and Machine Learning Models for IoMT Security

The AI/ML models powering the A-IDSs necessitated in Section 4 can be classified into three main paradigms: Supervised, Unsupervised, and Deep Learning. Each approach has distinct advantages and disadvantages regarding different aspects of IoMT security.

5.1. Supervised Learning Models

Supervised learning is an approach where a model is trained using data previously labeled as "normal" and "attack" (Binbusayyis et al., 2022). Based on these labeled examples, the model learns to predict which class (normal or attack) new, unlabeled data belongs to.

Models: Popular algorithms in this category include Support Vector Machines (SVM) (Binbusayyis et al., 2022), K-Nearest Neighbors (k-NN) (Binbusayyis et al., 2022), and ensemble learning methods like Random Forest (RF) (Binbusayyis et al., 2022).

Evaluation (IoMT): The greatest advantage of supervised models is their extremely high accuracy rates in detecting *known* attack types. Studies report that these models reach impressive accuracy levels, such as 99.86% (Binbusayyis et al., 2022), particularly on labeled datasets like Bot-IoT. This makes them excellent for *compliance* and defense against known threats. However, their major disadvantage is their complete blindness to *new* (unlabeled) or *zero-day* attacks, as discussed in Section 4.

5.2. Unsupervised Learning Models

Unsupervised learning is generally seen as a more suitable approach for IoMT security because it works on *unlabeled* data. The goal of these models is to learn the structure or distribution of "normal" data and flag *outliers* or *novelties* (Al-Amri et al., 2025) that do not fit this structure as anomalies.

The models can be listed as follows: Models:

Isolation Forest (IF): A model that quickly detects outliers (anomalies) based on the ease of "isolating" data points by randomly splitting them. It relies on the assumption that anomalies can be isolated with fewer splits than normal points. Its lightweight nature makes it suitable for real-time systems and resource-constrained environments (Al-Amri et al., 2025).

One-Class SVM (OCSVM): This model is trained only on data considered "normal" and learns a boundary (hyperplane or hypersphere) enveloping this normal dataset. Any new data point falling outside this boundary is classified as an anomaly. However, it can be inefficient and slow on large, non-linear datasets.

Autoencoders (AE): While often classified under deep learning, they are fundamentally used for unsupervised reconstruction tasks. An Autoencoder is a neural network that learns to compress data into a lower-dimensional representation (encoding) and then reconstruct the original data from this representation (decoding). When trained *only on normal data*, the model can reconstruct normal data with low error. However, when it encounters abnormal data it has never seen (e.g., an attack packet), it fails to reconstruct it properly, producing a high *reconstruction error* (Aldhaheri et al., 2025). This high error threshold allows for anomaly detection.

The primary advantage of unsupervised models is their potential to detect zero-day attacks (Al-Amri et al., 2025). This makes them critical for *resilience*. However, their most distinct disadvantage is their potential oversensitivity to small, harmless changes in "normal" behavior (e.g., a device receiving a software update), leading to *high false positive rates* (Adil et al., 2025; Aldhaheri et al., 2025).

5.3. Deep Learning Models

Deep learning utilizes multi-layer neural networks to automatically learn complex, hierarchical, and especially *temporal* patterns inherent in IoMT data (network traffic or physiological signals).

The models can be listed as follows: Models:

Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM): LSTM is a specialized type of RNN that is far more successful than standard RNNs in learning long-term dependencies in *time-series* or *sequential* data. They are ideal for data such as sequences of packets in a network traffic flow or a patient's continuous ECG signal. They can capture temporal patterns of an attack (e.g., a slow scan).

Convolutional Neural Networks (CNN): While known for image processing, they can be used to capture *spatial* patterns in IoMT data (e.g., relationships

between different sensor data). They are often combined with LSTM in hybrid models (e.g., CNN-LSTM) for temporal analysis.

They generally demonstrate superior performance over supervised and unsupervised models in analyzing complex attack scenarios and high-dimensional physiological data (Binbusayyis et al., 2022). However, they have two significant drawbacks: 1) They typically have high computational costs, making them unsuitable for resource-constrained devices (without Edge AI optimizations). 2) Their "black-box" nature, making it difficult to understand how decisions are made, reduces interpretability which is critical for clinical and security contexts (Aldhaheri et al., 2025).

In summary, there is no single "best" AI model; there is the *right* AI model for the task. A hybrid strategy is necessary for IoMT security. For Network-level detection, models analyzing *time-series* data like network traffic such as LSTM or hybrid CNN-LSTM are most suitable for capturing temporal dependencies. Conversely, for Sensor-level detection (monitoring device metrics or physiological data), Unsupervised Models that learn "normal" behavior (e.g., lightweight Isolation Forest or Autoencoders) are generally more appropriate (Aldhaheri et al., 2025).

6. Case Studies: Performance of AI Models on Real-World Datasets

The practical efficacy of the AI models classified theoretically in Section 5 can only be understood when tested on realistic and modern datasets.

We recommend the use of modern datasets instead of traditional datasets. Older cybersecurity datasets (e.g., KDD Cup 99) are insufficient to reflect the complex, high-volume, and multi-protocol IoT/IoMT network traffic and attack vectors of today. Therefore, researchers have focused on developing realistic datasets specific to IoMT.

Bot-IoT Dataset: A widely used, labeled modern dataset containing botnet attacks (DDoS, scanning, etc.) in IoT networks (Binbusayyis et al., 2022).

CICIoMT2024 Dataset: One of the newest generation datasets designed specifically for IoMT security. It reflects the heterogeneous nature of IoMT, containing *multi-protocol* (MQTT, WiFi, Bluetooth) and *device-specific* (e.g., wearable sensors, heart monitors) data. It also hosts realistic attack scenarios targeted at IoMT, such as DoS, DDoS, and spoofing (Aldhaheri et al., 2025).

6.1. Performance Analysis (Bot-IoT Dataset)

Studies on the Bot-IoT dataset report that AI-based IDSs demonstrate exceptionally high accuracy rates in detecting known attack types.

Neural network-based models achieved nearly perfect accuracy of 99.9% on Bot-IoT. A comparative analysis found that deep learning models (RNN and LSTM) outperformed traditional supervised models (SVM), providing superiority in detecting known IoT attacks (Binbusayyis et al., 2022).

This confirms AI's capability to classify *known* threats with high accuracy.

6.2. Performance Analysis (CICIoMT2024 Dataset)

The CICIoMT2024 dataset offers a more complex and challenging benchmark due to its inclusion of multiple protocols and zero-day scenarios. Recent case studies on this dataset reveal the limitations of singular models and the superiority of *hybrid* or *ensemble* models (Aldhaheri et al., 2025).

A notable case study tested a *stacking ensemble model* combining different AI paradigms described in Section 5:

Base Learners: Autoencoder (AE) (unsupervised, zero-day detection), Isolation Forest (IF) (unsupervised, lightweight anomaly detection), LSTM, and CNN-LSTM (deep learning, temporal patterns) were used to capture different data features.

Meta-Learner: An XGBoost model (advanced supervised ensemble model) was used to take the predictions (outputs) of the base learners as input and make the final decision.

The performance of this hybrid model on the CICIoMT2024 dataset yielded the following results compared to singular models (Aldhaheri et al., 2025):

Ensemble (XGBoost): Showed the highest and most balanced performance. It achieved accuracy between 95% and 100% on attack-specific data. However, the most critical finding is that this model *significantly reduced false positive rates* (Aldhaheri et al., 2025).

Autoencoder (AE): While performing well in general anomaly detection (on unlabeled data), it led to *significantly more false positives* in test data.

LSTM: Was effective in attacks requiring temporal patterns like MQTT spoofing, but its general accuracy on time-series data (76%) lagged behind the ensemble model.

This case study proves that there is no "single best model" for IoMT security, but an *intelligent combination of models* (ensembling) yields the best results. The Stacking approach offers a robust practical strategy by taking the strengths of different models (e.g., AE's zero-day detection, LSTM's temporal analysis) and correcting their weaknesses (e.g., AE's high false positives) with a meta-learner (Aldhaheri et al., 2025).

High accuracy alone is not a sufficient metric. Especially in a life-critical field like IoMT, a *false positive* alarm (e.g., shutting down a device unnecessarily or constantly occupying a doctor with false alarms) leads to "alert fatigue" (Olatunji et al., 2025), potentially paralyzing clinical operations. The success of ensemble models lies in making the system *practically usable* by reducing false positives.

7. Future Perspectives: New Horizons in IoMT Security

While the AI models examined in Sections 5 and 6 provide significant advancements in IoMT security, fundamental challenges remain. These include high computational costs, patient data privacy, and the "black box" nature of AI decisions. This section addresses four key advanced paradigms developed to overcome these challenges: Edge AI, Federated Learning, XAI, and Zero Trust Architecture.

7.1. Challenge1: Resource Constraints with Solution: Edge AI and Lightweight Models

Advanced deep learning models (e.g., LSTM, CNN-LSTM) require high computational power for high accuracy. This makes them "heavy" and unsuitable for the restricted computation, memory, and energy resources of IoMT devices or local gateways (Adil et al., 2025; Al-Amri et al., 2025).

Edge AI: An approach where AI models (specifically the *inference* stage for anomaly detection) are run *at the edge*, near the data source (device or gateway), rather than in the centralized cloud.

Significance: This architecture provides three key advantages: 1) **Low Latency:** Since data does not need to travel to the cloud and back for analysis, an anomalous condition (attack or clinical event) can be detected instantly (Sethi & Sarangi, 2017). 2) **Privacy:** Sensitive patient data is analyzed without leaving the device or local network. 3) **Bandwidth:** Network bandwidth usage decreases as only processed, meaningful data (e.g., alerts) is sent to the cloud.

7.2. Challenge2: Data Privacy with Solution: Federated Learning (FL)

Training high-performance AI models requires massive amounts of diverse data. However, strict data privacy regulations like HIPAA or GDPR make aggregating sensitive patient data (ePHI) on a centralized server for analysis legally and ethically nearly impossible (Makhdoom et al., 2025).

Federated Learning (FL): A decentralized training mechanism that takes the AI model to the data, rather than bringing the data to the model (Gupta et al., 2023).

Significance: In FL architecture, raw patient data *never leaves* the local network of the device or hospital. Instead, a global AI model is sent to local

devices/hospitals, trained on local data, and only the model "updates" (e.g., gradients or weights) are encrypted and sent back to the central server. The server aggregates these updates to create a smarter global model.

Advanced FL Security: FL itself can be vulnerable to data leakage or model poisoning attacks. Advanced frameworks like SAFE-MED use Privacy-Enhancing Technologies (PETs) such as *neural cryptography* or *differential privacy* to secure these model updates (Makhdoom et al., 2025).

7.3. Challenge3: The "Black Box" Problem with Solution: Explainable AI (XAI)

Deep Learning models, despite high performance, operate as "black boxes" where it is unclear *how* and *why* they make decisions.

Significance: This is unacceptable in both *clinical* and *security* contexts.

Clinical Trust: When an AI model flags an "anomaly", a doctor must know "why" (e.g., "due to T-wave inversion and ST-segment elevation in patient data") to trust and act on the decision.

Security Trust: When an A-IDS alarm triggers, a security analyst needs to know if it is a real threat or a *false positive*. Understanding which features (e.g., "unusual packet sizes from a specific port") the model focused on is critical to prevent "alert fatigue" (Olatunji et al., 2025).

Solution: XAI: A set of techniques explaining complex model decisions in terms understandable to humans. Techniques like LIME and SHAP are used to increase trust, transparency, and accountability in AI-based IDSs and clinical decision support systems.

7.4. Challenge4: Architectural Security with Solution: Zero Trust Architecture (ZTA)

Traditional perimeter-based security allows an attacker free lateral movement once they breach the network. In IoMT, the concept of "inside" vs. "outside" is virtually meaningless.

Solution: ZTA: A security model based on the principle "Never Trust, Always Verify" (Olatunji et al., 2025).

Significance: ZTA demands *continuous verification* and *granular access control* for *every* user, device, and application access request, regardless of network location. Combined with *micro-segmentation*, this prevents an attacker from spreading through the network even if a device is compromised (Olatunji et al., 2025).

7.5.Synthesis of Paradigms: The Future Architecture of Secure IoMT

The four "future" solutions presented (Edge, FL, XAI, ZTA) are not independent; they form a holistic architecture:

ZTA Needs AI: Continuous verification cannot be achieved with static rules. Trust scores must be determined dynamically via *AI-based anomaly detection*.

ZTA Needs Edge AI: This dynamic trust engine cannot run in the cloud due to latency requirements. It must be deployed as *Edge AI*.

Edge AI Needs FL: To train these Edge models without moving privacy-sensitive data, *Federated Learning* is required.

ZTA Needs XAI: When ZTA blocks access, **XAI** is needed to explain *why* to clinicians or analysts, ensuring accountability.

8. Conclusion

The transformation of the healthcare sector through the Internet of Medical Things (IoMT) represents a paradigm shift that promises to revolutionize patient care; however, this potential is shadowed by the emergence of deep cybersecurity risks that threaten to turn interconnectivity into a vulnerability. As explored throughout this chapter, the security imperative in this ecosystem has fundamentally transitioned from traditional data privacy to a critical focus on patient safety, recognizing that cyber vulnerabilities in medical devices can create a "cyber-physical kill chain" capable of causing direct physical harm. Consequently, traditional signature-based defense mechanisms, which are inherently reactive and blind to zero-day attacks, are rendered obsolete by the dynamic nature of IoMT and the severe resource constraints of medical sensors. This reality establishes that the integration of Artificial Intelligence (AI) and Machine Learning (ML) into Anomaly-based Intrusion Detection Systems (A-IDS) is not merely a technological enhancement but an operational necessity for the survival of the ecosystem.

Our analysis of AI methodologies reveals that while Supervised Learning models excel in compliance against known threats and Unsupervised Learning provides resilience against unknown anomalies, no singular model suffices on its own. Recent case studies on modern datasets, such as CICIoMT2024, demonstrate that hybrid and ensemble architectures which combine the temporal analysis capabilities of Deep Learning with the lightweight efficiency of unsupervised methods offer the most robust defense. These intelligent combinations are critical not only for achieving high detection accuracy but also for significantly reducing false positive rates, thereby mitigating the risk of "alert fatigue" that can paralyze clinical operations.

Looking toward the future, the security of the IoMT ecosystem cannot rely solely on isolated AI models but must instead embrace a holistic "Cyber-Physical Safety" architecture that converges multiple advanced paradigms. This next-generation framework necessitates the deployment of Edge AI to process data at the source, effectively resolving latency and bandwidth constraints while keeping sensitive data local. Simultaneously, to address the conflict between data-hungry AI training and strict privacy regulations, Federated Learning must be adopted to enable decentralized model updates without exposing raw patient records. Furthermore, to bridge the trust gap inherent in "black box" algorithms, Explainable AI (XAI) must be integrated to provide clinicians and security analysts with transparent, interpretable decision-making processes. Finally, these intelligent systems must operate within a Zero Trust Architecture (ZTA), which enforces a policy of continuous verification and granular access control, ensuring that trust is never assumed but always validated. Ultimately, the synthesis of these technologies defines the future of medical device security, moving beyond reactive defense to a proactive, resilient, and patient-centric safety model.

References

- Adil, M., et al. (2025). A systematic literature review on security vulnerabilities of IoMT devices. arXiv preprint arXiv:2501.07703.
- Aldhaheeri, S., et al. (2025). An advanced, real-time anomaly detection framework for IoMT network traffic. arXiv preprint arXiv:2502.11854.
- Al-Amri, R., et al. (2025). A-IDS for resource-constrained devices in IoMT Networks. *Sensors*, 25(4), 1216. <https://doi.org/10.3390/s25041216>
- American Hospital Association. (2018). The Importance of Cybersecurity in Protecting Patient Safety. AHA Center for Cybersecurity and Risk Advisory Services.
- Binbusayyis, A., Alaskar, H., Vaiyapuri, T., & Dinesh, M. (2022). An investigation and comparison of machine learning approaches for intrusion detection in IoMT network. *The Journal of Supercomputing*, 78, 1-20. <https://doi.org/10.1007/s11227-022-04568-3>
- BSI Group. (2018). Cybersecurity of medical devices. BSI Group White Paper.
- Deb, S., Lupu, E., Mic Drakakis, E., Bharath, A. A., Leung, Z. K., Ma, G. R., & Chattopadhyay, A. (2025). A taxonomy of potential attacks targeting IoMT. arXiv preprint arXiv:2507.19609.
- Ghayvat, H., et al. (2024). Systematic mapping of machine learning with edge computing and wearable devices. *BMC Medical Informatics and Decision Making*, 24(1), 1-20.
- Gupta, R., et al. (2023). Federated Learning for the Internet-of-Medical-Things: A Survey. *Mathematics*, 11(1), 151. <https://doi.org/10.3390/math11010151>
- IEEE. (2024). IEEE/UL 2933™-2024, Standard for Clinical Internet of Things (IoT) Data and Device Interoperability with TIPPSS – Trust, Identity, Privacy, Protection, Safety, Security. IEEE Standards Association.
- Makhdoom, I., et al. (2025). SAFE-MED for Privacy-Preserving Federated Learning in IoMT via Adversarial Neural Cryptography. *Mathematics*, 13(18), 2954. <https://doi.org/10.3390/math13182954>
- Olatunji, A., Alozie, E., Olagunju, H. I., & Udensi, F. O. (2025). Zero-Trust Architecture in IoMT: Applications, Issues, and Further Research Directions. ResearchGate.
- Seh, A. H., Zarour, M., Alenezi, M., Sarkar, A. K., Agrawal, A., Kumar, R., & Khan, R. A. (2020). Locations from where the protected health information is breached. *Journal of Medical Internet Research*, 22(7), e18418. <https://doi.org/10.2196/18418>
- Sethi, P., & Sarangi, S. (2017). Internet of things: architectures, protocols, and applications. *Journal of Electrical and Computer Engineering*, 2017.
- Thigpen, B. L. (2020). Strategies to Lower Security Risks Involving Medical Devices in Patient Care. (Doctoral dissertation, Walden University). ScholarWorks.

CHAPTER 11

AI-Assisted Semantic Evaluation of EMC and Electrical Safety Standards: A Review and Framework Proposal

Oğuz Gündüz¹ & Gökhan Demirkıran²

¹ Electrical and Electronics Engineer, Yaşar University, ORCID: 0009-0003-7770-6159

² Assoc. Prof. Dr., Yaşar University, Faculty of Engineering, Department of Electrical and Electronics Engineering, İzmir, Turkey, ORCID: 0000-0002-0076-6036

1. INTRODUCTION

Electromagnetic compatibility (EMC) is the ability of electrical and electronic equipment to function properly in their defined electromagnetic environment—such as home, industrial, or automotive—without emitting excessive noise that affects other equipment and without being unduly disturbed by external interference. In practice, this requirement is achieved through EMC and electrical-safety standards that specify limits, test set-ups, and measurement procedures. Manufacturers determine compliance by performing EMC tests and revising their designs accordingly.

In reality, compliance is rarely achieved in a single attempt. Instead, manufacturers follow an iterative design–test–redesign loop: engineers build a prototype, send it to an EMC lab, analyze any failures, modify the hardware, and schedule another EMC test session. In each loop, one should find relevant clauses, resolve conflicting standards, figure out the details of the test set up, document both the problems and the solutions. As products become more complex and regulations more fragmented, this process has become increasingly labor-intensive and cognitively demanding.

We argue that AI technologies today —especially large language models— can help with this process at three distinct levels:

- **Level 1 – Standards-aware assistant.** A retrieval-augmented LLM that lets engineers upload standards and guidance documents, ask questions in natural language, and obtain clause-linked answers, summaries, or rewritten documentation.
- **Level 2 – Test-analysis copilot.** A system that takes in measurement logs, immunity/emission tables, and test notes, then links them to relevant clauses, highlights likely root causes of failures, and suggests new designs or follow-up tests.
- **Level 3 – Non-language and multimodal tasks.** LLM-based agents that work with incomplete or noisy signals, simulation results, or control logs—tasks where statistical learning is difficult because data are scarce, proprietary, or fragmented, but where **semantic description** of the setup and expected behaviour is available.

Existing EMC literature has largely focused on standards, measurement techniques, and numerical simulation; AI-related work has so far targeted specific tasks such as EMI prediction or anomaly detection. There is very limited work that combines EMC with LLMs, and no end-to-end framework that treats standards interpretation, test reasoning, and interoperability as one coherent semantic workflow.

To address this gap, we introduce an AI-assisted semantic framework (Figure 1) for EMC and electrical safety, which treats standards, measurement data, and interoperability protocols as a unified knowledge space. Rather than proposing yet another isolated AI model, our goal is to:

1. **Define the main functional blocks** needed for LLM-based compliance support.
2. **Map existing literature** onto these blocks to show what is already available and what is missing.
3. **Articulate a research agenda** for turning the framework into a validated, standards-aligned tool.

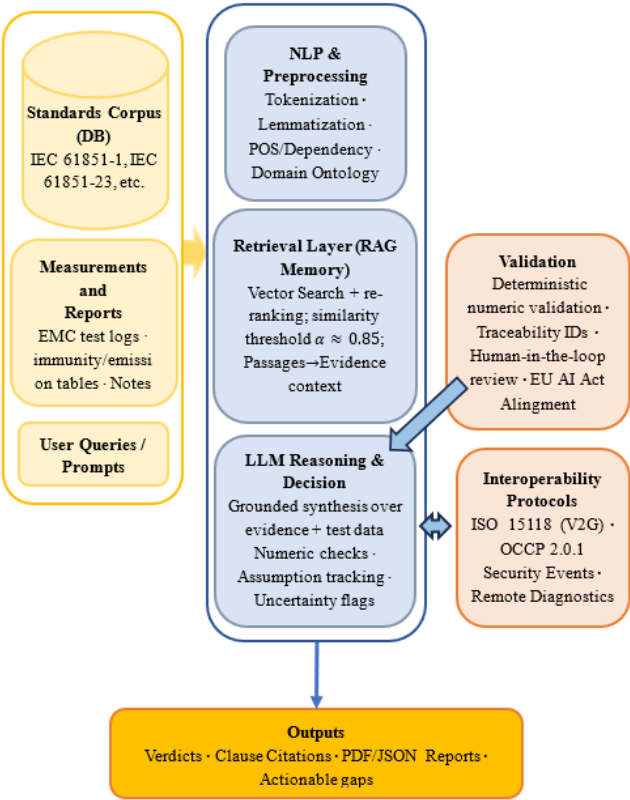


Figure 1. Proposed AI-assisted semantic evaluation framework for EMC and electrical safety standards

The architecture (Figure 1) combines (i) a standards and test-report corpus, (ii) NLP-based preprocessing and ontology mapping, (iii) a retrieval layer (RAG memory), (iv) LLM-based reasoning and numeric checking, and (v) validation and interoperability modules aligned with EU AI Act and domain protocols such

as ISO 15118 and OCPP. To position this framework, we provide a structured literature review along its components: (1) complexity and fragmentation of EMC and safety standards; (2) NLP and ontology work for regulatory text; (3) retrieval-augmented and knowledge-graph-based reasoning; (4) LLM use in other regulated domains (finance, law, privacy, safety); (5) validation, benchmarking, and human-in-the-loop oversight; and (6) AI for non-language technical tasks such as EMC simulation, anomaly detection, and multimodal interpretation.

Across these strands, we identify three levels of LLM support for EMC: Level 1, standards-aware assistants that answer questions and generate documentation; Level 2, test-analysis copilots that link failures to clauses and propose design modifications; and Level 3, models that interact with non-linguistic artefacts such as waveforms, logs, and control signals. The review shows that many ingredients of our framework already exist in the literature but remain unintegrated for EMC. We conclude by outlining research priorities and concrete evaluation milestones needed to transform the proposed architecture into a robust, auditable tool for AI-assisted EMC and safety compliance.

The remainder of this paper is organized as follows. Section 2 explains the proposed framework and its components. Section 3 presents a literature review organized around these components. Section 4 discusses limitations and future research directions.

2. FRAMEWORK FOR AI-ASSISTED SEMANTIC EVALUATION

Figure 1 depicts the proposed architecture. It is conceptual rather than fully implemented, but each component is grounded in existing research studies.

2.1 Input Layer: Standards Corpus, Measurements, and Queries

On the left, the Standards Corpus contains EMC and safety documents such as IEC 61851-1/-23/-21-2, CISPR and IEC 61000-4 series, EU directives (e.g., 2014/30/EU), and domain-specific protocols, thus vast and fragmented. Alongside it, a Measurements and Reports store holds EMC test logs, immunity/emission tables, numerical simulations, and lab notes. A User Queries/Prompts channel provides natural-language questions or task descriptions from engineers, auditors, or regulators.

2.2 NLP & Preprocessing

This block takes raw standards (often as PDFs), EMC test reports, and turns them into smaller, structured sections. It splits the documents into clauses and sections, extracts tables and figures, and attaches metadata such as standard number, clause ID, test type and limit value. It also normalises units and notation (e.g. different dB μ V / V/m conventions) and maps synonyms in an EMC ontology so that terms like “EUT”, “device under test” or “equipment under test”

are treated as the same concept. In practice, these steps can be implemented with standard document-processing and NLP toolboxes (e.g. PDF parsers such as PyMuPDF or pdfplumber, table-extraction tools like Camelot or tabula-py, and light-weight ontology/phrase-matching libraries such as spaCy), but the framework itself is toolbox-agnostic. Importantly, this is not low-level NLP preprocessing (tokenisation, lemmatisation, etc.) already performed inside the LLM; instead, it focuses on standards-aware structuring and domain-specific normalisation.

Using ideas from work on requirements analysis, legal clause extraction and ontology building (Section 3.2), this block can also flag ambiguous or inconsistent wording and suggest more precise formulations. In short, it prepares a clean, structured view of the standards and reports that the retrieval and LLM reasoning layers can reliably build on.

2.3 Retrieval Layer (RAG Memory)

The **Retrieval Layer** is a hybrid memory: vector embeddings for semantic similarity, plus symbolic indices over clause IDs, ontology nodes, and test artefacts. It supports:

- Vector search with similarity thresholds (e.g., $\alpha \approx 0.85$) and re-ranking.
- Passage-to-evidence mapping: retrieved text is augmented with the relevant tables, figures, or test logs.
- Cross-document linking (e.g., between IEC, CISPR, FCC, and national rules).

This layer turns user queries and test descriptions into grounded context for the LLM. Section 3.3 reviews RAG, knowledge-graph, and enterprise knowledge-base work that motivates this component.

2.4 LLM Reasoning & Decision

At the core sits the LLM, which performs grounded synthesis over retrieved evidence and test data. Instead of answering freely, it:

- Cites specific clauses and figures.
- Carries out basic numeric checks (limits, margins, worst-case combinations).
- Tracks assumptions and flags uncertainty.
- Supports multiple task types: question answering, failure analysis, report drafting, harmonization across standards, and—at Level 3—

reasoning about non-linguistic artefacts such as waveforms or simulation outputs.

Section 3.4 surveys LLM usage in other regulated domains (finance, GDPR, EU law, safety-critical engineering) that inform how such reasoning should be structured and constrained.

2.5 Validation and Interoperability

On the right, Validation and Interoperability Protocols ensure that LLM outputs are usable in real EMC workflows.

- **Validation** includes deterministic numeric checks, traceability IDs that link each statement to its evidence, human-in-the-loop review, and audit logs suitable for EU AI Act-style governance. It also covers benchmark datasets and question suites for quantitative evaluation.
- **Interoperability Protocols** include standards such as ISO 15118 (V2G), OCPP 2.0.1, security events, and remote diagnostics—interfaces through which the framework can exchange data with EV chargers, test equipment, or monitoring systems.

The literature on LLM evaluation, MLOps, ethical filtering, accountable design, and industrial LLM deployment (Section 3.5) provides the methodological background for these blocks.

2.6 Outputs

Finally, the system produces verdicts, clause-linked justifications, PDF/JSON reports, and actionable gaps (e.g., missing tests, contradictory requirements, or design risks). In an EMC lab context, this corresponds to structured test reports; in design and certification contexts, it can drive design-review checklists or automatic generation of technical files.

3. LITERATURE REVIEW ALONG THE FRAMEWORK COMPONENTS

Rather than listing works chronologically, we organize the literature according to the functional blocks of Figure 1. This shows how the proposed framework is anchored in existing research and where the gaps remain.

3.1 Standards Corpus: Complexity, Fragmentation, and EMC Challenges

A first group of works explains why a structured standards corpus is necessary.

Tajima et al. (2025) summarize ongoing CISPR 37 standardization efforts on in-situ measurement methods for large, difficult-to-test systems. They point out

that standards landscape is becoming large and fragmented. More recent work echoes this picture: Keller (2022) describes EMC regulations and standards as a constantly changing, country-dependent landscape, while practical guides for global market access emphasise regional variations and overlapping frameworks between IEC/CISPR, FCC and national schemes (e.g. CompEng, 2023; OECD, 2010; IB-Lenhardt, 2025). Despite this growing complexity, there is still no unified educational resource, so engineers must actively track evolving requirements across labs and countries (Aciu et al., 2018, 2024).

Even though significant efforts have been made to structure and harmonize EMC knowledge—such as long-term technology roadmaps and unified research agendas (van Doorn, 2006; Vision 2020 – The EMC Technology Roadmap) and sector-specific databases that collect EMC know-how to support more consistent limits and procedures (TREND project, 2014)—these initiatives remain fragmented. There is still no cross-domain, machine-readable “standards corpus” that unifies EMC and exposure requirements across countries and product families (CEN-CENELEC EMC webinar, 2021).

These works together necessitates the "Standards Corpus (DB)" block, which is a machine-readable storage space that brings together documents, versions, and jurisdictions that are spread out. This is necessary for any AI-assisted semantic evaluation.

3.2 NLP & Preprocessing: Terminology, Ambiguity, and Ontologies

For the NLP & Preprocessing block, the key challenges are domain terminology (e.g. mapping EMC-specific abbreviations such as *EUT*, *CDN* or detector types like *quasi-peak* to a shared ontology), ambiguity (e.g. vague phrases such as “large EUT”, “sufficient number of test points”, or “minimize unintended coupling”), and converting free text into structured representations—for example, turning the sentence “*the immunity test shall be performed at 3 V/m from 80 MHz to 1 GHz with 1% step size and 3-second dwell time*” into a machine-readable JSON-like form such as {"test_type": "radiated_immunity", "field_strength": "3 V/m", "frequency_range": "80 MHz–1 GHz", "step_size": "1%", "dwell_time": "3 s"}. Addressing these issues in the preprocessing stage, the framework aims to create a clean, well-structured layer of meaning that the later blocks in the framework can reliably use.

These issues are well recognised in both EMC practice and the requirements-engineering literature. The IEEE C63.14-2023 dictionary consolidates EMC/E3/EMP/ESD terminology, yet still leaves practitioners with a substantial cognitive load, suggesting a role for AI-based terminology extraction and normalisation as a complementary mechanism. The characteristic problems of EMC standards—ambiguous wording, incomplete phrasing and inconsistent

terminology—align closely with established requirements-quality defects: a systematic mapping of 105 empirical studies reports that ambiguity, completeness and consistency account for more than half of all documented problems in natural-language requirements (Montgomery et al., 2022). Recent work on prompt underspecification indicates that missing or underspecified instructions can lead to unstable or degraded LLM behaviour (Yang et al., 2025), while studies on automated reasoning from natural-language requirements show that ambiguity and incompleteness substantially constrain both rule-based and LLM-based interpretation unless a structured preprocessing and clarification phase is introduced (Ocleppo, 2025). Collectively, these findings motivate the systematic detection, normalisation and rewriting of EMC requirements as a prerequisite to LLM-based semantic processing.

Evidence from related fields indicates that such preprocessing layers are technically achievable with existing methodologies. Abualhaija et al. (2025) use an LLM+RAG pipeline to extract and structure detailed GDPR (European General Data Protection Regulation)-related privacy requirements. Marczak-Czajka et al. (2025) employ an LLM-driven assistant to identify relevant laws and map them to design decisions and underlying societal values. Finally, ontology-extraction studies—such as LLM-based knowledge-graph approaches to automating regulatory compliance (Chattoraj & Joshi, 2025) or LLM-KG integration for regulatory compliance in safety-management texts (Xiang et al., 2025) —show that standards, clauses, and test artefacts can be mapped automatically into structured knowledge bases, which in turn supports the domain-ontology mapping in Figure 1.

3.3 Retrieval Layer (RAG Memory) and Knowledge Bases

The framework's retrieval layer follows the broader trend of combining LLMs with knowledge bases. Reflecting this trend, Zhang (2024) proposes a decision-making assistant in which a relational/knowledge-graph data model, LLM-based semantic matching, and RPA automation are combined to answer user queries over large, heterogeneous document sets. Gao and Liu (2024) integrate a large language model with a flight-safety knowledge graph, showing that graph-structured domain knowledge can improve the reliability of safety-event analysis and support fault diagnosis in civil aviation. Ouerghemmi and Ertz (2025) review applications of generative LLMs in digital manufacturing and propose an integrative framework that links manufacturing process optimisation, structured data pipelines, and domain knowledge integration in smart-factory workflows. These works collectively justify a RAG memory for EMC, where RAG contains vector search, standards, test reports and ontological entities before the LLM reasoning block.

3.4 LLM Reasoning and Decision in Regulated Domains

The LLM Reasoning & Decision is studied in other highly regulated fields. For example, Cao and Feinstein (2024) show that, for the Basel III *Minimum Capital Requirements for Market Risk* rules, GPT-4 can interpret the regulation and recover the underlying mathematical relations when guided by carefully engineered prompts. Furthermore, loading the regulatory text as images rather than PDFs significantly improves accuracy, especially for formulas, tables, and footnotes, which are also common in EMC regulations. Abualhaija et al. (2025) show that their XTRAREG LLM+RAG approach uses LLM-based reasoning over retrieved GDPR passages to generate access and portability requirements with accuracy above 80%. Marczak-Czajka et al. (2025) present the ADA assistant, an LLM-driven tool that identifies relevant laws and underlying societal values and maps them to candidate system design choices. Nascimento & Lucena (2023) integrate GPT-4 into the MAPE-K loop, where the LLM performs analysis, planning, and reasoning for autonomous agents, enabling adaptive behaviour within a multi-agent system. These studies collectively indicate that LLMs can perform semantics-aware reasoning over regulatory text, provided they are properly preprocessed. This is exactly the role envisioned for the reasoning block in our framework for EMC.

3.5 Validation, Benchmarking, and Human-in-the-Loop Governance

For safety-critical applications, validation is very important. Since LLMs are limited by shallow reasoning and hallucinations, fully autonomous operation is not acceptable, and human-in-the-loop control is unavoidable. Several works emphasise secure deployment as well as control points (Gupta et al. 2023). To strengthen validation, a dual-LLM architecture in which one model generates the content and the other checks it, seems one practical approach as suggested by (Zimin, 2025). Mahr's (2024) industrial reference architecture includes secure deployment, control points. These studies all emphasise the importance of human-in-the-loop review in safety-critical deployments, and safety-critical deployments cannot be delegated to an LLM alone. In our framework, the validation block therefore combines automated checks with mandatory human review steps: it incorporates deterministic numeric checks, traceability IDs linking outputs to clauses, graded benchmark question sets (e.g., ~50 questions per standard family with controlled difficulty), and structured human-review workflows aligned with EU AI Act expectations.

3.6 Interoperability and System-Level Integration

Interoperability is essential for any practical deployment of LLM-based assistants, because these systems must interface with the existing protocols and data channels in industry. For example, an LLM-based industrial process

controller must read sensor data, make decisions, and adapt control policies within the existing control infrastructure (Wang, 2024). LLM-based models such as DistilBERT can match or even exceed traditional ML methods for network intrusion detection when working jointly with structured logs and unstructured textual data (Mercan, 2025). In realistic web-security settings, a two-phase multimodal LLM-based phishing webpage detector achieves high detection rates at high precision and outperforms a state-of-the-art brand-based vision model (Lee et al., 2024). To sum up, these studies demonstrate that LLM-based systems can interface existing protocols in industry. By analogy, similar architectures could interface with EMC-relevant protocols without disrupting existing infrastructure.

3.7 AI for EMC Simulation, Testing, and Non-Language Tasks

Finally, several EMC-specific works point to Level-3 (non-language) opportunities. Several recent works in power electronics, electromagnetics, and materials science point to Level-3 (non-language) opportunities. LLM-assisted surrogate modelling for engineering optimisation already shows that ChatGPT can help select and train regression models that replace expensive simulation iterations while retaining competitive accuracy (Rios et al., 2025). LLM-based optimisation has also been demonstrated directly on power-integrity problems, where a prompting scheme guides decoupling-capacitor placement on complex power distribution networks (Kim et al., 2025). In electromagnetics, LLMs have been embedded into the modelling and inverse-design loop: ChatGPT has been used to learn metamaterial response and perform inverse design from target spectra (Lu et al., 2025). LLM can extract design requirements of equivalent-circuit pipelines for automatic inverse design of frequency selective surface structure (Shen et al., 2025). Survey and perspective papers similarly emphasise that LLM-centred scientific agents can orchestrate simulations, optimisation, and experiments in a closed loop (Ren et al., 2025; Zhang et al., 2025).

These trends are particularly relevant for EMC because high-fidelity, continuous, and repeatable datasets are fundamentally scarce: test setups vary, measurements contain gaps, and re-collecting data is often impossible. Classical statistical models, including deep learning methods—such as LSTM/Transformer predictors—require long, clean, contiguous time series that EMC workflows rarely provide (Li et al., 2024). This motivates the use of LLM as scientific regression tool in EMC, since they rely on semantic descriptions. However, the use of LLM as scientific regression tool is still being developed (Wang et al., 2025).

4. DISCUSSION AND FUTURE WORK

The literature reviewed above shows that many building blocks of the proposed framework already exist, but maybe further developed. For example,

- EMC and safety standards have been extensively analyzed from legal, technical perspectives, creating an opportunity for the creation of a consolidated Standards Corpus.
- NLP and ontology work in the fields of law and requirements engineering shows how ambiguous, complicated regulations can be turned into machine-readable structures. This shows the feasibility of our NLP & Preprocessing block.
- RAG, knowledge-graph, and enterprise knowledge-base systems is similar to Retrieval Layer.
- LLM use in safety-critical engineering demonstrates the feasibility—but also the fragility—of LLM. To overcome fragility, human-in-the-loop design is unavoidable.
- Industrial control deployments point toward integrating LLMs with interoperability protocols and non-language artefacts, enabling Level-3 functionality.

At the same time, the literature also makes clear that no end-to-end system yet exists for EMC and electrical safety that combines all these elements. In particular:

- There are no publicly available benchmark datasets that link EMC standards to test data and evaluation questions.
- The resolution of conflicts of overlapping standards (e.g., IEC 61000-4-4 vs. CISPR 22, or EU vs. national rules) remains largely manual.
- Using LLMs for non-language tasks is still at the research-prototype stage.

Building on these observations, we outline the following research directions, referencing Figure 1:

1. **Corpus and Ontology Construction.** Develop an EMC ontology that captures entities, relations, and temporal provisions. It should also include the information of jurisdictional regions.
2. **Benchmark and Evaluation.** Design a gold-standard question set (e.g., ~50–100 items per standards) that assesses clause retrieval, conflict resolution etc.

3. **Prototype RAG+LLM System.** Implement the core mechanism, that is NLP → Retrieval → LLM reasoning, using open-source tools. At the beginning, only target Level-1 tasks, such as question-answering and report drafting.
4. **Extension to Level-2 Test Analysis.** Develop tools for explainability of reasoning, failures, and for suggestions of redesign, and missing measurements.
5. **Validation and Governance Layer.** Add deterministic numeric validators—such as traceability IDs, uncertainty scoring—and tools for human-in-the-loop review and align these with EU AI Act transparency requirements.
6. **Exploration of Level-3 Non-Language Tasks.** Explore LLM-guided simulation workflows where the model proposes parameter sweeps and interprets results.

Ultimately, the goal is not to replace EMC engineers but to augment them: to transform standards and test data into the semantic domain, where LLMs provide fast, clause-grounded reasoning while humans retain control over safety-critical decisions. The framework proposed in Figure 1, together with the literature-based positioning presented here, provides a structured starting point for a research direction.

5. REFERENCES

- Abualhaija, S., Ceci, M., Sannier, N., Bianculli, D., Lannier, S., Siclari, M., ... & Tosza, S. (2025, September). LLM-assisted extraction of regulatory requirements: A case study on the GDPR. In 2025 IEEE 33rd International Requirements Engineering Conference (RE) (pp. 142-154). IEEE.
- Aciu, L. E., Ghiță-Pîrnuță, O.-A., & Ogrutan, P. L. (2024). Methods used to increase and promote the academic interest of the electromagnetic compatibility course content. *TEM Journal*, 13(4), 3250–3258. <https://doi.org/10.18421/TEM134-59>
- American National Standard Dictionary of Electromagnetic Compatibility (EMC) Including Electromagnetic Environmental Effects (E3)," in ANSI/USEMCSC C63.14-2023 (Revision of ANSI C63.14-2014) , vol., no., pp.1-89, 7 June 2024, doi: 10.1109/IEEESTD.2024.10550174
- Cao, Z., & Feinstein, Z. (2024, October). Large language model in financial regulatory interpretation. In 2024 IEEE Symposium on Computational Intelligence for Financial Engineering and Economics (CIFer) (pp. 1-7). IEEE.
- CEN-CENELEC. (2021, May 4). CEN-CENELEC webinar on EMC: EMC standardization and harmonisation [Webinar slides]. CEN-CENELEC. https://www.cencenelec.eu/media/CEN-CENELEC/Events/Webinars/2021/2021-05-04_emc-webinar_final-presentation.pdf
- Chao, G., Xiaoyuan, L., & Dongfang, Z. (2024, October). The integration of large language model and knowledge graph and its application in flight safety. In 2024 IEEE 6th International Conference on Civil Aviation Safety and Information Technology (ICCASIT) (pp. 491-494). IEEE.
- Chattoraj, S., & Joshi, K. P. (2025). LLM based Knowledge Graph Approach to Automating Medical Device Regulatory Compliance. *IEEE Big Data 2025*.
- CompEng. (2023, July 18). Navigating EMC testing for global markets: A comprehensive guide. *CompEng Blog*. <https://www.compeng.com.au/navigating-emc-testing-global-markets-guide/>
- European Commission. (2015, March 10). Test of Rolling Stock Electromagnetic Compatibility for cross-Domain interoperability (TREND) [FP7 project fact sheet and results]. *CORDIS*. <https://cordis.europa.eu/project/id/285259>
- Fliess, B., Gonzales, F., Kim, J., & Schonfeld, R. (2010). The use of international standards in technical regulation (OECD Trade Policy Papers, No. 102).

- IB-Lenhardt AG. (2025). EMC compliance guide: Standards & requirements. IB-Lenhardt Knowledge Base. <https://ib-lenhardt.com/kb/faq/emc-compliance>
- ISO/IEC. (2018). ISO/IEC Directives, Part 2: Principles and rules for the structure and drafting of International Standards (8th ed.). International Organization for Standardization and International Electrotechnical Commission.
- Keller, R. B. (2022). Design for electromagnetic compatibility—In a nutshell: Theory and practice. Springer.
- Li, Z., Fan, S., Gu, Y., Li, X., Duan, Z., Dong, B., ... & Wang, J. (2024, March). Flexkbqa: A flexible llm-powered framework for few-shot knowledge base question answering. In Proceedings of the AAAI conference on artificial intelligence (Vol. 38, No. 17, pp. 18608-18616).
- Marczak-Czajka, A., Dearstyne, K., & Cleland-Huang, J. (2025, April). Assessing Compliance of Software System Designs to Laws, Regulations, and their Underlying Values. In 2025 IEEE/ACM International Workshop on Designing Software (Designing) (pp. 47-54). IEEE.
- Kim, K., Shin, T., Park, H., Choi, S., Kim, H., Lee, J., ... & Kim, J. (2025, August). Large Language Model-Based Decoupling Capacitor Placement Optimization. In 2025 IEEE International Symposium on Electromagnetic Compatibility, Signal & Power Integrity (EMC+ SIPI) (pp. 291-296). IEEE.
- Lee, J., Lim, P., Hooi, B., & Divakaran, D. M. (2024, September). Multimodal large language models for phishing webpage detection and identification. In 2024 APWG Symposium on Electronic Crime Research (eCrime) (pp. 1-13). IEEE.
- Li, D., Chen, H., Zhou, S., & Zhang, Y. (2024). Electromagnetic Compatibility Test Signal Prediction Method Based on LSTM Network. Journal of Signal Processing Systems, 96(12), 837-848.
- Lu, D., Deng, Y., Malof, J. M., & Padilla, W. J. (2025). Learning Electromagnetic Metamaterial Physics With ChatGPT. IEEE Access.
- Mercan, Ö. B., Toprak, A. G., & Osmanca, M. S. (2025, June). Traditional Machine Learning and Large Language Model Approach in Network Intrusion Detection. In 2025 33rd Signal Processing and Communications Applications Conference (SIU) (pp. 1-4). IEEE.
- Montgomery, L., Fucci, D., Bouraffa, A., Scholz, L., & Maalej, W. (2022). Empirical research on requirements quality: a systematic mapping study. Requirements Engineering, 27(2), 183-209.

- Nascimento, N., Alencar, P., & Cowan, D. (2023, September). Self-adaptive large language model (llm)-based multiagent systems. In 2023 IEEE International Conference on Autonomic Computing and Self-Organizing Systems Companion (ACSOS-C) (pp. 104-109). IEEE.
- Ocleppo, V. (2025). Enhancing requirements engineering with large language models: From elicitation and classification to traceability, ambiguity management and API recommendation (Master's thesis). Politecnico di Torino, Department of Computer Engineering, Cinema and Mechatronics.
- Ouerghemmi, C., & Ertz, M. (2025). Integrating Large Language Models into Digital Manufacturing: A Systematic Review and Research Agenda. *Computers*, 14(8), 318.
- Ren, S., Jian, P., Ren, Z., Leng, C., Xie, C., & Zhang, J. (2025). Towards scientific intelligence: A survey of llm-based scientific agents. *arXiv preprint arXiv:2503.24047*.
- Rios, T., Lanfermann, F., & Menzel, S. (2024, June). Large language model-assisted surrogate modelling for engineering optimization. In 2024 IEEE Conference on Artificial Intelligence (CAI) (pp. 796-803). IEEE.
- Shen, J., Liu, J., Qu, C., Wang, B., Wang, Q., & Wei, Z. (2025). Automatically Inverse Design of FSSs by Conditional Variational Autoencoder and Equivalent Circuit Models with Large Language Models. *IEEE Transactions on Antennas and Propagation*.
- Tajima, K., Mitsuzuka, N., Takabe, M., Kobayashi, E., & Ono, T. (2025). Standardization activities of IEC/CISPR 37 on in situ measurement method for system EMC. In 2025 Asia-Pacific International Symposium and Exhibition on Electromagnetic Compatibility (APEMC). IEEE. <https://doi.org/10.1109/apemc62958.2025.11051816>
- Vagnoni, S., & Palmirani, M. (2025, June). Detecting Transitional Provisions in EU Legislation with Hybrid AI for Better Regulation. In 2025 Eleventh International Conference on eDemocracy & eGovernment (ICEDEG) (pp. 190-197). IEEE.
- van Doorn, M. J. C. M. (2006, February). EMC technology roadmapping: a long-term strategy. In 2006 17th International Zurich Symposium on Electromagnetic Compatibility (pp. 156-159). IEEE.
- Wang, G., Hu, J., Zhou, J., Liu, S., Li, Q., & Sun, Z. (2025). Knowledge-guided large language model for material science. *Review of Materials Research*, 100007.
- Wang, Z., & Qin, H. (2024, June). Intelligent industrial production process automatic regulation system based on LLM agents. In 2024 5th International

Conference on Artificial Intelligence and Electromechanical Automation (AIEA) (pp. 133-137). IEEE.

- Xiang, Y., Lin, P., Luo, Y., Ning, Z., Liu, Y., Liu, K. (2025). Integrating Knowledge Graph and Large Language Model for Safety Management Regulatory Texts. In: Zhou, K. (eds) Computational and Experimental Simulations in Engineering. ICCES 2024. Mechanisms and Machine Science, vol 175. Springer, Cham. https://doi.org/10.1007/978-3-031-81673-4_76
- Yang, C., Shi, Y., Ma, Q., Liu, M. X., Kästner, C., & Wu, T. (2025). What Prompts Don't Say: Understanding and Managing Underspecification in LLM Prompts. arXiv preprint arXiv:2505.13360.
- Zhang, J., Chen, X., Ye, X., Yang, Y., & Ai, B. (2025). Large Language Model in Materials Science: Roles, Challenges, and Strategic Outlook. Advanced Intelligent Discovery, 202500085.
- Zhang, Q., Zhang, Y., Yang, M., & Zhu, Q. (2024, May). Design of Decision-Making Assistant Knowledge Base System Based on Large Language Model. In 2024 IEEE 6th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC) (Vol. 6, pp. 941-945). IEEE.

CHAPTER 12

Assessment of Additive Impact on Fluidized Bed Combustion of Biofuels

Zuhal Akyürek¹

¹ Prof. Dr., Burdur Mehmet Akif Ersoy University, ORCID: 0000-0003-3102-4278

Energy plays a crucial role in achieving sustainable development of countries. In the contemporary world, rising population levels, rapid industrial growth, and continuous developments in technological progresses have all driven energy demand upward. World energy is mainly supplied by fossil fuels (Figure 1). The environmental repercussions associated with fossil-fuel-based energy systems have raised global concerns, as they significantly contribute to climate change and global warming. Although various initiatives aim to reduce these emissions, global CO₂ levels continue to rise and more effective precautions are required to prevent the most severe impacts of climate change. Coal and oil are the largest source of global carbon dioxide emissions with share of 45 % and 33 %, respectively (IEA, 2023a).

Renewable energy sources are becoming increasingly significant to reduce global CO₂ emissions. Renewables are primarily used for electricity generation, and also for heating purposes. Biomass has emerged as a notable energy resource owing to its carbon-neutral nature, ease of production and storage, wide availability, contribution to waste management, and relatively lower cost compared to other renewable energy sources (Millinger et al., 2025). In addition, renewable biofuels represent a growing technological pathway for reducing emissions in the transportation sector that are difficult to decarbonize (IRENA, 2024).

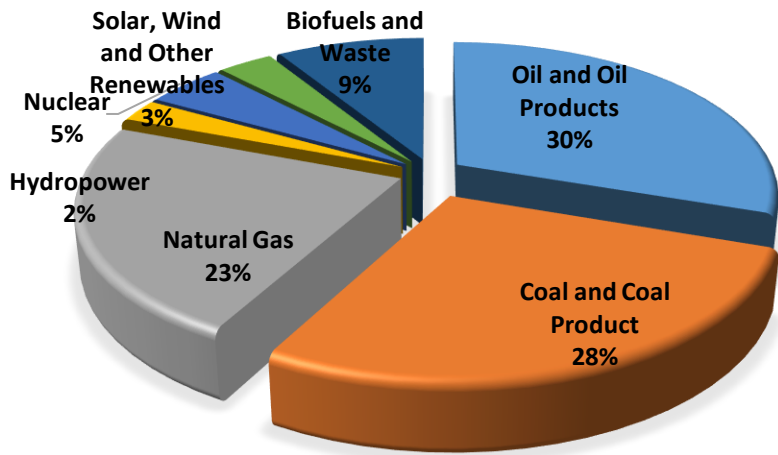


Figure 1. World total energy supply, (IEA, 2023b)

Diverse benefits of biomass/waste make it a promising option for addressing rising energy needs while also supporting countries in achieving their greenhouse gas reduction targets and complying with related policy directives (Rather, 2022).

Unlike other renewable sources such as solar and wind, biomass is a continuous source due to storage capacity. Biomass can be used as a coal substitute in combustion systems to reduce the emission of pollutant release (Wu et al., 2025). Its utilization in electricity generation systems can help to reduce dependency on fossil fuels, provide renewable energy and support the sustainable development (Ersoy and Ugurlu, 2024).

BIOFUEL COMBUSTION IN FLUIDIZED BED COMBUSTION (FBC) SYSTEMS

The combustion of biomass is a multifaceted process that proceeds through a sequence of both homogeneous and heterogeneous reactions. The main stages of this process consist of moisture removal, volatiles release, gasification, char combustion, and reactions occurring in the gas phase (Khan et al., 2009). Fluidized bed combustion systems are recognized as an effective, fuel flexible and high efficiency system among the available technologies for biomass combustion.

Biomass composition and physiochemical properties; and operational parameters are significant factors for biomass utilization in FBC systems. Biofuels often contain considerable amounts of alkali metals, potassium, sulfur, nitrogen, silica, chlorine, and other inorganic constituents, its ash exhibits relatively low melting temperature (Kubay and Akyürek, 2025). Low energy density, high moisture contents and low melting point constituents in biomass can lead to several operational challenges during combustion, including agglomeration, slagging, fouling and corrosion in combustion systems and particulate matter (PM) emissions which hinder large-scale utilization of biomass (Nosek et al., 2020).

Additives on the other hand, can alter the ash characteristics of fuels. In the context of biomass combustion, additives are used to decrease the concentration of problematic inorganic compounds and to raise the ash melting temperature, thereby improving combustion performance, decrease the emission of inorganic particulate matter (PM), regulate greenhouse gas emissions and reducing operational issues in biomass-fired systems (Akyürek, 2024; Wang et al., 2012).

The main categories of additives include those based on calcium, phosphorus, sulfur, and aluminum-silicate compounds. These additives are commonly employed to modify ash behavior, improve combustion performance, and mitigate emissions in biomass-fired systems (Carrol and Finnan, 2015; Akyürek, 2024).

FUEL INDEX EVALUATION

Fuel indices are useful tools to anticipate potential combustion-related issues that could occur when introducing biomass fuel into combustion system. These indices depend on the chemical analysis of fuels and they are generally validated through measurements conducted in both laboratory- and full-scale combustion systems. They allow to produce mitigation strategies before large-scale application of biofuels.

Agglomeration of bed material disrupts the stable operation of fluidized-bed reactors by causing the formation of relatively large particle clusters, in other words agglomerates, which interfere with fluidization during biomass combustion processes (He et al., 2022). Alkali metals in biomass ash are main contributors to agglomeration.

Agglomeration of bed material can be predicted by bed agglomeration index (Bapat et al., 1997):

- Bed Agglomeration Index (BAI)

$$\text{Bed Agglomeration Index, BAI} = \frac{\% Fe_2O_3}{\%(K_2O + Na_2O)} \quad (1)$$

BAI lower than 0.15 indicates agglomeration risk of the fuel during FBC (Zhang et al., 2024).

Slagging and fouling on heat transfer surfaces can be anticipated by some indices as follows (Zhang et al., 2024):

- Alkali Index (AI)

$$\text{Alkali Index, AI} = \frac{(K_2O + Na_2O) kg}{(LHV) GJ} \quad (2)$$

AI lower than 0.17 indicate safe combustion, AI between 0.17 and 0.34 refers to probable slagging/fouling, AI greater than 0.34 refers to certain slagging/fouling risk.

- Base to Acid Ratio ($R_{B/A}$) (Jenkins et al., 1998)

$$R_{B/A} = \frac{\%(Fe_2O_3 + CaO + MgO + K_2O + Na_2O)}{\%(SiO_2 + TiO_2 + Al_2O_3)} \quad (3)$$

Ash deposition tendency is low when $R_{B/A}$ ratio lower than 0.5. $R_{B/A}$ ratio is in between 0.5 and 1, deposition tendency is moderate. $R_{B/A}$ ratio greater than 1 indicates high slagging, fouling risk.

- Slagging Index (SI)

$$SI = \frac{\%(Fe_2O_3 + CaO + MgO + K_2O + Na_2O)}{\%(SiO_2 + TiO_2 + Al_2O_3)} \times \%S \text{ (dry)} \quad (4)$$

Slagging tendency is low when SI is lower than 0.6. SI is in between 0.6 and 2, slagging is moderate. When SI is 2 greater there occurs high slagging, fouling risk.

- Fouling Index (FI)

$$FI = \frac{\%(Fe_2O_3 + CaO + MgO + K_2O + Na_2O)}{\%(SiO_2 + TiO_2 + Al_2O_3)} \times (K_2O + Na_2O) \quad (5)$$

Fouling tendency is low when FI is lower than 0.6. SI is in between 0.6 and 40, fouling is moderate. When FI is higher than 40, high fouling risk expected.

- Slag Viscosity Index (SVI)

$$SVI = \frac{\%(SiO_2 * 100)}{\%(Fe_2O_3 + CaO + MgO + SiO_2)} \quad (6)$$

SVI value higher than 72 indicates low slagging, SVI value between 65 and 72 refers to moderate slagging and SVI lower than 65 indicates high slagging risk.

- Corrosion Index, C_i (Vamvuka et al., 2008)

$$\text{Corrosion Index, } C_i = \frac{\%S}{\%Cl} \quad (7)$$

C_i is lower than 2 indicate possible corrosion. When C_i is greater than 4, there is no corrosion risk.

EFFECT OF ADDITIVES ON FBC OF BIOFUEL

In this study, the possible operational problems during combustion of a biofuel blend composed of poultry waste and pine wood chips were investigated through different fuel indices. The fuel samples were provided from a previous study,

where the effect of addivation on particulate matter and gaseous emissions during combustion and oxyfuel combustion of the fuel blend were examined (Kubay and Akyürek, 2025).

The fuel blend (RBF) composed of 50% poultry manure and 50% pine wood sawdust by weight. Figure 2 shows the major and minor elements examined in raw biofuel (RBF), ZnO addivated biofuel (RBF-ZnO) and MgCO₃ addivated biofuel (RBF-MgCO₃). Table 1 presents the trace element concentrations in raw and addivated biofuel blend ashes. Biofuel ashes composed of a mixture of minerals, salts, and organic compounds.

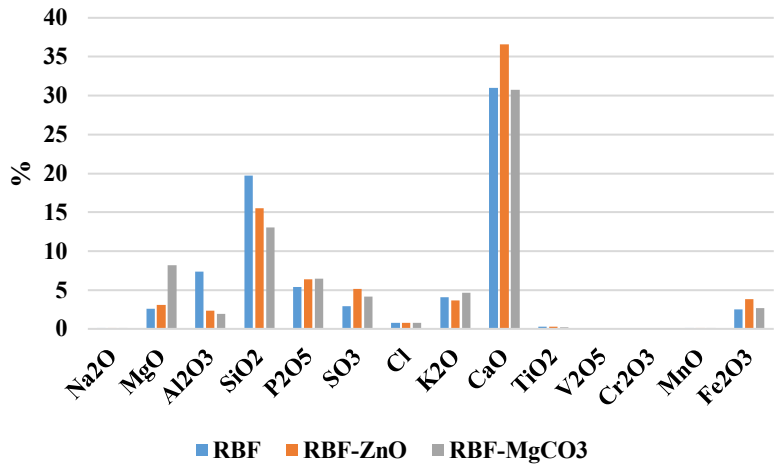


Figure 2. Major and minor elements in biofuel ashes.

Table 1. Trace element concentrations in biofuel ashes with and without additives (Kubay and Akyürek, 2025).

Trace Elements	RBF	RBF-ZnO	RBF-MgCO3
Co	38,5	59	19,4
Ni	95,9	111,2	91,2
Cu	111,4	117,4	121,9
Zn	64,5	3740	53,4
As	6,2	7,5	8,2
Br	6,6	5,5	7,4
Rb	50,3	47,2	39,2
Sr	313	364	333,3
Zr	50	53,6	47,9

Cd	1,8	2,2	1,2
Sn	1,6	1,8	3,1
Cs	8,4	5,9	7
Ba	215,1	223	208,3
Pb	6	5,8	4,7

According to the ash characteristics for raw and addivated biofuels, fuel indices were calculated and evaluated for safe operation in FBC systems. Table 2 shows results obtained for raw and addivated biofuels.

Table 2. Fuel Indices calculated for RBF, RBF-ZnO, RBF-MgCO₃

Fuel Index	RBF	RBF-ZnO	RBF-MgCO₃
BAI	0,59	1,01	0,57
AI	0,29	0,26	0,33
R _(B/A)	1,47	2,61	3,07
SI	1,71	5,32	5,06
FI	6,18	9,81	14,53
SVI	35,33	26,25	23,85
C _i	1,55	2,72	1,12

Bed agglomeration arises when the bed agglomeration index (BAI) falls below 0.15. As can be seen from the table all biofuel blends have low bed agglomeration risk.

Alkali Index (AI) is a measure to determine fouling and slagging formation during combustion processes. All the AI values of biofuel blends are within 0,17-0,34 range which indicate slagging/fouling probability.

Base to acid ratio of fuel blends above 0.50 indicates slagging potential, hence all the fuel blends have shown slagging risk according to R_(B/A) index.

When slagging index is considered, RBF, RBF-ZnO and RBF-MgCO₃ has high slagging potential. FI values of all the fuel blends show moderate fouling potential. According to SVI examination, all the biofuel blends have shown high slagging indication.

Corrosion risk was evaluated with corrosion index (C_i). As can be seen from the table, C_i value of RBF is lower than 2, which refer to possible corrosion, whereas C_i values of addivated biofuels have higher than 2, depicts lower corrosion risk compared to raw biofuel during combustion process.

CONCLUSION

The prediction of biomass combustion mechanism and occurrence of operational problems is challenging process due to complex structure of biomass. Fuel indices are useful tools for rapid evaluation of possible ash-related problems during biomass combustion. These indices can be used as a supporting assessment method for decision making process during adaptation of biofuels in FBC systems.

This study has evaluated influence of ash characteristics on operation performance of fluidized bed combustion of biofuel blend composed of poultry waste and pine wood chips (50/50 wt. %) with/out additives. The results of fuel index analyses revealed there is low agglomeration and corrosion risk exist, on the other hand possible occurrence of slagging/fouling during combustion of biofuel in FBC systems. Addivation has seen to improve ash characteristics of the biofuel blend.

ACKNOWLEDGEMENT

This work was supported by Burdur Mehmet Akif Ersoy University through a research project BAP-0902-YL-23.

REFERENCES

- Akyürek, Z. (2024) The Effect of Additives on Particulate Matter Emissions from Biomass Combustion, Aerosol Optical Depth and Precipitation - Measuring particle concentration, health risks and environmental impacts. Springer Nature, Eds. Sneha Gautam, Roshini Praveen Kumar, Cyril Samuel, ISBN:978-3-031-55835-1.
- Bapat D.W., Kulkarni S.V., Bhandarkar V.P. (1997). Design and operating experience on fluidized bed boiler burning biomass fuels with high alkali ash. Proceedings of the 14th International Conference on Fluidized Bed Combustion, 165.
- Carroll J.P., Finnan J.M. (2015) The use of additives and fuel blending to reduce emissions from the combustion of agricultural fuels in small scale boilers. *Biosystems Engineering* 129, 127-133. <https://doi.org/10.1016/j.biosystemseng.2014.10.001>
- Ersoy A.E., Ugurlu A. (2024). Bioenergy's role in achieving a low-carbon electricity future: A case of Türkiye. *Applied Energy*, 372, 123799. <https://doi.org/10.1016/j.apenergy.2024.123799>
- He Z.M., Cao J.P., Zhao X.Y. (2022). Review of Biomass Agglomeration for Fluidized-Bed Gasification or Combustion Processes with a Focus on the Effect of Alkali Salts. *Energy & Fuels*, 36(16), 8925–8947. <https://doi.org/10.1021/acs.energyfuels.2c01183>
- International Energy Agency (IEA), CO2 emissions from fuel combustion in World (2023a) <https://www.iea.org/world/emissions>
- International Energy Agency (IEA), World total energy supply, (IEA, 2023b) <https://www.iea.org/world/energy-mix>
- International Renewable Energy Agency (IRENA) Decarbonising Hard-to-debate with renewables (2024). https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2024/Apr/IRENA_G7_Decarbonising_hard_to_abate_sectors_2024.pdf
- Jenkins B.M., Baxter L.L., Miles Jr. T.R., Miles T.R. (1998). Combustion properties of biomass. *Fuel Processing Technology*, 54 (1–3), 17. [https://doi.org/10.1016/S0378-3820\(97\)00059-3](https://doi.org/10.1016/S0378-3820(97)00059-3)
- Khan A.A., de Jong W., Jansen P.J., Spliethoff H., (2009). Biomass combustion in fluidized bed boilers: Potential problems and remedies, *Fuel Processing Technology* 90, 21-50. <https://doi.org/10.1016/j.fuproc.2008.07.012>
- Kubay, H.H., Akyürek, Z. (2025) The effect of additives on particulate matter and gaseous emissions from combustion and oxy-fuel combustion of a biofuel

- blend composed of poultry manure and pine wood chips. *Environ Sci Pollut Res* 32, 4090–4102 <https://doi.org/10.1007/s11356-025-35955-x>
- Millinger, M., Hedenus, F., Zeyen, E. et al. Diversity of biomass usage pathways to achieve emissions targets in the European energy system. *Nat Energy* 10, 226–242 (2025). <https://doi.org/10.1038/s41560-024-01693-6>
- Nosek, R., Werle, S., Borsukiewicz, A., Żelazna, A., & Łagód, G. (2020). Investigation of Pellet Properties Produced from a Mix of Straw and Paper Sludge. *Applied Sciences*, 10(16), 5450. <https://doi.org/10.3390/app10165450>
- Rather R. A., Wanib A. W., Mumtaz S., Padder S.A., Khand A.H., Almohanae A.I., Almojile S.F., Alam S.S., Babab T.R. (2022). Bioenergy: a foundation to environmental sustainability in a changing global climate scenario, *J King Saud Univ Sci*, vol. 34. <https://doi.org/10.1016/j.jksus.2021.101734>
- Sommersacher P., Brunner T., Obernberger I. (2011). Fuel Indexes: A Novel Method for the Evaluation of Relevant Combustion Properties of New Biomass Fuels.
- Vamvuka D., Zografos D., Alevizos G. (2008) Control methods for mitigating biomass ash-related problems in fluidized beds, *Bioresource Technology*, 99, 3334–3344. <https://doi.org/10.1016/j.biortech.2007.07.049>
- Wang, L., Skjevrak, G., Hustad, J., Grønli, M. (2021). Sintering characteristics of sewage sludge ashes at elevated temperatures. *Fuel Process. Technol.* 96, 88–97. <https://doi.org/10.1016/j.fuproc.2011.12.022>
- Wu, M., Wei, K., Jiang, J. et al. (2025). Advancing green sustainability: A comprehensive review of biomass briquette integration for coal-based energy frameworks. *Int J Coal Sci Technol* 12, 44. <https://doi.org/10.1007/s40789-025-00779-0>

CHAPTER 13

Effects of Solar Panel Cooling Techniques on Energy Efficiency

Ibrahim Gunes¹

¹ Assoc. Prof. Dr., Istanbul University-Cerrahpasa, Engineering Faculty, Electrical and Electronics Engineering Department, Avcilar,34320 Istanbul.
ORCID: 0000-0003-1032-1134

Introduction

Photovoltaic (PV) technology has become one of the most rapidly expanding renewable energy sources worldwide. As solar deployment increases, understanding the performance limitations of PV modules under real environmental conditions becomes increasingly important. Among these limitations, temperature rise stands out as a critical factor that directly reduces power output and system efficiency.

PV modules convert only a fraction of incident solar irradiance into electrical energy; the remaining portion is transformed into heat, causing module temperatures to rise significantly above ambient conditions. Elevated temperature conditions affect semiconductor material properties, shift I–V characteristics, and accelerate degradation mechanisms such as encapsulant browning, solder bond failure, and delamination. The thermal challenges vary with climate, panel configuration, mounting structure, and system design. Furthermore, thermal performance has become more crucial with the global adoption of bifacial modules, floating PV (FPV), building-integrated PV (BIPV), and non-planar PV structures, all of which have unique cooling behaviors.

Understanding the fundamental thermal challenges of PV systems provides the foundation for developing effective cooling strategies. As PV installations diversify across climates and applications, addressing temperature-induced losses becomes essential for maximizing energy yield and long-term system reliability.

The thermal behavior of PV cells is governed by a balance between absorbed solar radiation, electrical energy conversion, and heat dissipation. A detailed understanding of heat transfer phenomena is essential for analyzing cooling strategies.

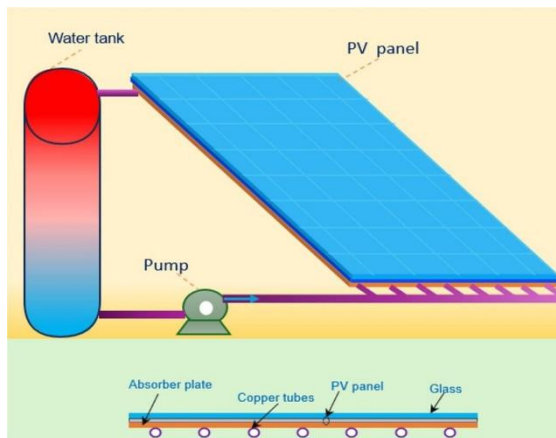


Figure1. Schematic view of a closed-loop PV cooling system

PV cells absorb photons across a broad spectrum, but only photons above the bandgap generate electron–hole pairs. Excess photon energy and sub-bandgap radiation are converted into heat. Thermal processes include conduction within the module layers, convection to the surrounding air, and radiation exchange with the sky. The magnitude of these processes depends on module construction, mounting configuration, and environmental parameters such as wind speed and irradiance.

The thermal behavior of PV cells reveals the importance of engineering designs that enhance heat dissipation. Recognizing the basic heat transfer mechanisms allows researchers to evaluate cooling methods more systematically.

Impact of Temperature on PV Electrical Performance

Temperature has a well-established negative effect on the electrical performance of PV modules, predominantly manifested through voltage reduction. As temperature increases, the intrinsic carrier concentration of semiconductor materials rises, causing a decrease in the open-circuit voltage (V_{oc}) and modest changes in short-circuit current (I_{sc}). The combined effect results in a significant reduction in maximum power point (MPP) and overall system efficiency. High temperatures also induce mismatch losses, affect inverter performance, and accelerate material degradation. Long-term exposure to elevated temperatures contributes to hotspots, microcrack propagation, and encapsulant discoloration. The reduction in electrical performance under elevated temperatures justifies the integration of cooling strategies to preserve energy yield and system lifetime.

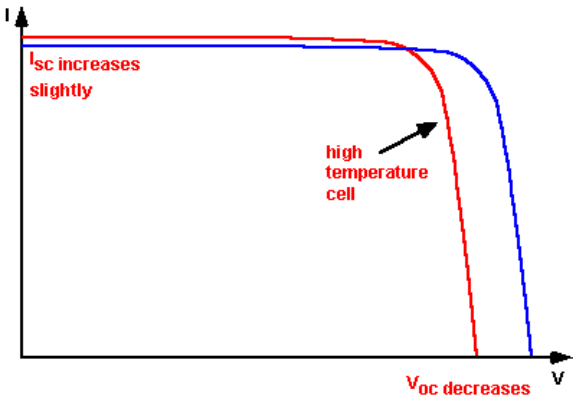


Figure2. Effect of temperature on the I-V characteristics of PV.

Environmental and Climatic Factors Influencing PV Heating

Environmental conditions directly influence PV module operating temperature and cooling behavior. Major climatic factors include solar irradiance,

ambient temperature, wind speed, humidity, atmospheric pollution, and albedo. Mountain regions, deserts, coastal environments, and urban areas each exhibit unique thermal profiles for PV modules. Dust and pollution reduce electrical output while increasing module temperature by altering surface optical properties. High humidity environments reduce cooling by limiting convective exchanges.

On floating PV (FPV) systems, water temperature, humidity gradients, and evaporation significantly influence the thermal environment. A comprehensive understanding of climatic effects is crucial for selecting and optimizing cooling techniques in different geographic regions.

Overview of Solar Panel Cooling Techniques

PV cooling methods can be grouped into passive, active, and hybrid categories. Passive cooling techniques rely on natural heat dissipation through improved airflow, radiative cooling, and enhanced thermal conduction. Active cooling methods employ external energy sources—such as fans, pumps, or refrigeration systems—to remove heat from the panel surface or rear side. Hybrid cooling systems combine passive and active mechanisms to achieve higher thermal control while optimizing energy consumption. This classification provides the structural basis for analyzing the advantages and limitations of each cooling group in subsequent chapters.

Passive Cooling Methods: Design, Materials, and Performance, Passive cooling aims to enhance heat dissipation without consuming external energy.

Common passive strategies include:

- Heatsinks and fins to increase surface area
- Rear-side ventilation enhancements
- High-conductivity backsheet materials
- Radiative cooling coatings designed to emit heat efficiently
- Thermally optimized mounting structures
- These techniques are attractive for large-scale deployments due to cost-effectiveness and reliability.

Passive cooling is an economically viable approach, particularly suitable for distributed systems, rooftops, and large utility-scale plants.

Active Cooling Methods: Technologies and Energy Consumption Trade-Offs:

Active cooling systems provide more effective temperature control at the cost of energy consumption. Techniques include forced-air cooling, liquid cooling with heat exchangers, sprinkler systems, and thermo electric coolers. Active methods can significantly lower module temperature but require detailed evaluation of net energy gain, operational cost, maintenance requirements, and system complexity. Active cooling is most appropriate for concentrated PV (CPV), research-oriented platforms, and environments with extreme temperatures.

Hybrid Cooling Approaches Integrating Passive and Active Methods: Hybrid cooling seeks to balance performance gains with energy efficiency. Examples include liquid-cooled systems with natural convection enhancement, PCM-integrated modules supplemented with airflow systems, and radiative coatings combined with evaporative cooling. These systems offer optimized performance while reducing energy consumption relative to fully active solutions. Hybrid cooling provides a promising middle ground with strong applicability for modern large-scale installations.

The electrical performance of photovoltaic (PV) modules is strongly influenced by operating temperature, with elevated temperatures causing significant reductions in power output and long-term reliability. As global solar energy deployment accelerates, particularly in high-irradiance and warm climatic regions, interest in advanced cooling strategies has increased. This paper provides a comprehensive review of solar panel cooling techniques, including passive, active, and hybrid approaches. The thermal behavior of PV modules is examined, along with the thermodynamic mechanisms that govern heat generation and dissipation. Various cooling solutions—ranging from heat sinks, airflow optimization, and phase change materials (PCMs) to liquid-based cooling and thermoelectric systems—are evaluated in terms of effectiveness, energy consumption, cost, and scalability. The study highlights the potential benefits of advanced cooling technologies for improving energy yield, reducing degradation rates, and extending system lifetime. Additionally, challenges associated with implementation, economic feasibility, and integration into large-scale PV farms are discussed. The review concludes by outlining emerging trends and future research directions in PV thermal management.

Solar photovoltaic technology has become one of the fastest-growing renewable energy sources worldwide, driven by decreasing module prices, policy

incentives, and advancements in manufacturing. Despite this progress, the efficiency of PV systems remains constrained by thermal losses. Under standard test conditions (25°C cell temperature), PV modules exhibit rated performance; however, in real-world environments, module temperatures often reach 45–80°C, depending on irradiance level, mounting configuration, ambient temperature, and wind speed. Cooling techniques have therefore become an essential research area for optimizing PV performance, especially in regions such as the Middle East, North Africa, South Asia, and Southern Europe, where PV modules frequently operate above their nominal temperature range. This article reviews the state-of-the-art cooling strategies, categorizing them into passive, active, and hybrid approaches, and assesses their potential contributions to energy generation efficiency.

Thermal Behavior of Photovoltaic Modules

The thermal behavior of PV modules is governed by a balance between absorbed solar radiation, electrical conversion, and heat dissipation. Typically, only 15–22% of incident solar irradiance is converted into electrical energy, while the remainder becomes thermal energy. The primary heat transfer mechanisms affecting module temperature include:

Conduction, Heat is conducted through module layers including the glass, encapsulant, cells, and backsheets. The thermal conductivity of these materials influences temperature gradients within the module.

Convection, Natural or forced convection dissipates heat to the surrounding air. The magnitude of convective cooling depends heavily on wind speed, module inclination, mounting height, and geometry.

Radiation, PV modules emit thermal radiation toward the sky, which has an effective temperature significantly lower than ambient conditions. Radiative heat loss is enhanced by materials with high emissivity in the infrared spectrum.

Understanding these mechanisms is essential for designing cooling strategies that target the dominant heat transfer routes under specific environmental conditions.

Cooling techniques vary significantly in performance, cost, reliability, and scalability. Passive methods are most suitable for widespread deployment due to their simplicity and low cost. Active methods deliver superior cooling but require optimization to ensure positive net energy gain. Hybrid strategies provide a balanced solution and are increasingly explored in modern PV system design. Economic considerations are crucial. In many cases, a temperature reduction of

just 5–8°C can justify the investment in cooling technologies by improving annual energy yield by 3–6%. However, cost–benefit analyses must account for installation complexity, maintenance costs, and climatic suitability.

Environmental factors also play a key role. Dusty or polluted regions may benefit more from cooling systems that double as cleaning mechanisms (e.g., sprinklers). Regions with water scarcity must prioritize passive or hybrid methods with minimal water use. Additionally, cooling contributes to long-term reliability. Minimizing thermal cycling reduces mechanical stresses within the module, lowering the risk of long-term degradation.

Solar panel cooling techniques represent a vital solution for mitigating temperature-related efficiency losses in photovoltaic systems. Passive, active, and hybrid cooling approaches each offer unique advantages depending on climate, application, and economic constraints. Passive systems are cost-effective and scalable, active systems provide strong cooling performance, and hybrid systems offer optimized energy use.

Future research should focus on: Development of nanomaterial-based coatings, advanced PCM formulations with improved thermal stability, AI-driven thermal management systems, and integrated cooling solutions for bifacial, floating, and non-planar PV structures. As global PV deployment expands into increasingly challenging environments, effective thermal management will become an essential component of solar system design, ensuring higher efficiency, longer module lifespan, and improved return on investment.

Cooling Solutions for Floating Photovoltaic (FPV) Systems

Floating photovoltaic (FPV) systems have emerged as a rapidly expanding segment of the global solar energy market due to their advantages in land-use efficiency, enhanced cooling potential, and increased power generation. While FPV installations benefit from the inherent cooling effect of water bodies, module temperatures can still rise significantly under high irradiance, low wind conditions, and stagnation of water surfaces. As the scale and complexity of FPV installations increase, the need for advanced cooling solutions becomes more pronounced. This chapter provides a comprehensive review of cooling mechanisms applicable to FPV systems, including water-based passive cooling, convective enhancement through hydrodynamic effects, hybrid cooling structures, thermally optimized float designs, and active water circulation systems. Environmental factors influencing thermal performance—such as water temperature, humidity, turbidity, and evaporative processes—are analyzed in detail.

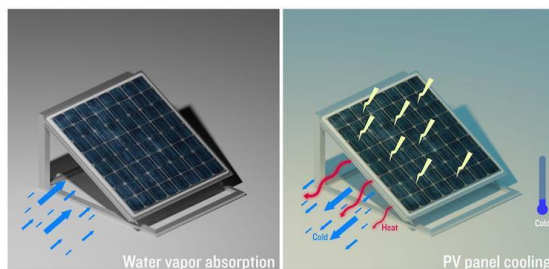


Figure3: Cooling Systems on Floating PV

The chapter also identifies research gaps, highlights engineering challenges, and proposes future directions for improving thermal management in FPV systems.

Floating photovoltaic (FPV) technology has gained significant traction as a promising alternative to conventional land-based solar installations. By placing PV arrays on lakes, reservoirs, quarry basins, hydropower dams, and other water surfaces, FPV systems offer unique operational benefits. These include reduced land-use conflicts, minimized shading losses due to surface uniformity, and potential synergies with hydropower infrastructure.

A defining characteristic of FPV systems is their interaction with the aquatic environment, which strongly influences module temperature. Since operating temperature is a key determinant of PV performance, understanding and optimizing thermal behavior becomes essential. While early FPV research assumed that water provides sufficient passive cooling, extensive experimental evidence now reveals that thermal performance is highly dependent on flotation design, environmental conditions, and system-scale dynamics.

This chapter reviews state-of-the-art cooling solutions for FPV systems, addressing the fundamental thermal mechanisms and engineering strategies that aim to maximize the natural cooling advantage of water surfaces while mitigating temperature fluctuations under varying climatic conditions.

Thermal Behavior of Floating Photovoltaic Systems

Heat Transfer Mechanisms in FPV

FPV systems are governed by three primary pathways of heat exchange:

- (1) Convection to air, influenced by wind speed and module elevation;
- (2) Conduction to the float body and underlying water;

(3) Radiation to the sky, moderated by humidity, cloud coverage, and atmospheric temperature.

Compared to land-based PV systems, the proximity to water significantly modifies these pathways. Water bodies typically absorb heat slowly and maintain a stable temperature, providing a constant thermal sink for heat dissipation from PV modules. This results in a typical reduction of module temperature by 5–12°C relative to ground-mounted systems. However, water temperature varies seasonally and diurnally, particularly in shallow reservoirs.

Environmental Factors Affecting FPV Cooling,

Multiple environmental parameters influence FPV thermal behavior:

Water temperature: A key variable determining the magnitude of conductive cooling.

Humidity: High humidity reduces radiative cooling at night but enhances evaporative cooling.

Wind speed: Low wind environments weaken natural convection.

Water movement: Currents, waves, and turbulence increase convective heat transfer at the module underside.

Turbidity and albedo: Influence both thermal absorption and light reflection.

Evaporative processes: Provide cooling but depend on air–water vapor gradients.

Understanding these variables is crucial to designing effective FPV cooling strategies tailored to local environmental conditions.

Passive Cooling Solutions for FPV Systems

Passive cooling is especially attractive for FPV applications due to its low cost, reliability, and compatibility with large-scale installations.

Natural Water-Based Cooling: FPV modules are positioned on floating structures, maintaining a small air gap between the module and water surface. The water mass acts as a thermal reservoir, stabilizing temperatures and providing steady cooling through conduction. Thermal measurements show significant morning–afternoon temperature moderation, improving daily energy yield.

Advantages: No energy consumption, High reliability, Effective in moderate climates, Reduces thermal cycling and long-term degradation Limitations, Reduced effectiveness in warm, stagnant waters, Strongly dependent on seasonal temperature variations, Limited control over cooling intensity

Optimized Float Designs for Enhanced Cooling, The geometry and material of floating platforms play a critical role in thermal performance. FPV floats constructed from high-density polyethylene (HDPE) exhibit moderate thermal conductivity, which limits conduction. Research suggests:

Metal-reinforced float structures improve conductive pathways.

Thermally porous floats, allowing water splashing, increase convective cooling.

Underside open-channel designs leverage natural water flow beneath panels.

These structural modifications can reduce module temperatures by 3–7°C beyond natural cooling.

Ventilation-Enhanced FPV Platforms, Increasing air circulation beneath FPV modules significantly improves convective cooling. This can be achieved by:

Increasing float height

Designing slotted or perforated float surfaces

Configuring panel tilt to maximize airflow

Such designs are particularly effective in low-wind environments.

Radiative Cooling Coatings in FPV, Selective thermal-emissive coatings applied to the rear or front surfaces of FPV modules enhance nighttime cooling and reduce thermal buildup during the day. When combined with FPV's naturally high humidity environment, radiative cooling can decrease evening module temperatures faster than on land-based systems.

Active and Hybrid Cooling Strategies for FPV

While FPV benefits from inherent passive cooling, additional mechanisms may be necessary in hot climates or during periods of high irradiance.

Active Water Circulation Systems, Circulating water beneath the module array or through integrated channels increases heat transfer efficiency. Approaches include:

- Pumped water loops
- Submerged cooling plates
- Slow-flow cooling channels integrated into float structures

These systems reduce module temperatures by up to 15°C. However, energy consumption and mechanical complexity limit scalability.

Sprinkler-Based Evaporative Cooling

Water sprinkling on FPV module surfaces creates evaporative cooling. FPV installations have an advantage here because water is readily available. Performance improvements include:

- 8–12°C temperature reduction
- Increased module cleaning
- Decreased soiling losses

Challenges include mineral deposits, increased maintenance, and potential water wastage.

Hybrid PCM–Water Cooling Systems, Phase change materials (PCMs) integrated into float structures or module backsheets can store thermal energy during peak irradiance and release it at night. When combined with water exposure, PCM systems demonstrate improved thermal cycling stability.

Integrated PV/T (Photovoltaic–Thermal) in FPV Systems, FPV offers a unique platform for hybrid PV/T cooling systems, where water acts as both coolant and heat storage medium. Such systems enable:

- Simultaneous electricity and thermal energy production
- Enhanced cooling through liquid circulation
- Potential integration with aquaculture or water purification applications

Environmental and Operational Considerations

Cooling performance must be evaluated in the context of environmental impacts, system durability, and operational constraints.

Water Quality and Biofouling, FPV systems in natural water bodies may face challenges such as algae growth, sediment accumulation, and biological fouling, which can reduce heat transfer efficiency and affect float longevity.

Mechanical Reliability Under Thermal Cycling, Thermal gradients between modules, float bodies, and water may induce material fatigue. Cooling solutions must therefore minimize abrupt thermal transitions.

Variability in Water Levels and Hydrodynamics, Reservoir drawdown, wave patterns, and seasonal water fluctuations influence cooling effectiveness and require adaptive float designs.

Environmental Impact Mitigation, Cooling interventions should not negatively affect aquatic ecosystems. For example: Sprinkler discharge should avoid chemical contamination, Active cooling should minimize noise and energy consumption.

Discussion

The unique characteristics of FPV systems present both opportunities and challenges for efficient thermal management. Natural water-based cooling offers significant advantages compared to terrestrial PV, but it is not sufficient under extreme climatic conditions. Optimized float structures, enhanced ventilation designs, and hybrid active–passive cooling techniques show strong potential for improving performance.

Economic and environmental assessments are crucial. While active cooling methods deliver strong thermal control, their energy consumption can offset electrical gains. Passive approaches remain the most feasible for large-scale FPV deployments due to scalability, low cost, and reduced maintenance.

Future FPV cooling strategies will likely integrate:

Real-time thermal monitoring

AI-based cooling control algorithms

Advanced materials with higher thermal conductivity

Multi-functional platforms combining cooling, energy storage, and water resource management

Cooling solutions for FPV systems are essential for maximizing performance, reliability, and long-term sustainability. FPV installations inherently benefit from proximity to water bodies, but optimized cooling strategies can further enhance module efficiency by reducing thermal stress and improving daily energy yield.

Passive cooling methods—float optimization, ventilation enhancement, and radiative coatings—offer robust and cost-effective solutions. Active cooling techniques, while effective, must be evaluated for net energy gain and operational feasibility. Hybrid approaches hold promise for high-temperature environments and specialized FPV applications.

With FPV deployment continuing to expand globally, effective cooling strategies will play a central role in unlocking the full potential of floating solar technologies, ensuring enhanced performance, durability, and integration with broader water–energy management systems.

Acknowledgments

Dr. Gunes acknowledges support of Turkish Petroleum Foundation under grant International Research Support which will support his continued research in this area and Dr. Gunes wishes to thank Texas A&M University and the Renewable Energy & Advanced Power Electronics Research (REAPER) Lab for hosting him as a Visiting Scholar, which made this book chapter possible.

References

1. Kjeldstad, T., Lindholm, D., Marstein, E., & Selj, J. (2021). Cooling of floating photovoltaics and the importance of water temperature. *Solar Energy*, 218, 544–551.
2. Ranjbaran, P., & Souri, A. (2019). A review on floating photovoltaic (FPV) power generation units. *Renewable and Sustainable Energy Reviews*, 110, 332–347.
3. Essak, L., & co-authors. (2022). Floating photovoltaics: A review. *Proceedings*, 4(3), 46. (MDPI)
4. International Energy Agency – Photovoltaic Power Systems (IEA-PVPS). (2025). Floating PV Plants – Review and Technical Guidance (T13-31 report). IEA-PVPS.
5. González, R. C., & co-authors. (2025). Energy production enhancement and evaporation reduction by floating PV systems: a field and modeling assessment. *Applied Energy / Energy Policy* (article).
6. Razeghi Jahromi, D., & co-authors. (2024). A comprehensive analysis of floating photovoltaic systems: performance comparison and deployment considerations. *Journal of Renewable and Sustainable Energy (JREE)*, (2024), Article.
7. Dörenkämper, M., & co-authors. (2023). Modeled and measured operating temperatures of floating PV arrays: comparisons and implications. *Energies*, 16(20), 7153.
8. Kjeldstad, T., & Lindholm, D. (2021). Cooling of floating photovoltaics and the importance of water temperature — case study of membrane technology. *Solar Energy*, 218, 544–551.
9. Amin, M. M., & co-authors. (2023). Floating photovoltaic performance evaluation using a novel cooling system: case study and CFD analysis. [Conference/Journal article — Simulation and experimental results].
10. Lebbi, M., & co-authors. (2025). Solar photovoltaic–thermal (PV-T) hybrid systems: detailed review with emphasis on floating configurations. *Renewable Energy Reviews/Elsevier* (review article).
11. Wang, C., & co-authors. (2025). A comprehensive review of research works on cooling photovoltaic systems: immersion, floating and hybrid cooling. *Energies*, 18(16), 4305.
12. Liu, G., & co-authors. (2024). Review of recent offshore floating photovoltaic systems: design and thermal aspects. *Journal of Marine Science and Engineering*, 12(11), 1942.

13. Solomin, E., & co-authors. (2021). Hybrid floating solar plant designs: a review of configurations and thermal management. Conference/Technical Review.
14. Marini, V., & co-authors. (2025). Analysis of site suitability for floating photovoltaic plants: climatic and thermal constraints. *Sustainability*, 17(13), 5681.
15. Razeghi Jahromi, D., et al. (2024). Thermal advantages and design considerations for FPV: a systematic evaluation. *Journal of Renewable Energy Engineering*.
16. Dörenkämper, M., et al. (2023). Measured temperature fields of floating PV arrays and implications for cooling strategies. *Energies*, 16(20), 7153. (similar coverage as #8)
17. Santos, F. R., & co-authors. (2022). Field study of FPV impacts on reservoir evaporation and PV efficiency. *Energies*, 15(17), 6274. (field data on water–energy interactions)
18. Kjeldstad, T. (2021). Assessment of Ocean Sun membrane FPV technology: thermal contact, water temperature, and yield. Technical paper / Solar Energy study.
19. IEA-PVPS (2025). Floating PV Plants: Technical and Operational Considerations. IEA Technical Report.
20. Dörenkämper, M., (2023). Modeling convection and conduction under floating PV modules: influence of water movement. *Applied Thermal Engineering / Energies*.
21. Wang, X., & co-authors. (2024). Experimental evaluation of evaporative cooling for improving PV efficiency: implications for FPV sprinkler applications. *Renewable Energy / Experimental Studies*.
22. Fekih, A., & co-authors. (2025). A review of thermal behavior of floating photovoltaic under tropical climates. *IEEE LATAM Transactions / Special Issue*.
23. González, R. C., & co-authors. (2025). Comparative field trials of ground vs floating PV: thermal, soiling and evaporation metrics. *Energy Conversion and Management*.
24. Ocean Sun AS / Industry report. (2019). Floating Solar membrane technology: thermal contact and demonstrated cooling performance. (Industry white paper / case study).

25. Solomin, E., & co-authors. (2021). Floating PV hybridization with hydro and PV/T: thermal synergy and cooling strategies. Journal or Conference Proceedings.
26. EPJ PV / KTH thesis (2024). The water-cooling effect on floating photovoltaic plants: master's thesis and measurements. KTH Royal Institute of Technology.
27. Amin, M. M., et al. (2023). Novel natural cold water cooling design for floating PV: ANSYS CFD simulations and field tests. Engineering Simulation Journal / Conference.
28. Santos, F. R. (2022). FPV system environmental interactions: water temperature, humidity and evaporative cooling. *Energies*, 15(17), 6274. (repeats authorship focus useful for environmental context)
29. Comprehensive industrial and academic FPV bibliography. (2022–2025). Collection of FPV design, cooling and environmental studies compiled from MDPI, Elsevier, IEA-PVPS, and peer-reviewed journals.

CHAPTER 14

Evaluation of Environmental Recalcitrance of Biochar

Zuhal Akyürek¹

¹ Prof. Dr., Burdur Mehmet Akif Ersoy University, ORCID: 0000-0003-3102-4278

Climate change has become a crucial global problem toward the end of twentieth century. The rapid progress of climate change has intensified the actions for reducing atmospheric carbon dioxide emissions (Hanson et al., 2025). Anthropogenic greenhouse gas emissions resulted from industry and energy production have triggered the global warming problem and led to air pollution and many ecological hazards (Dubey and Arora, 2022; Saravanan et al., 2022).

With the Paris Agreement it is aimed to hold the global temperature rise to well below 2 °C above conditions prior to industrialization. This target necessitates applying strict strategies for reduction of carbon dioxide emissions from fossil fuel energy generation and from industrial activities. Figure 1 shows the variation in anthropogenic carbon dioxide emissions in the last decades. As can be seen from the figure, the CO₂ emissions have continuously risen. Especially, in developing economies, the rise in energy-related CO₂ emissions is mainly associated high demand for energy due to growth in population and economy. The global CO₂ emissions has reached to 37.8 Gt in 2024 (IEA, 2025).

There are mitigation pathways for anthropogenic CO₂ emissions and achieving carbon neutrality such as applying energy efficiency measures, using low carbon fuels, utilizing nuclear and renewable energy sources, and carbon capture and sequestration (CCS) (IPCC, 2005). Carbon sequestration must provide long-lasting and substantial impact (Lehmann, 2007). Among the alternative solutions, bioenergy with carbon capture and storage (BECCS) functions as a clean energy solution that lowers emissions and sequesters carbon and offset greenhouse gas emissions (Deng et al., 2025).

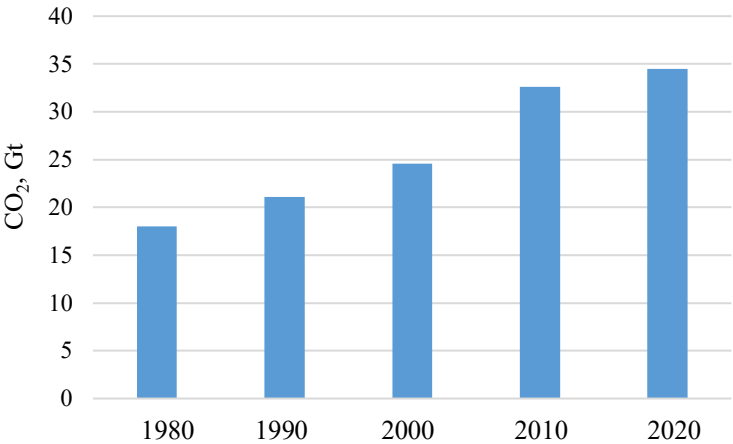


Figure 1. The variation of global CO₂ emissions from energy generation and industrial processes (IEA, 2025)

BIOCHAR AS A CARBON NEUTRAL SOURCE

Biochar, a carbon-rich material generated through the thermal decomposition of biomass in oxygen-free medium. Biochar can directly and indirectly contribute lowering carbon emissions. It provides carbon capture and sequestering from the atmosphere and hence supports mitigation of global warming and climate change impacts (Mishra et al., 2023).

Biochar can be considered a carbon-neutral material, as carbon is blocked in biochar structure during pyrolysis. Biochar can sequester more carbon from the atmosphere with respect to raw plant material, and, thus, it becomes is a carbon-negative material rather than carbon neutral source (Lehmann, 2007). Biochar carbon sequestration can reduce 12% of anthropogenic CO₂ emissions by 2050 (Fawzy et al., 2021).

Biochar is a multi-beneficiary product due to its unique properties of high nutrients content, porosity and water holding capacity (Akyürek et al., 2021). Therefore, it can be used for soil amendment purposes and in water decontamination process as bio-sorbent. Soil management to enhance soil organic carbon (SOC) storage mitigates rising CO₂ emissions from land use (Song et al., 2012). Biochar promotes soil quality and improves soil organic carbon (SOC) sequestration. Due to their long-term stability, these fractions can remain in the soil for millennia, thereby playing a significant role in shaping the terrestrial carbon sink (Xiang et al., 2022).

RECALCITRANCE AND SOIL CARBON SEQUESTRATION

Recalcitrance refers to resistance of carbon to abiotic and/or biotic degradation, and it is a significant indicator of soil carbon sequestration (Harvey et al., 2012). Biochar carbon typically exhibits greater recalcitrance than the raw biomass carbon due to its enhanced aromatic structure and stability in soil (Windeatt et al., 2014). These materials possess greater chemical durability and are more resistant to biological decomposition. Biochar can be used as carbon storage material due to its relatively high chemical recalcitrance. Recalcitrant organic carbon (ROC) and mineral-associated organic carbon (MAOC) constitute the stable SOC fractions that resist biological degradation (Shi et al., 2023; Enders et al., 2012).

Environmental recalcitrance of biochar can be assessed through recalcitrance index (R_{50}), which is used to evaluate the biochar suitability for carbon sequestration process.

The index reflects how the thermal stability of a biochar measures against that of graphite. (Harvey et al., 2012). Thermogravimetric analysis, thermogravimetry/derivative thermogravimetry (TG/DTG) and differential

scanning calorimetry (DSC), are effective tools for soil organic matter characterization and in this study, the recalcitrance index is determined by thermograms from thermogravimetric analysis.

The recalcitrance index used to determine the relative thermal degradability of biochar can be calculated as (Harvey et al., 2012);

$$R_{50,x} = \frac{T_{50,biochar}}{T_{50,graphite}} \quad (1)$$

where $T_{50,x}$ and $T_{50, graphite}$ are the temperature values corresponding to 50% oxidation/volatilization of biochar and graphite, respectively. T_{50} for graphite is 886 °C. These temperatures can be obtained from TGA thermograms corrected for moisture and ash contents.

In this study, the environmental recalcitrance of biochar derived from cattle manure was evaluated. Cattle manure biochar (BCM) was produced at 300 °C in muffle furnace with a heating rate of 10°C/min and a residence time of 1 h. Figure 2 shows pictures of the raw cattle manure and biochar produces.

In this study, cattle manure sample was put in a crucible and the lid of the crucible was strictly closed in order to avoid air penetration. After thermochemical process, the furnace was cooled to room temperature and biochar sample stored in the desiccator. Thermogravimetric (TG) analysis of the BCM was carried out using a TG analyzer (Seiko SII TG/DTA 7200) under air atmosphere, heated from room temperature to 1000°C, at heating rate of 30°C/min. Lower H/C and higher O/C ratios of biochar compared to raw biomass indicates greater recalcitrance (Ling et al., 2022). Nutrient content of biochar is illustrated in Figure 3.

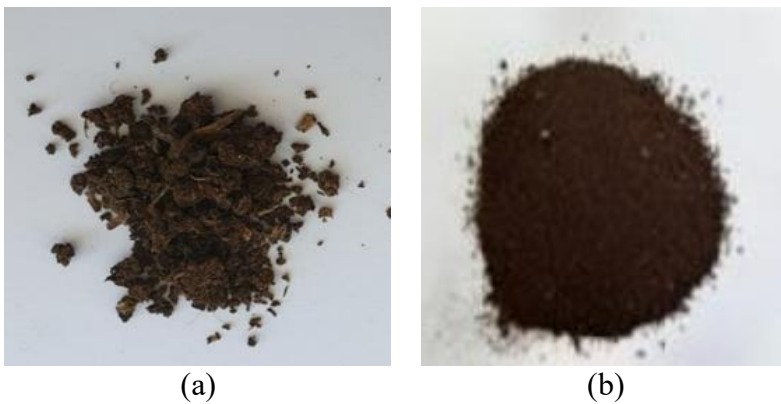


Figure 2. (a) Raw cattle manure (RCM), (b) Cattle manure biochar (BCM).

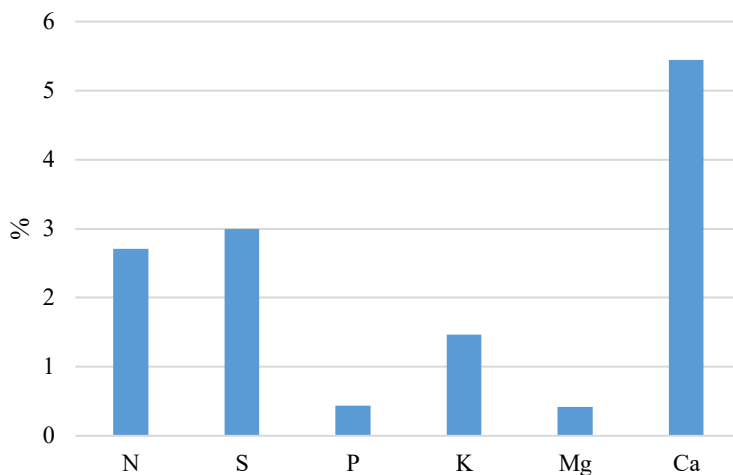


Figure 3. Major nutrient concentrations in biochar.

Original and moisture-ash content corrected TGA thermograms of cattle manure biochar are presented in Figure 4 and Figure 5 , respectively.

According to their recalcitrance value, biochar can be categorized into three groups—Class A, Class B, and Class C—each exhibiting different levels of susceptibility to environmental degradation (Harvey et al., 2012):

- Class A carbons ($R_{50} \geq 0.70$) are the most recalcitrant;
- Class B carbons ($0.50 \leq R_{50} < 0.70$) exhibit minimal degradation,
- Class C carbons ($R_{50} < 0.50$) are more degradable.

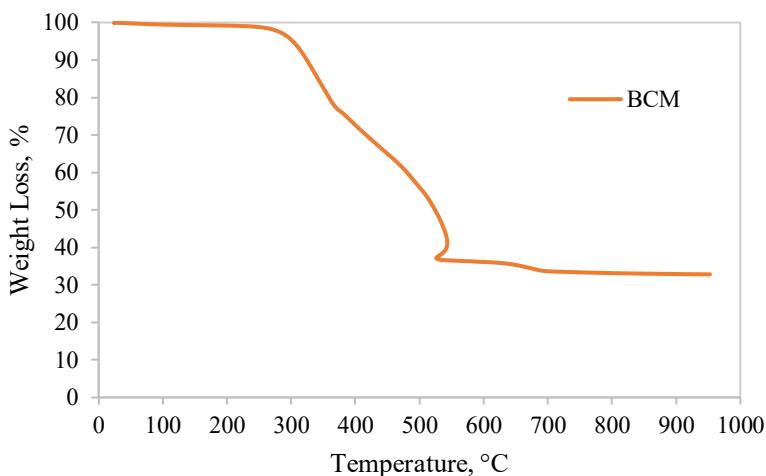


Figure 4. Uncorrected thermogravimetric thermogram of BCM

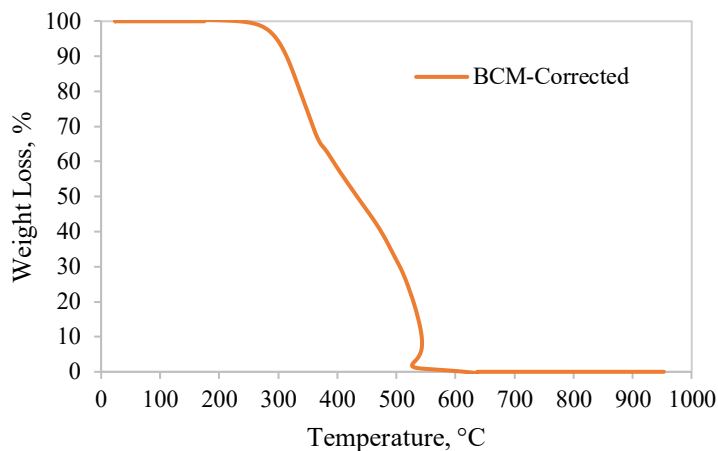


Figure 5. Moisture and ash content-corrected thermogravimetric thermogram of BCM

Using the R_{50} recalcitrance classification system, cattle manure biochar produced at 300 °C is found to have $R_{50} = 0.52$ recalcitrance, which indicates moderate degradability in soil. The recalcitrance of different biochar types was presented in Figure 5.

Biochar with the high lignin contents generally have higher recalcitrance. As can be seen from Figure 5, higher pyrolysis temperatures could also increase the recalcitrance of the biochar (Zhao et al., 2013).

It could be more useful to consider other parameters such as water holding capacity, porosity, cation exchange capacity, etc. together with recalcitrance index, when soil amendment is aimed.

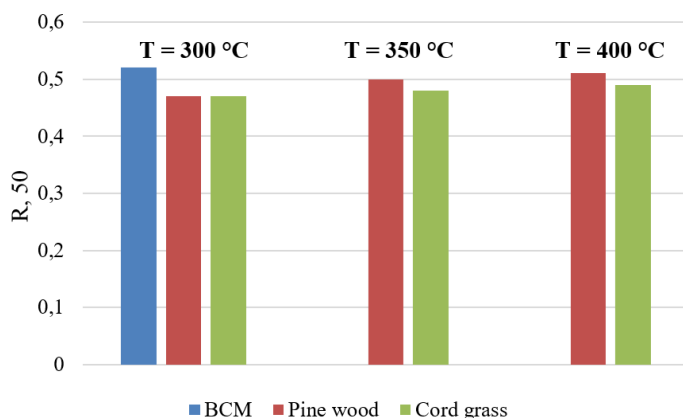


Figure 6. Recalcitrance index values of BCM, pine wood and cord grass. (Kuo et al., 2008; Zimmerman, 2010; Harvey et al., 2012)

CONCLUDING REMARKS

Increasing global temperature that results in climate change and ecosystem deterioration require sustainable solutions to mitigate excess release of atmospheric greenhouse gas emissions. Biochar is a solid material produced from thermochemical degradation of biomass under oxygen-deficit environment. Recalcitrance of biochar promotes its long-term stability in soil. Biochar utilization for soil amendment purposes is a promising solution that enhances soil productivity and soil carbon sequestration.

REFERENCES

- Akyürek, Z. (2021). Carbon Dioxide Sequestration Assessment through Pyrogenic Biomass: Case for Turkey. *Pol. J. Environ. Stud.* 30(3):2467-2475. <https://doi.org/10.15244/pjoes/129688>
- Deng X., Teng F., Fan J.L., Forsell N., Reiner D.M. (2025). Co-deploying biochar and bioenergy with carbon capture and storage improves cost-effectiveness and sustainability of China's carbon neutrality. *One Earth* 8, 101172. <https://doi.org/10.1016/j.oneear.2024.12.008>
- Dubey A., Arora A. (2002). Advancements in carbon capture technologies: a review, *J. Clean. Prod.*, 373, 133932, <https://doi.org/10.1016/j.jclepro.2022.133932>
- Enders A., Hanley K., Whitman T., Joseph S., Lehmann J. (2012). Characterization of biochars to evaluate recalcitrance and agronomic performance. *Bioresource Technology* 114, 644-653. <https://doi.org/10.1016/j.biortech.2012.03.022>
- Global Energy Review 2025, CO2 Emissions. <https://www.iea.org/reports/global-energy-review-2025/co2-emissions>
- Harvey O., Kua L.J., Zimmerman A.R., Louchouart P., Amonette J.E., Herbert B.E. (2012). An Index-Based Approach to Assessing Recalcitrance and Soil Carbon Sequestration Potential of Engineered Black Carbons (Biochars). *Environ. Sci. Technol.* 2012, 46, 1415–1421 <https://doi.org/10.1021/es2040398>
- Hanson E., Nwakile C., Hammed V.O. (2025). Carbon capture, utilization, and storage (CCUS) technologies: Evaluating the effectiveness of advanced CCUS solutions for reducing CO2 emissions. *Results in Surfaces and Interfaces* 18, 100381. <https://doi.org/10.1016/j.rsurfi.2024.100381>
- Fawzy, S., Osman, A.I., Yang, H. et al. Industrial biochar systems for atmospheric carbon removal: a review. *Environ Chem Lett* 19, 3023–3055 (2021). <https://doi.org/10.1007/s10311-021-01210-1>
- IPCC Special Report on Carbon Dioxide Capture and Storage (2005) https://www.ipcc.ch/site/assets/uploads/2018/03/srcs_wholereport-1.pdf
- Kuo, L.-J.; Herbert, B. E.; Louchouart, P. (2008) Can levoglucosan be used to characterize and quantify char/charcoal black carbon in environmental media? *Org. Geochem.* 39, 1466–1478. <https://doi.org/10.1016/j.orggeochem.2008.04.026>
- Lehmann, J. (2007) A handful of carbon. *Nature* 447, 143–144. <https://doi.org/10.1038/447143a>

- Ling L., Luo Y., Jiang B., Jitao L., Meng C., Liao Y., Reis B.J., Ding F., Lu Z., Kuzyakov Y., Xu J. (2022). Biochar induces mineralization of soil recalcitrant components by activation of biochar responsive bacteria groups. *Soil Biology and Biochemistry* 172, 1087. <https://doi.org/10.1016/j.soilbio.2022.108778>
- Mishra G., Danoglidis P.A., Shah S.P., Konsta-Gdoutos M.S., (2023) Carbon capture and storage potential of biochar-enriched cementitious systems. *Cement and Concrete Composites*, 140, 105078. <https://doi.org/10.1016/j.cemconcomp.2023.105078>
- Paris Agreement. <https://unfccc.int/process-and-meetings/the-paris-agreement>
- Saravanan A., Kumar P.S. (2022). Biochar derived carbonaceous material for various environmental applications: systematic review, *Environ. Res.*, 214, 113857, <https://doi.org/10.1016/J.ENVRES.2022.113857>
- Shi J., Yang L., Zhao F., Wu J., Li J., Yu Z., Li A., Shanguan Z., Deng L. (2023) Recalcitrant organic carbon plays a key role in soil carbon sequestration along a long-term vegetation succession on the Loess Plateau. *CATENA* 233, 107528. <https://doi.org/10.1016/j.catena.2023.107528>
- Song X., Li L., Zheng J., Pan G., Zhang X., Zheng J. Hussain Q., Han X. (2012). Sequestration of maize crop straw C in different soils: Role of oxyhydrates in chemical binding and stabilization as recalcitrance. *Chemosphere* 87(6), 649-654. <https://doi.org/10.1016/j.chemosphere.2012.01.038>
- Xiang H., Luo H., Zhang L., Hou E., Li J., Zhu Q., Wen D. (2022). Forest succession accelerates soil carbon accumulation by increasing recalcitrant carbon stock in subtropical forest topsoils. *Catena*, 212, 106030. <https://doi.org/10.1016/j.catena.2022.106030>
- Windeatt J.H., Ross A. B., Williams P.T., Forster P.M., Nahil M.A., Singh S. (2014). Characteristics of biochars from crop residues: Potential for carbon sequestration and soil amendment. *Journal of Environmental Management*, 146, 189-197. <https://doi.org/10.1016/j.jenvman.2014.08.003>
- Zhao L., Cao X., Masek O., Zimmerman A. (2013). Heterogeneity of biochar properties as a function of feedstock sources and production temperatures. *J. Hazard. Mater.*, 256, 1-9. <https://doi.org/10.1016/j.jhazmat.2013.04.015>
- Zimmerman, A. R. (2010) Abiotic and microbial oxidation of laboratory-produced black carbon (biochar). *Environ. Sci. Technol.* 44, 1295–1301. <https://doi.org/10.1021/es903140c>

CHAPTER 15

Buldan Fabric Defects: Description, Analysis and Detection Using Image Processing Based Intelligent Models

**Erdem Dilmen¹ & Emine Utkun Aktener² &
Selami Beyhan³ & Batuhan Bilgi⁴**

¹ Asst. Prof. Dr., Department of Mechatronics Engineering, Pamukkale University, Denizli, Turkey, ORCID: 0000-0003-3092-0927

² Assoc. Prof. Dr., Buldan Vocational School, Pamukkale University, Buldan, Turkey, ORCID: 0000-0002-6550-4168

³ Prof. Dr., Department of Artificial Intelligence and Data Engineering, Yıldız Technical University, İstanbul, Turkey, ORCID: 0000-0002-9581-2794

⁴ Department of Electrical and Electronics Engineering, Pamukkale University, Denizli, Turkey

1 Introduction

Buldan city is an important local weaving center in west Anatolia, in the city of Denizli, where significant textile production of Turkey is done. According to the inventory research conducted by (Ertugrul and Utkun, 2009), it was confirmed that in Buldan city, there were 45 units of hand looms, 127 units of semi-automatic looms and 102 units of automatic looms operational by 2009. Nowadays, the most of the products are manufactured in workshop-type small enterprises and semi-automatic weaving looms in the country. Note that history of Buldan fabric dates back to centuries before (Ozturk and Coruh, 2013). Buldan has been an important textile production center in Anatolia for so long, which makes Buldan fabric worth investigating.

Actions or environmental facts may cause defects (or faults) on the manufactured fabric. To get a high quality production, detection of the defects is crucial. The faulty products should not be released to the market and they should be detected on time. Such effort can also give manufacturers the ability to analyze the reasons for defects and improve the production process. We can simply introduce some of the studies relating to detection of fabric defects using different techniques. In (Wu et al, 2017), dictionary learning is utilized in fabric defect detection. Least squares method is applied to learn the dictionary, where discrete cosine transform is performed to obtain the initiation dictionary. In (Lv and Long, 2015), defects in the color-textured fabric are classified employing an Elman neural network (Elman, 1990). However, the authors modified it by adding an extra context layer which they call 'memory layer'. Processed images of the fabric are obtained using color ring projection and they are fed as input to the modified Elman neural network. In the detection phase of fabric defects, they need some morphology processing parameters and optimal values are found using genetic algorithm (Holland, 1984). In (Yuen et al, 2009), automatic detection of fabric stitching defects is investigated. The images of the fabric are segmented into three classes of stitching defects using a segmentation window technique. Using those segmented images, nine characteristic variables are obtained and they are fed as input to a standard feedforward artificial neural network (ANN) to detect and distinguish between three different types of defects. In (Bu et al, 2010), power spectral density analysis of the fabric images is performed and features are extracted such that, those features can represent well the fabric data in a suitable form which can be differentiated as normal and defective types. Using support vector data description method (SVDD) (Scholkopf et al, 2001; Tax and Duin, 2004) together with the extracted features, the detection of the fabric defects can be achieved. Note that in SVDD, the defective texture can be detected by outlier rejection, which is quite different than the conventional binary classification that needs labeled training data as 'normal' and 'defective' in a supervised learning

framework. In (Wong et al , 2009), stitching defect detection and classification is achieved using wavelet transform and backpropagation (BP) neural network. In (Tabassian et al, 2011), classification of circular knitted fabric defect is investigated using uncertain information of labeling and their combination with the Dempster–Shafer theory of evidence.

Considering the literature studies mentioned, it is worth noting that approaching the problem of fabric defect detection by considering the machine learning techniques such as ANN, support vector machine (SVM) (Vapnik, 1998), SVDD, etc., can yield intelligent models capable of capturing the fabric properties effectively. This fact makes use of such models an attractive alternative for the researchers and motivates us to use new intelligent models. It should be indicated that, ANN and SVM are widely used by the researchers and they are quite popular in the fabric defect detection community. However, extreme learning machine (ELM) (Huang et al, 2004; Huang et al, 2006) has not been comprehensively investigated in this topic yet. Thus, in this study, we propose the use of ELM in defect detection in Buldan fabric and we compare its performance with traditional models such as ANN and SVM. Based on that, let us briefly give a description of the reason why we involve ELM in this study. It is slightly different than the standard feedforward ANN. Its hidden layer weights are randomly generated and based on those weights, output layer weights are obtained via least squares estimation. Thus, ELM is derived originally from the standard ANN. The main idea is to use random learning approach for tuning the hidden layer weights while avoiding the local minima. Its performance on fabric defect detection task is definitely worth investigating, especially comparing it to its ancestor, standard ANN. Note that, details of all three models employed in this study (ANN, ELM and SVM) will be given in Sect. 3.

In this chapter, first, types of the defects in Buldan fabric and possible causes are explained. Then, intelligent models ANN, ELM and SVM are employed for detection of the defects in Buldan fabric. From a weaving loom, 35 defective and 20 normal samples of Buldan fabric are imaged. First fabric images are divided into subimages then the eccentricity, extent and orientation features of subimages are extracted from the fabric images and these features are fed as input to the intelligent models. From 55 samples of the fabric, 33 samples are used for training and 22 samples are used for testing the intelligent models. In order to compare the models fairly, the training and testing samples are randomly selected in 50 different combinations where the performance of each defect detection experiment is recorded. Then, obtained application results are discussed over the mean values. To the best of authors' knowledge, the analysis of defects in Buldan fabric, and the detection of those defects has not been investigated yet in the literature.

Organization of the paper is as follows. In Sect. 2, defects and possible causes to those defects in Buldan fabric are investigated briefly. Intelligent models adopted for fabric defect detection are introduced in Sect. 3. Application results and discussion are given in Sect. 4. Finally, Sect. 5 concludes the paper.

2 Buldan Fabric and Defect Types

The woven items in Buldan are divided into two: parts production and meter production. The parts items are towel, sheet, handkerchief, loin cloth, robe, halves and table-top. Meter items are Buldan cloth, interlining, tent fabric, canvas and flannel (Ertugrul and Utkun, 2009; Utkun and Kirtay, 2009; Utkun et al, 2009). The characteristics of Buldan cloth can be given as follows. It is made from 100% cotton yarn, it is a fabric that has a plain weave texture produced by shuttle weaving method and the weft yarn has a higher twist value than the warp yarn. It has a curly surface as it is woven with twisting yarn, it is flexible and tight-fitting. The airspace between the twists keeps the body cool, its sweat absorbency is sanitary because of its cotton quality. Buldan cloth has the quality of being a very unique type with its compliance with the market conditions, its traditional value and technical properties. After Buldan cloth is woven, it is not processed with any chemical substance and it is just subjected to desizing process (Utkun and Ondogan, 2012; Ozdemir and Utkun, 2014; Utkun and Gundogan, 2014). The technical properties of Buldan cloth used in this study is given in Table 1.

Table 1. The technical specifications of Buldan cloth.

	Number (Ne)	Density (wrap/cm)	Twist coefficient 1 (αe)	Twist coefficient 2 (αe)
Warp yarn	80/2	16	3.7 (Z way)	3.1 (S way)
Weft yarn	40/1	11	3.7 (Z way)	-

We can briefly present the defect types in Buldan fabric as follows.

- **Defect 1. Foreign material:** It is a kind of defect that is seen when the dust or foreign materials are spilled from the workbench into the fabric. The working environment of the weaving loom must be clean. Defect 1 is shown in Fig. 1(a).

- **Defect 2. Weft accumulation-1:** When the yarn is consumed for weft, the empty shuttle is replaced with a full one. At this time, such a defect is caused by the jamming of excess yarns. When the excess yarns in the gap are properly

drawn, fabric will take on its normal appearance and will be defect-free. Defect 2 is shown in Fig. 1(b).

- **Defect 3. Weft accumulation-2:** This type of defect can be seen when the weft yarn comes from the shuttle improperly and causes accumulation of the yarns. It may also be encountered when the shuttle runs out of the yarn and the weft yarn is left half-way. In this case, the new weft yarn is fed into the remaining one. It is not possible to undo this defect like weft accumulation-1. Defect 3 is shown in Fig. 1(c).

- **Defect 4. Weft opening:** If the stroke of the comb is not properly adjusted, while the fabric is woven, feeding skewed weft yarns can cause this defect. Defect 4 is shown in Fig. 1(d).

- **Defect 5. Frequent and infrequent weft :** When the current weft is laid, the new weft is pushed by the comb. In case of this defect, the beaten amount of comb is not balanced, sometimes frequent and sometimes infrequent. Defect 5 is shown in Fig. 1(e).

- **Defect 6. Warp broken :** One or more of the warp yarns of the fabric is broken. Defect 6 is shown in Fig. 1(f).



(a) Defect 1. Foreign material.



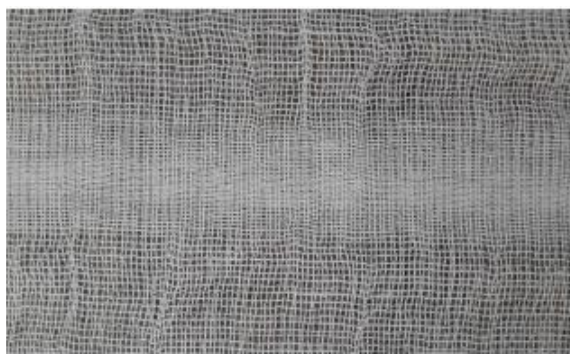
(b) Defect 2. Weft accumulation-1.



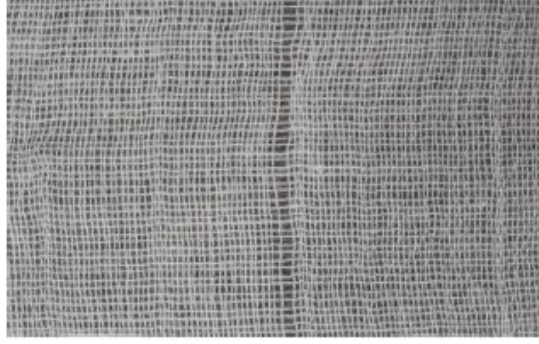
c) Defect 3. Weft accumulation-2.



(d) Defect 4. Weft opening.



e) Defect 5. Frequent and infrequent weft.



(f) Defect 6. Warp broken.

Figure 1. The defect types in Buldan fabric.

3 Intelligent Models for Detection of Defects

This section introduces the artificial intelligent models employed in this study, which have been used in many scientific research topics and industrial applications. Consider that we have fabric samples $\{\mathbf{x}_i, y_i\}_{i=1}^N$ obeying an unknown input-output relation which represents a mathematical expression for the fabric properties. Let it be in the form of

$$y_i = f(\mathbf{x}_i) \quad (1)$$

where $y_i \in -1, +1$ is the output of the unknown function $f(\cdot): \mathcal{R}^d \rightarrow \mathcal{R}$. Here, y_i denotes a binary class label indicating whether the i_{th} fabric sample is defective or not. Note that, -1 corresponds to a defective sample and +1 corresponds to a normal one. Also, $\mathbf{x}_i \in \mathcal{R}^d$ is the input vector which contains the extracted features from image of the i_{th} fabric sample.

Since we do not have a mathematical model for the defect detection function $f(\cdot)$, we employ artificial intelligent models instead. In fact, those models are defined as nonlinear regressors which approximate the input-output relation $f(\cdot)$ by learning the fabric properties in the data. To do this, fabric image data are divided into training and testing parts. Using the training data, intelligent models are trained in such a way that a priori cost function of the class label prediction errors is minimized. The prediction error for a sample is the difference between the true class label of that sample and the predicted output by the model. Note that, prediction by a model is done via the approximated input-output relation which is written implicitly as follows.

$$\begin{aligned} \hat{y}_i &= \hat{f}(\mathbf{x}_i) \\ \hat{y}_{i_l} &= \text{sat}(\hat{y}_i) \end{aligned} \quad (2)$$

$$e_i = y_i - \hat{y}_i \quad (3)$$

where, $\hat{y}_i \in \mathcal{R}$ is the predicted output by the model and $\hat{f}(\cdot)$ denotes the approximated input-output relation. In (3), e_i is the prediction error for the i_{th} sample. $\hat{y}_{i_t} \in \{-1, +1\}$ denotes the actual class label prediction by the model. It returns a binary class label passing the predicted output value $\hat{y}_i$ through the saturation function $sat(\cdot)$ where

$$\begin{aligned} \hat{y}_{i_t} &= sat(\hat{y}_i) \\ sat(\hat{y}_i) &= \begin{cases} +1, & \hat{y}_i \geq 0 \\ -1, & \hat{y}_i < 0 \end{cases} \end{aligned} \quad (4)$$

Note that, training a model means finding the optimum weights for that model which minimizes a priori cost function of the errors. Once the model is trained using the training data, its performance is evaluated on the test data. Note that, test data belong to the samples that the artificial model did not see in the training phase. Class label corresponding to the test samples are predicted using the trained model with optimum weights. Then, the model is assessed with a performance metric, e.g., classification accuracy. Lower number of misclassified samples leads to a higher classification accuracy, which is the desired case in machine learning applications. In other words, a model is expected to make a correct prediction for a fabric image which the model has not seen in the training phase. Such a model will have a good generalization ability and can be a reliable candidate as a defect detection system.

In this study, the approximated input-output relation $\hat{f}(\cdot)$ is obtained by adopting ANN, ELM and SVM models for defect detection. The models are detailed in the following subsections.

3.1 Artificial Neural Network

Artificial neural network (ANN) is one of the well-known function approximators that are structured by artificial neurons. The weights of neurons store information about the interaction between input-output. Using enough number of neurons, a black-box model (Ljung, 2001) can be designed for identification, classification, control etc. Suppose that we have a set of collected data from the unknown system given by (1). Mathematically, the ANN model output in response to an input vector x_i corresponding to the i_{th} sample can be written as

$$\hat{y}_i = \mathbf{w}_o^T \boldsymbol{\phi}(\mathbf{w}_h x_i + \mathbf{b}_h) \quad (5)$$

where $\mathbf{w}_o \in \mathcal{R}^{1 \times s}$ and $\mathbf{w}_h \in \mathcal{R}^{s \times d}$ are the output layer and hidden layer weight matrices, respectively. Also, $\mathbf{b}_h \in \mathcal{R}^s$ is the vector of hidden layer bias values. These bias values prevent the neurons from sleeping. $\boldsymbol{\phi}(\cdot): \mathcal{R}^s \rightarrow \mathcal{R}^s$ denotes the activation function which fires the neurons. In general, the activation

function can be chosen among the types: sigmoidal, tangent hyperbolic, Gaussian, trigonometric, orthogonal functions, etc. A tangent hyperbolic,

$$\phi(a) = \frac{e^a - e^{-a}}{e^a + e^{-a}} \quad (6)$$

is used as the activation function in this paper. A one hidden layer, multi-input multi-output artificial neural network model is shown in Fig. 2 as a general representation. However, in the applications, we employ a multi-input 'single'-output type of ANN since the studied task in this paper is binary classification. Levenberg-Marquardt (LM) algorithm is adopted for training the ANN model, thus finding the optimum model weights.

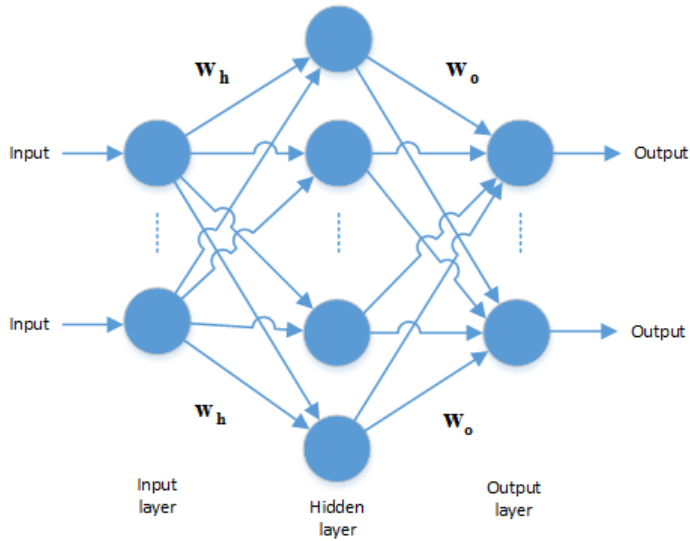


Figure 2. ANN model with one hidden layer.

3.1.1 Levenberg-Marquardt Optimization

The conventional LM optimization actually performs nonlinear least squares estimation. It is an iterative gradient based method that is one of the fastest optimization methods (Scales, 1985). Let us have a quadratic cost function given by

$$E = \frac{1}{2} \sum_{i=1}^N e_i^2 \quad (7)$$

where $E: \mathfrak{R}^N \rightarrow \mathfrak{R}$ denotes the sum of squared errors (SSE) of the data samples obeying the unknown input-output relation in (1) and e_i is given by (3). Let $\boldsymbol{\theta} \in \mathfrak{R}^n$ denote the vector holding parameters of the function approximator model $\hat{f}(\cdot)$ in (2). Hence, predicted output by the model can be rewritten as follows.

$$\hat{y}_i = \hat{f}(x_i, \boldsymbol{\theta}) \quad (8)$$

When we consider (3) and (8), we see that the error term e_i involves $\hat{f}(x_i, \boldsymbol{\theta})$. We already know from (7) that, E is a function of the error terms $e_i, i = 1, \dots, N$. Therefore, E is also a function of the parameters $\boldsymbol{\theta}$ ($E \rightarrow E(\boldsymbol{\theta})$). What LM method actually does is to find the optimum values for $\boldsymbol{\theta}$ iteratively that will lead us to a local minimum of the function $E(\boldsymbol{\theta})$. Here, although the cost function in (7) is a quadratic one, each error term is nonlinear in the parameters $\boldsymbol{\theta}$ since each involves a nonlinear function of the parameters such as $\hat{f}(x_i, \boldsymbol{\theta})$ in (8). This is the reason why cost function $E(\boldsymbol{\theta})$ is not in the form of a function which has only a unique global minimum. Conversely, $E(\boldsymbol{\theta})$ has several local minima. LM method converges to one of those local minima which is the nearest one to the point defined by the initial parameter vector in n -dimensional space. To do this, it updates the parameters in $\boldsymbol{\theta}$ at each iteration by adding a correction term. Let E_k denote the cost function of the errors at the k_{th} iteration. As LM runs through iterations, cost function of the errors tends to decrease ($E_{k+1} < E_k$). When the decrease in the cost function between two sequential iterations gets smaller than a predefined threshold, LM reaches the nearest local minimum of E . This is the point where optimum values for the parameter vector $\boldsymbol{\theta}$ is found. Note that, the local minimum that LM will reach strongly depends on the initial parameter values. This is the main reason for 'getting stuck at the local minimum' problem encountered in gradient based optimization methods.

Optimization by LM is based on the linearized form of E_k about $\boldsymbol{\theta}_k$, where $\boldsymbol{\theta}_k$ denotes parameters of the model $\hat{f}(\cdot, \cdot)$ at the k_{th} iteration. Let $\Delta\boldsymbol{\theta}_k \in \mathfrak{R}^n$ be the correction term to the parameters as $\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k + \Delta\boldsymbol{\theta}_k$. Then,

$$\Delta\boldsymbol{\theta}_k = -(\nabla^2_{E_k})^{-1} \nabla_{E_k} \quad (9)$$

gives us the optimal correction term for the parameter update. In (9), $\nabla_{E_k} \in \mathfrak{R}^n$ is the gradient vector and $\nabla^2_{E_k} \in \mathfrak{R}^{n \times n}$ is the Hessian matrix of the cost function E_k wrt to the parameters in $\boldsymbol{\theta}_k$.

$$\nabla_{E_k} = \frac{\partial E_k}{\partial \boldsymbol{\theta}_k}, \quad \nabla^2_{E_k} = \frac{\partial^2 E_k}{\partial \boldsymbol{\theta}_k^2} \quad (10)$$

Considering (3), (7) and (8), gradient vector can be formulated as follows.

$$\begin{aligned}
\nabla_{E_k} &= \frac{\partial E_k}{\partial \theta_k} \\
&= \sum_{i=1}^N e_{i_k} \frac{\partial e_{i_k}}{\partial \theta_k} \\
\frac{\partial e_{i_k}}{\partial \theta_k} &= \frac{\partial (y_i - \hat{y}_{i_k})}{\partial \theta_k} = -\frac{\partial \hat{y}_{i_k}}{\partial \theta_k} = -\frac{\partial \hat{f}(x_i, \theta_k)}{\partial \theta_k}
\end{aligned} \tag{11}$$

In a more compact form, gradient vector can be rewritten as

$$\begin{aligned}
\nabla_{E_k} &= \mathbf{J}_k^T \mathbf{e}_k \\
\mathbf{e}_k &= [e_{1_k}, e_{2_k}, \dots, e_{N_k}]^T \\
\mathbf{J}_k &= - \begin{bmatrix} \frac{\partial \hat{f}(x_1, \theta_k)}{\partial \theta_{1_k}} & \dots & \frac{\partial \hat{f}(x_1, \theta_k)}{\partial \theta_{n_k}} \\ \vdots & \ddots & \vdots \\ \frac{\partial \hat{f}(x_N, \theta_k)}{\partial \theta_{1_k}} & \dots & \frac{\partial \hat{f}(x_N, \theta_k)}{\partial \theta_{n_k}} \end{bmatrix}
\end{aligned} \tag{12}$$

where $\mathbf{J}_k \in \mathbb{R}^{N \times n}$ is the Jacobian matrix and $[\theta_{1_k} \dots \theta_{n_k}]^T$ denote the individual parameters in the vector θ_k . Now, let us obtain the Hessian matrix considering (10)-(11).

$$\begin{aligned}
\nabla^2_{E_k} &= \frac{\partial^2 E_k}{\partial \theta_k^2} \\
&= \frac{\partial}{\partial \theta_k} \left(\sum_{i=1}^N e_{i_k} \frac{\partial e_{i_k}}{\partial \theta_k} \right)
\end{aligned} \tag{13}$$

which is equal to

$$\nabla^2_{E_k} = \sum_{i=1}^N \frac{\partial e_{i_k}}{\partial \theta_k} \left(\frac{\partial e_{i_k}}{\partial \theta_k} \right)^T + \sum_{i=1}^N e_{i_k} \frac{\partial^2 e_{i_k}}{\partial \theta_k^2} \tag{14}$$

It is experimentally observed that, second order partial derivative term $\frac{\partial^2 e_{i_k}}{\partial \theta_k^2}$ in (14) takes very small values and increases the computational load of the optimization. Thus, it can be ignored. The Hessian expression can be derived approximately as follows.

$$\nabla^2_{E_k} \approx \sum_{i=1}^N \frac{\partial e_{i_k}}{\partial \theta_k} \left(\frac{\partial e_{i_k}}{\partial \theta_k} \right)^T \tag{15}$$

which can be rewritten in a more compact form considering (11)-(12) as follows.

$$\nabla^2_{E_k} \approx \mathbf{J}_k^T \mathbf{J}_k \tag{16}$$

Combining (12) and (16), the correction term in (9) can be rewritten as

$$\Delta \boldsymbol{\theta}_k = -(\mathbf{J}_k^T \mathbf{J}_k)^{-1} \mathbf{J}_k^T \mathbf{e}_k \quad (17)$$

Note that, in order to maintain minimization of the cost function, the Hessian matrix must be positive definite. However, sometimes Hessian may not be positive definite due to negative eigenvalues. In such case, positive numbers are added gradually to the diagonal elements until the positive definiteness is achieved.

$$\Delta \boldsymbol{\theta}_k = -(\mathbf{J}_k^T \mathbf{J}_k + \mu \mathbf{I}_{n \times n})^{-1} \mathbf{J}_k^T \mathbf{e}_k \quad (18)$$

In (18), μ is the damping coefficient and $\mathbf{I}_{n \times n}$ is the identity matrix. μ does not only prevent the Hessian from being negative definite but also it provides the algorithm with a variable convergence speed. This is achieved by adaptively changing the value of μ between iterations. This is the main idea proposed by LM algorithm which makes it very usefull in many learning tasks where batch training data is used.

While training the ANN model using LM algorithm, the function approximator model $\hat{f}(\mathbf{x}_i, \boldsymbol{\theta})$ in (8) will be substituted by ANN model output equation in (5). Also, parameters of ANN which are contained in $\mathbf{w}_o$, $\mathbf{w}_h$ and $\mathbf{b}_h$ will be transferred into the parameter vector $\boldsymbol{\theta}$. Note that, this approach can be extended to use of any intelligent model which processes batch training data. The crucial part here is to compute the partial derivative of the model output wrt the model parameters correctly to obtain the Jacobian matrix $\mathbf{J}$ in (12). Pseudoce of LM algorithm is given in Algorithm 1 in Appendix A1.

3.2 Extreme Learning Machine

An extreme learning machine (ELM) is constructed based on the one hidden layer feedforward neural network. While training this model, the hidden layer weighting and biasing parameters are randomly generated (Huang et al, 2004; Huang et al, 2006) and based on that, optimal output weighting parameters are obtained using least squares estimation (LSE). This type of training provides high degree of input-output modelling which is based on the random learning theory of creatures (Huang, 2014; Beyhan and Kavaklioglu, 2015).

More clearly, ELM model is trained combining random and deterministic learning approaches while ANN model adopts only deterministic learning as we saw in the previous subsection (learning based on LM algorithm). Random learning of both the hidden layer weights and the number of neurons bring some advantages compared to the conventional ANN model. First, random assignment of the hidden layer weights overcomes the challenge of local minima up to some degree. Note that, this challenge is frequently faced when using gradient descent

optimization. Second, computational time required for model training in the conventional ANN case is decreased. This is due to the fact that, most of the computational time required for training is spent by mathematical operations performed by the deterministic training method. When some portion of the computational time belonging to the deterministic training part is devoted to random learning, total computational time for training the model decreases.

Let us have a set of data samples which satisfy the unknown input-output function given by (1). Output prediction by the ELM model for the i_{th} data sample is computed similar to the ANN model output as follows.

$$\hat{y}_i = \boldsymbol{\theta}^T \boldsymbol{\phi}(\mathbf{w}_h x_i + \mathbf{b}_h) \quad (19)$$

In (19), $\mathbf{x}_i$ is the i_{th} data sample input vector and $\hat{y}_i$ is the model output in response to x_i . Note that, $\boldsymbol{\phi}(\cdot)$ denotes the activation function which is tangent hyperbolic, while $\mathbf{w}_h$ and $\mathbf{b}_h$ stand for the hidden layer weights and biases. Due to the hybrid training approach, $\mathbf{w}_h$ and $\mathbf{b}_h$ are generated randomly and output layer weights contained in the vector $\boldsymbol{\theta} \in \mathbb{R}^{1 \times s}$ are obtained via LSE. Note that, ELM and ANN are very similar in structure. Here we call the output layer weights as $\boldsymbol{\theta}$ where they were referred to as $\mathbf{w}_o$ in ANN model.

Now, we will give the detail of how the parameters in $\boldsymbol{\theta}$ is obtained. Let us denote all the output samples collected from the system given by (1) in a vector, call it the observed output vector: $\mathbf{y} \in \mathbb{R}^N$. Using the given input-output data and randomly generated weights of the hidden layer, an expression relating the ELM model to the observed outputs can be written in a compact form as follows.

$$\mathbf{F}\boldsymbol{\theta} = \mathbf{y} \quad (20)$$

where $\mathbf{F}_{N \times s}$ is the regressor matrix given by

$$\mathbf{F} = \begin{bmatrix} \phi^T(\mathbf{w}_h \mathbf{x}_1 + \mathbf{b}_h) \\ \vdots \\ \phi^T(\mathbf{w}_h \mathbf{x}_N + \mathbf{b}_h) \end{bmatrix}_{N \times s} \quad (21)$$

Let the vector of prediction errors be $\mathbf{e} \in \mathbb{R}^N$, which is given by

$$\mathbf{e} = \mathbf{F}\boldsymbol{\theta} - \mathbf{y} \quad (22)$$

Then, the optimization problem is based on the idea of minimizing the quadratic cost function

$$\min_{\boldsymbol{\theta}} E = \frac{1}{2} \mathbf{e}^T \mathbf{e} \quad (23)$$

One may notice that, (23) is a representation in vector notation of the equivalent cost function given by (7). We know that, if a cost function is quadratic and the error term is linear in the parameters, then there is one minimum point which is global. At that global minimum, gradient will be zero. Equation (23) can be rewritten as

$$\min_{\theta} E = \frac{1}{2} (\mathbf{F}\theta - \mathbf{y})^T (\mathbf{F}\theta - \mathbf{y}) \quad (24)$$

When we expand the terms and solve for the parameters which satisfy zero-gradient condition, we obtain the optimal parameter vector as

$$\hat{\theta} = \mathbf{F}^{\dagger} \mathbf{y} \quad (25)$$

where

$$\mathbf{F}^{\dagger} = (\mathbf{F}^T \mathbf{F})^{-1} \mathbf{F}^T \quad (26)$$

is the Moore-Penrose inverse of $\mathbf{F}$.

3.3 Support Vector Machine

SVM is originally based on a linear binary classification problem (Vapnik, 1998). It does not only minimize the prediction error but also sets an upper bound on the generalization error. Thus, it provides both empirical and structural risk minimization. Assume that, we have a set of collected data samples from an unknown system given by (1), which is divided into training and testing parts. By the structural risk minimization principle, SVM avoids memorizing the training data and sets a balance between the training and testing performances. This is provided by the regularization term, which can make us interpret SVM technique as an improved ridge regression (with observed outputs in $\{-1, +1\}$). However, SVM method finds its foundations on the comprehensive statistical learning theory. In SVM methodology, a quadratic cost function with inequality constraints is minimized. As we will see below, it is in form of a function which has a unique global minimum and no other local minima. Thus, optimization of the model gives globally optimum parameters, which yields a promising generalization ability by the aid of structural risk minimization mentioned above. Note that, in this paper, we will cover the nonlinear classifier with slack variables, aka soft margin classifier in the literature. It has a simple mathematical expression as

$$\hat{y}_i = \mathbf{w}^T \phi(\mathbf{x}_i) + b \quad (27)$$

where $\mathbf{w}$ denotes the SVM weights and $b \in \mathfrak{R}$ is the bias while $\hat{y}_i \in \{-1, +1\}$ shows the model output in response to the input vector $\mathbf{x}_i \in \mathfrak{R}^d$ corresponding to the i_{th} sample. $\phi(\cdot)$ denotes a nonlinear mapping from the input space to a higher dimensional feature space. Due to this mapping, nonlinear data in the input space can be linearly separated in a higher dimensional space. It is desired to find an optimal hyperplane that minimizes the cost function below. Let us call the region between $\hat{y} = [-1, +1]$ as margin. It is desired that, i) margin will be maximized (structural risk minimization) and ii) prediction error will be minimized (empirical risk minimization). Since soft margin classifier is investigated, errors are allowed but penalized. This is called the use of slack variables.

$$\begin{aligned} \min_{\mathbf{w}, b} E &= \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{i=1}^N \xi_i \\ \text{s. t. } \text{cnst.} &: \\ &y_i (\mathbf{w}^T \phi(\mathbf{x}_i) + b) \geq 1 - \xi_i, \quad i = 1, \dots, N \\ &\xi_i \geq 0, \quad i = 1, \dots, N \end{aligned} \quad (28)$$

where ξ_i are the so-called slack variables. In (28), which gives us the primal formulation, C is the regularization term and determines the trade off between prediction error performance and generalization capability. When we construct the Lagrangian equation, we obtain the dual formulation.

$$\begin{aligned} \min_{\alpha} E &= \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \alpha_i \alpha_j y_i y_j \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j) - \sum_{i=1}^N \alpha_i \\ \text{s. t. } \text{cnst.} &: \\ &\sum_{i=1}^N \alpha_i y_i = 0 \\ &0 \leq \alpha_i \leq C, \quad i = 1, \dots, N \end{aligned} \quad (29)$$

where α_i are the Lagrange (dual) variables. Note that, nonlinear classification originally requires the explicit knowledge of the mapping function $\phi(\cdot)$. However, practically it is a very difficult task and mostly avoidable. This is solved by substituting the inner product of the input vectors by a kernel function which satisfies the Mercer conditions, denoted by $K(\cdot, \cdot)$ as follows. This is called the kernel trick.

$$K(\mathbf{x}_i, \mathbf{x}_j) = y_j \phi(\mathbf{x}_i)^T \phi(\mathbf{x}_j) \quad (30)$$

There are several kernel functions in the literature such as linear, polynomial, tangent hyperbolic, trigonometric, orthogonal functions, etc. In this paper, we use Gaussian kernel

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-0.5 \frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{\sigma^2}\right) \quad (31)$$

where $\|\cdot\|$ denotes the L_2 norm of a vector and σ is the width parameter.

Advantage of using the dual formulation with an adequate kernel function is; it avoids the necessity of explicit formulation of the nonlinear mapping ϕ . Optimal parameters (dual variables α_i and bias b) can be solved using quadratic programming techniques (QP). Among the dual variables, which satisfy $\alpha_i > 0$ are called the support vectors and contribute to the model output given below.

$$\hat{y}(\mathbf{x}) = \sum_{i \in SV} \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b \quad (32)$$

In (32), SV denotes the set of support vectors. $\hat{y}(\mathbf{x})$ is the response of SVM to an arbitrary test input $\mathbf{x}$. Note that $\hat{y}$ is passed through the saturation function in (4) to obtain a binary output.

4 Fabric Defect Detection Results

This section presents design steps of the intelligent models and application results. For the defect detection problem, output of the defective fabric samples is considered as -1 while it is $+1$ for the normal samples. Then, predicted class labels by the intelligent models are obtained by passing model outputs through the saturation function given by (4). In other words, continuous outputs of the models are transformed into binary outputs in $\{-1, +1\}$. Parameters of the intelligent models are optimized using the training data set while testing data samples are used to test them and assess generalization performance.

4.1 Design Aspects for Intelligent Models

Artificial neural network is constructed with a one hidden layer and 5 neurons. The number of neurons are determined via grid search of an interval $[2, 20]$. ANN model parameters are optimized employing Levenberg-Marquardt algorithm using training input-output data samples.

Extreme learning machine is constructed based on a random selection of hidden layer neuron number, which is searched between $[2, 30]$ and optimal number of neurons is found as 9. For each different neuron number, 100 different sets of hidden layer parameters are assigned randomly. Using a regressor matrix constructed by these random assignments, output parameters are determined via least squares estimation.

For the support vector machine model, a grid search was performed to find the optimum hyperparameters for the model. It was observed that $C = 100$ and $\sigma = 0.5$ provide the best model for fabric defect detection task. C was searched in $[0,$

1000] and σ was searched in [0, 5]. Note that, hyperparameters associated with the models are determined experimentally and found amenable for this study.

4.2 Feature Selection of Images

The feature values of defective and normal Buldan fabric images are first extracted and recorded for use as training and test inputs of the intelligent models since the pixel values of a complete image can not be utilized as the inputs of the intelligent models. A feature vector corresponding to an image has totally 85 values such that those features are obtained first dividing the image into subimages. The eccentricity, extent, orientation properties are located in the input feature vector. Those features are explained detailly in (Gaurav and Ramesh, 2012). Feature vectors of two sample images are shown in Fig. 3.

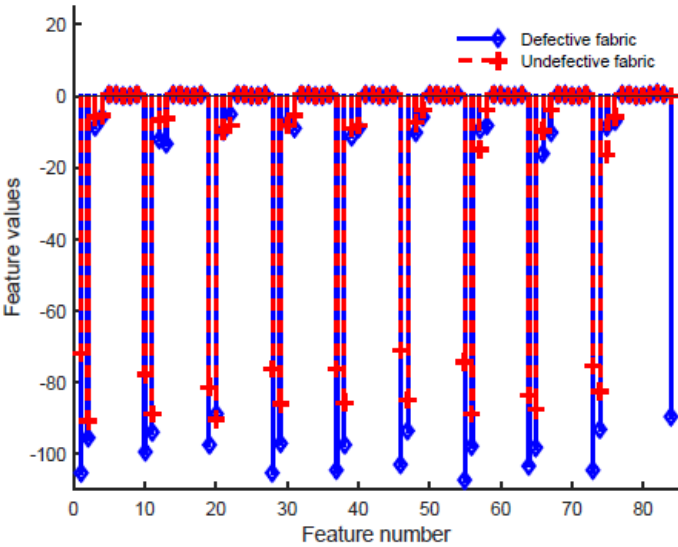


Figure 3. Feature map extracted from two sample images.

4.3 Application Results

After the hyperparameters of each model are determined optimally, the application results are statistically found. It means that, defect detection experiments are performed using 50 different random selections of the training and testing data sets. For each selection, 33 samples are used for training and 22 for testing among all data samples. These randomly selected data sets are used same for three models. Then, obtained results are listed considering three performance criteria which are classification accuracy, mean and variance. For a single experiment, 'accuracy' is calculated as the ratio of the number of truly detected outputs to the number of data samples in the set (training and testing separately). Based on that, overall accuracy is the mean of each single

experiment’s accuracy result obtained from 50 different experiments. Higher accuracy means better performance. The second performance criterion ‘mean’ is actually mean of the number of misclassified samples using 50 different data sets. Similarly, the last performance criterion ‘variance’ is the variance of that quantity. Obviously, larger values for mean and variance are not desired.

Table 2 gives the performance results of intelligent models for training phase. It is seen that the ANN model provides best performance result for accuracy, mean and variance. It is followed by ELM and SVM models respectively. ANN model yields %99.4 accuracy so that it learns very well the input-output relation which represents the fabric properties of the samples in the training data set.

Table 2. Training performances: mean of 50 random selections.

Performances	Accuracy (%)	Mean	Variance
ANN	99.40	0.50	0.58
ELM	96.11	3.26	1.17
SVM	95.19	4.04	5.14

Table 3. Testing performance: mean of 50 random selections.

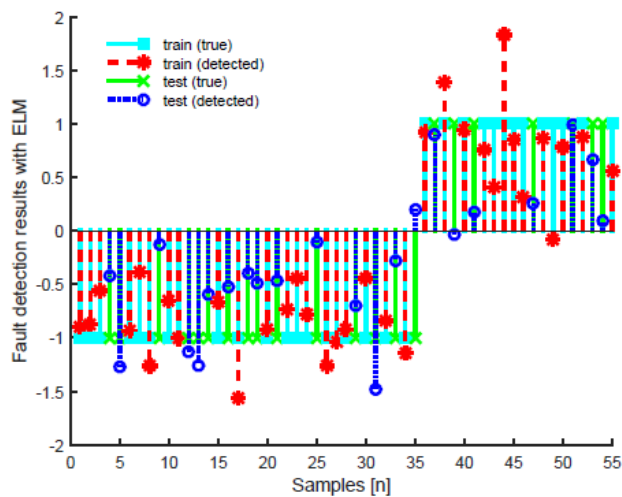
Performances	Accuracy (%)	Mean	Variance
ELM	96.66	2.80	1.42
SVM	91.57	7.08	5.48
ANN	90.85	7.68	4.02

Table 3 presents the testing performance results of the intelligent models. In fact, testing performances of the models are essential to us. For each experiment, intelligent models are tested by the samples which they did not see in the training phase. It is seen that, ELM model provides the best accuracy and the least mean and variance of the false predictions while SVM model is slightly better than the ANN. It can be deduced that, ANN memorizes the input-output relation of the training data samples, thus its prediction of the class labels belonging to the testing samples are not as accurate as for the training samples. Note that, both random and deterministic learning (parameter optimization using LSE) are adopted to train the ELM model. In random learning, parameters have a chance

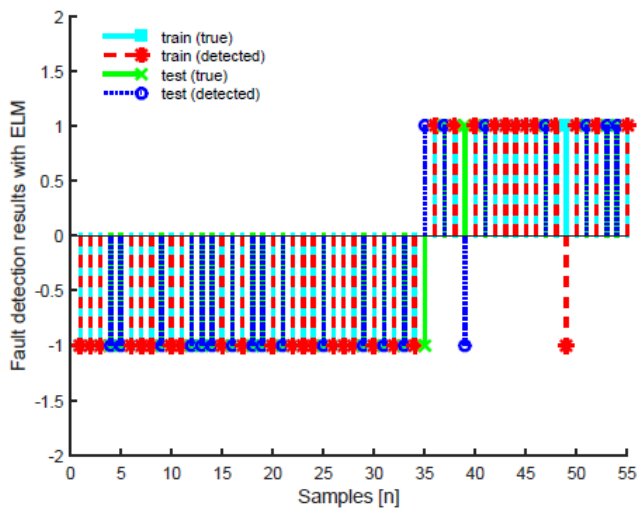
to escape from the local minima up to some degree, which is one reason for the superiority of ELM.

The other contribution to the superiority comes from the fact that, while using LSE, there is one global minimum and no local minima. Hence, the parameters found by LSE will always be globally optimum. In contrast, ANN model employs LM method to optimize its parameters, which is, optimum parameters may be obtained reaching a local minimum of the cost function of prediction errors. Since the minimized cost function is nonlinear in the parameters, LM performs iterative nonlinear optimization. Note that, SVM adopts iterative approaches for quadratic programming as well. However, the reason why test performance of SVM is slightly better than ANN is the structural risk minimization principle, which prevents SVM from focusing solely on the minimization of training error but brings balance to training/testing performances.

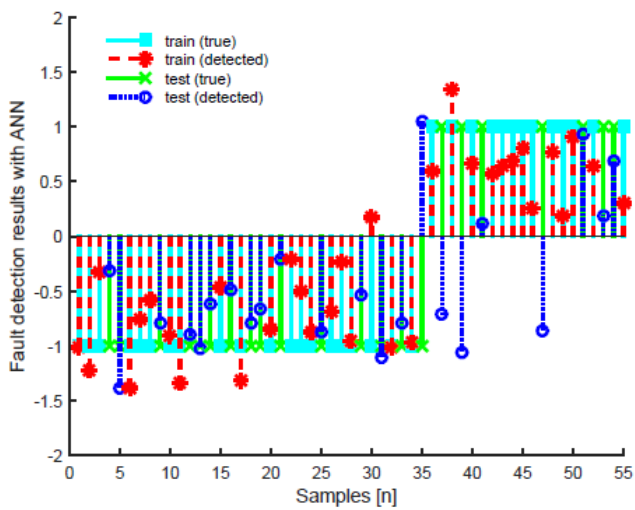
Fig. 4 shows the defect detection results with a randomly selected data set. These results are given to show the difference between the outputs of the models with and without saturation where the saturation function is given by (4). Note that, it is the usual way to see the outputs using a saturation function in classification problems since the classifier models are not expected to yield exact class labels of the samples in $\{-1,+1\}$.



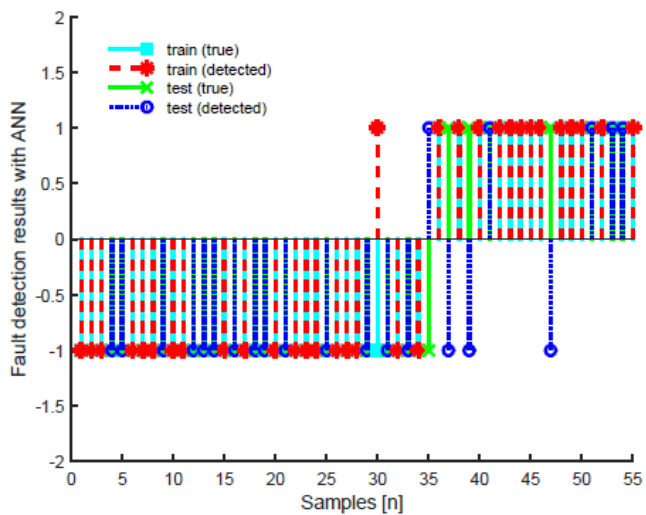
(a) ELM output without saturation.



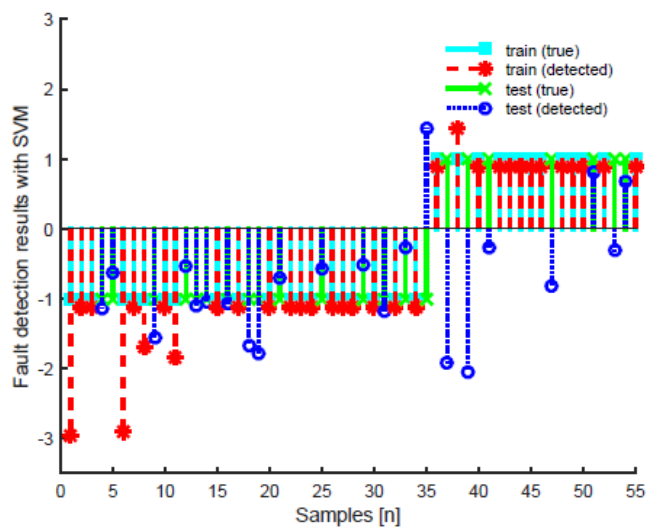
(b) ELM output with saturation.



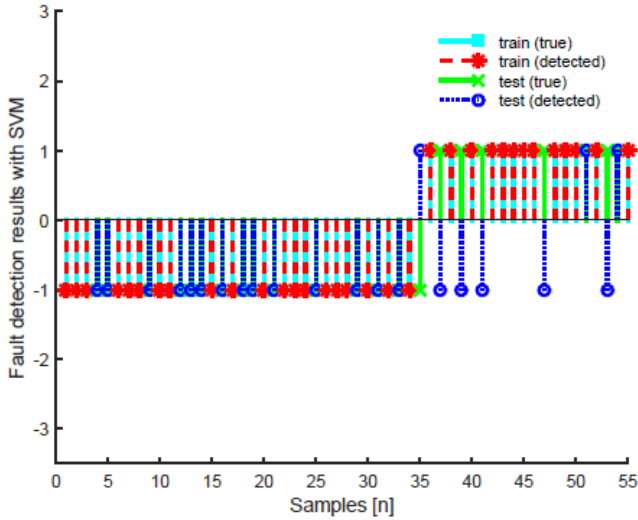
(c) ANN output without saturation.



(d) ANN output with saturation.



(e) SVM output without saturation.



(f) SVM output with saturation.

Figure 4. Fabric defect detection results with and without saturation. Results are obtained for a randomly selected training-testing data set which does not show complete performance of the intelligent models.

In this study, it is desired to analyze both performances and robustness of the models. Fig. 4 shows that, SVM performs output prediction in a more robust way compared to ELM and ANN models. SVM predictions without saturation do not locate as close to the zero decision boundary as those of ELM and ANN are, especially for the training data. Therefore, in case of noisy features, SVM is expected to achieve robust classification performance. If there is no significant noise in the data, due to the high performance of ELM, it can be used with saturation for the defect detection problem.

Another important result is that, designed intelligent models make false detections on the defective fabric samples of the training and especially testing data sets. In fact, it is due to the difficulty of defect detection in Buldan fabric. The extracted features from defective and normal fabric images are very similar. Performance of the ELM model is satisfactory but using different feature extraction methods, its performance can be improved.

5 Conclusion

In this chapter, defects in Buldan fabric were investigated in detail. Artificial neural network, extreme learning machine and support vector machine were designed optimally and employed in detection of defect in the fabric images. Results show that extreme learning machine provides the best performance which is in a satisfactory level. Use of extreme learning machine in detection of defect in fabric images contributes to both a growing topic in textile and particularly to Buldan fabric. As a future work, a different approach for binary classification will be adopted such that support vector data description method is planned to be employed in an unsupervised fashion, which does not need class labels in the training phase. In an other work, multiclassification case will be studied where the samples will be detected as normal or defective with a particular type of defect.

6 References

- Arpaci, Melek, Utkun, Emine, Yoresel Buldan Kumaslarinin Gecmisteki ve Gunumuzdeki Kullanım Alanlarinin Karsilastrilmesi, In: Uluslararası Katilimli Guzel Sanatlar ve Tasarım Sempozyumu, 2009.
- Beyhan, Selami, and Kadir Kavaklioglu. "Comprehensive Modeling Of U-Tube Steam Generators Using Extreme Learning Machines." *IEEE Transactions On Nuclear Science* 62.5 (2015): 2245-2254.
- Bu, H.-G., Huang, X.-B., Wang, J., and Chen, X. 2010. Detection of Fabric Defects by Auto-Regressive Spectral Analysis and Support Vector Data Description. *Textile Research Journal*, 80(7), 579–589.
- Elman, Jeffrey L. "Finding Structure In Time." *Cognitive Science* 14.2 (1990): 179-211.
- Ertugrul, Irfan, and Emine Utkun. Buldan Tekstil Sanayiinin Gelişimi Ve Envanter Araştırması. Ekin Basın Yayın Dağıtım-Bursa, 2009.
- Gaurav, Dinesh Dileep, and Renu Ramesh. "A Feature Extraction Technique Based on Character Geometry for Character Recognition." *Arxiv Preprint Arxiv:1202.3884* (2012).
- Holland, John H. "Genetic Algorithms and Adaptation." *Adaptive Control of Ill-Defined Systems*. Springer, Boston, Ma, 1984. 317-333.
- Huang, Guang-Bin, Qin-Yu Zhu, and Chee-Kheong Siew. "Extreme Learning Machine: A New Learning Scheme of Feedforward Neural Networks." *Neural Networks* 2 (2004): 985-990.
- Huang, Guang-Bin, Qin-Yu Zhu, and Chee-Kheong Siew. "Extreme Learning Machine: Theory and Applications." *Neurocomputing* 70.1-3 (2006): 489-501.
- Huang, Guang-Bin. "An Insight Into Extreme Learning Machines: Random Neurons, Random Features and Kernels." *Cognitive Computation* 6.3 (2014): 376-390.
- Ljung, Lennart. "Black-Box Models from Input-Output Measurements." *IMTC 2001. Proceedings Of The 18th IEEE Instrumentation and Measurement Technology Conference. Rediscovering Measurement in the Age of Informatics (Cat. No. 01ch 37188)*. Vol. 1. IEEE, 2001.
- Ozdemir, Serkan, and Emine Utkun. "Comparing The Profit Margin of Textiles Woven on a Hand Loom and Automatic Loom in the Textile Industry: Case of Buldan City." (2014).
- Oztürk, Fatma, and Esen Coruh. "An Application Concerning Utility of Buldan Weaves inn Fashion Industry." 2013 (Volume: 20)91 (2013).

- Scales, Laurence., E. Introduction to Non-Linear Optimization. Macmillan International Higher Education, 1985.
- Schölkopf, Bernhard; Platt, John C.; Taylor, John S.; Smola, Alex J.; Williamson, Robert C. "Estimating the Support of A High-Dimensional Distribution." *Neural Computation* 13.7 (2001): 1443-1471.
- Tabassian, Mahdi, Reza, Ghaderi, and Reza Ebrahimpour. "Knitted Fabric Defect Classification for Uncertain Labels Based on Dempster–Shafer Theory of Evidence." *Expert Systems with Applications* 38.5 (2011): 5259-5267.
- Tang-Jun, and Hai-Ru Long. "Feature Matching of Fancy Weft Knitted Stitch Based on Modified Surf Algorithm." *Textile Research Journal* 85.7 (2015): 751-758.
- Tax, David Mj, and Duin, Robert, Pw. "Support Vector Data Description." *Machine Learning* 54.1 (2004): 45-66.
- Utkun, Emine, and Kirtay, Erhan. "El Dokumalarında Kaliteyi Ve Fiyatı Etkileyen Faktörler, 1." *Uluslararası* 5 (2009): 623-633.
- Utkun, Emine, Oktay, Pamuk, and Ziyet, Ondogan. "Buldan Bezinden Tasarlanan Farklı Bayan Bluzlarının Optimum Kumaş Kullanım Miktarının Hesaplanması, 1." *Uluslararası* 5 (2009): 349-356.
- Utkun, Emine, and Arpacı, Gundogan M. "Buldan Dokumalarının Kullanım Alanının Arttırılmasına Yönelik Bir Uygulama Çalışması." *Ejovoc (Electronic Journal of Vocational Colleges)* 4.4 (2014): 73-7.
- Utkun, Emine, and Ziyet Ondogan. "Determination of the Ideal Fabric Width of the Classical Women's Blouse Models Made from Buldan Cloth." (2012).
- Vapnik, Vladimir, N. "The Nature of Statistical Learning Theory" Springer, New York (1999).
- Wong, W.K.; Yuen, C.W.M.; Fan, D.D.; Chan, Theresa, and Fung, Eric. "Stitching Defect Detection and Classification Using Wavelet Transform and BP Neural Network." *Expert Systems with Applications* 36.2 (2009): 3845-3856.
- Wu, Ying; Zhou, Jian; Akankwasa, Nicholas T.; Wang, Kai, and Wang, Jun. "Fabric Texture Representation Using the Stable Learned Discrete Cosine Transform Dictionary." *Textile Research Journal* 89.3 (2019): 294-310.
- Yuen, C.W.M.; Wong, W.K.; Qian, S.Q.; Fan, D.D., Chan, L.K., and Fung, E.H.K. 2009. Fabric Stitching Inspection Using Segmented Window Technique and BP Neural Network. *Textile Research Journal*, 79(1), 24–35.

APPENDIX

A1. Levenberg-Marquardt Pseudocode

LM algorithm is given in pseudocode below.

% LM optimization technique as an ALGORITHM %

% Definition of the algorithm hyperparameters

DEFINE scaling factor, μ_s ;

initial damper coefficient value μ_0 ;

minimum value for μ , μ_{min} ;

maximum value for μ , μ_{max} ;

cost function threshold, E_{trsh} ;

max number of iterations, k_{max} ;

threshold for stop criterion 1, $eps1$;

threshold for stop criterion 2, $eps2$;

threshold for stop criterion 3, $eps3$;

% Initialization of the algorithm

$k = 1$;

set initial values for the parameters, θ_0 ;

assign initial values to the parameter vector, $\theta_k = \theta_0$;

$\mu = \mu_0$;

Calculate e_k , E_k , J_k ;

% Logical flag

$innerloopcriterion \leftarrow true$;

% Iterative optimization begins

while $k \leq k_{max}$ **do**

while $innerloopcriterion$ **do**

 Calculate candidate correction term $\Delta\theta_k$;

 Calculate $candidateE_k$;

if $candidateE_k < E_k$ **then**

$\mu = \mu/\mu_s$;

```

        break;
    else
         $\mu = \mu \times \mu_s$ ;
        if ( $\mu > \mu_{max}$  or  $\mu < \mu_{min}$ ) then
             $\Delta \theta_k = 0$ ;
            innerloopcriterion  $\leftarrow$  false;
        end
    end
end
 $\theta_{kold} = \theta_k$ ;
 $\theta_k = \theta_k + \Delta \theta_k$ ;
Calculate  $e_k, E_k, J_k$ ;
if  $E_k < eps1$  then
    break;
if  $\|\theta_k - \theta_{kold}\| < eps2$  then
    break;
grad =  $J_k^T e_k$ 
if  $\|\text{grad}\| < eps3$  then
    break;
if !innerloopcriterion then
    break;
 $k = k + 1$ ;
end

```

Algorithm 1. Pseudocode of LM algorithm.

CHAPTER 16

Environmental Isotope Hydrology and Recharge Dynamics of the Süleymanlı-Ilıca Spring System

**Yusuf Uras¹ & Samet Ahmet Purdaş² &
Yağmur Uysal³**

¹ Prof. Dr., Kahramanmaraş Sutcu Imam University, Geological Engineering Department, Kahramanmaraş, ORCID: 0000-0001-5561-3275

² MSc., Kahramanmaraş Sutcu Imam University, Geological Engineering Department, Kahramanmaraş, Turkey.

³ Prof. Dr., Mersin University, Environmental Engineering Department, 33343, Mersin-Turkey. ORCID: 0000-0002-7217-8217

1. INTRODUCTION

Hydrogeology is the scientific discipline that examines the occurrence, distribution, and movement of groundwater, as well as the physical and chemical properties governing water flow within geological formations (Mutua et al., 2022). Water is fundamental for sustaining life, yet only about 0.74% of global water resources consist of potable freshwater that poses no health risks (Gleick, 1998). During the hydrological cycle, groundwater interacts with various geological units and therefore deviates from pure H₂O, acquiring dissolved ions and trace constituents whose concentrations depend on lithology, climatic conditions, recharge processes, and water–rock interactions (Dao et al., 2024). Stable and radiogenic environmental isotopes have become essential tools for solving numerous hydrogeological problems, including determining the origin and age of groundwater, identifying recharge elevations, and evaluating subsurface flow dynamics (Mohamed et al., 2021).

This study investigates the cold-water springs of the Süleymanlı–Ilıca region in Kahramanmaraş from the perspective of hydrogeochemistry and isotope hydrology. The investigated springs include Halimin Lake (ISS-1), Ziyaret Spring (ISS-2), Süleymanlı Spring (ISS-3), and Süleymanlı-2 Spring (ISS-4) (Figure 1). The primary objective is to characterize the hydrochemical and isotopic properties of these spring waters, infer their recharge mechanisms, evaluate groundwater flow paths and residence times, and identify the geological controls governing spring discharge. Within this framework, comprehensive assessments were conducted on the geology, hydrogeology, isotope hydrology, and water chemistry of the Süleymanlı–Ilıca region, located approximately 75 km north of the Kahramanmaraş city center (Figure 2).

Although the scope of this study is not medical geology, a concise assessment was made regarding the drinking-water suitability of the analyzed springs. Major ions, pH, electrical conductivity, hardness, and anion–cation balance values fall within acceptable ranges defined by the World Health Organization (WHO, 2022), the European Drinking Water Directive (EU, 2020), and Turkish Standards (TS 266). Trace metals were below detection limits (<DL), and fluoride, nitrate, and chloride concentrations remained well below health-based guideline values. These findings indicate that the Süleymanlı–Ilıca spring waters represent low-mineralized, high-quality meteoric waters that pose no detectable geochemical risk to public health (Deng et al., 2024).

A



D



B



B



C



C



Figure 1 General view of cold water springs in the research area (Süleymanlı-Ilıca) A: Halimin Lake spring (ISS-1), B: Ziyaret spring (ISS-2), C: Süleymanlı water spring (ISS3), D: Hacı İbrahim Uşağı Water Spring (ISS-4).

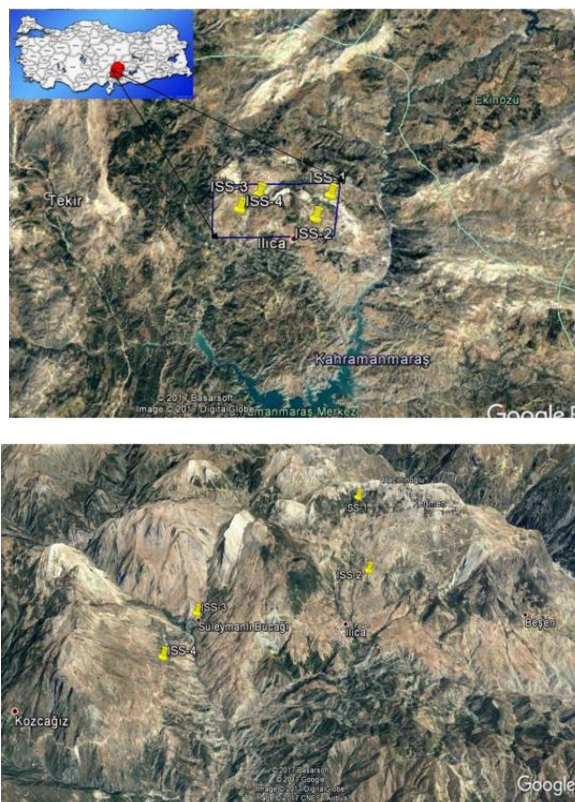


Figure 2 Locations of water resources in the study area and a map view of the study area (Web 1, 2017)

1.1. Geology of the Study Area

Kahramanmaraş Province and its surroundings are located within a region where the Arabian and Anatolian plates are sutured, resulting in a geologically complex setting shaped by major tectonic events associated with continent–continent collision. Owing to these significant geological processes, the area represents a structurally intricate zone and thus constitutes an attractive region for geoscientific investigations. In general, two major deformation phases are observed in the area: the first is related to the closure of the Neotethys Ocean during the Late Cretaceous, and the second is associated with continent–continent collision during the Eocene. Following the collision, the continued northward movement of the Arabian Plate induced a new compressional regime during the Miocene. This regime led to intense deformation, producing major faults and folds within the rock units deposited between the Late Maastrichtian and Early Miocene. The study area selected for this research occupies a key position, as it

represents these different deformation periods and the corresponding rock assemblages. The investigation area is located approximately 75 km north of the city center of Kahramanmaraş and lies within the boundaries of the 1:25,000-scale Gaziantep M37-b3 and Gaziantep M37-b4 map sheets (Figure 3).

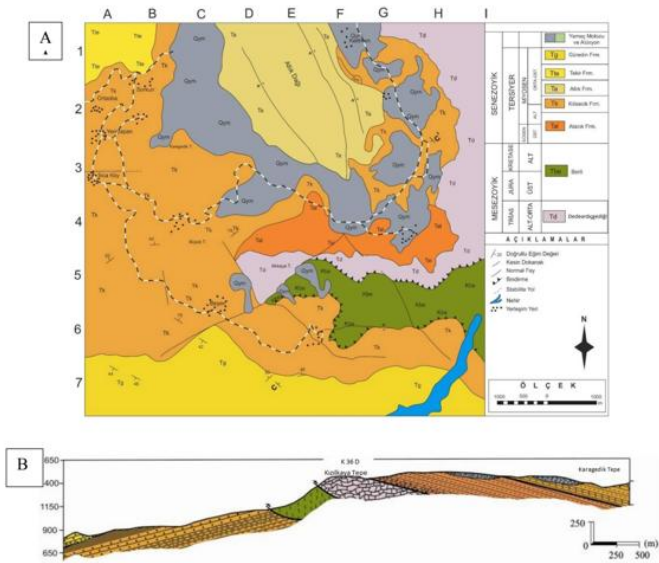


Figure 3 A: Geological map of the study area, B: Cross-section of the study area (Modified from Varol et al., 2012)

1.4. Stratigraphic and Structural Architecture of the Study Area

Morphologically, the study area is characterized by smaller mountains and hills that constitute the southern extensions of the high ranges such as Mount Berit to the north and Mount Engizek to the east, along with deep river valleys that dissect these elevated terrains. The principal topographic highs within the study area include Mount Atlık (2088 m), Ayar Hill (1643 m), Kızılkaya Hill (1512 m), Kale Hill (1412 m), Karagedik Hill (1388 m), Ayvacık Hill (1332 m), Büyükçat Hill (1285 m), Böyük Hill (1201 m), Demirkapı Hill (1162 m), Yılanlıkat Hill (1054 m), Çamsöken Hill (1034 m), and Yazlak Hill (946 m). The Ceyhan River is the largest fluvial system in the study area, while the Menzelet Dam, constructed on this river, represents the largest standing water body (Varol et al., 2012). The region experiences a typical Mediterranean climate, characterized by hot and dry summers and cold and wet winters. Due to its rugged topography, the area lacks suitable land for extensive agriculture. Consequently, local communities primarily rely on livestock breeding and viticulture as their main economic activities.

Within the Berit Nappe, which forms the exposed basement of the study area, the lowermost unit is the Early–Middle Miocene Kilisecik Formation (Tk), consisting of shale, claystone, and sandstone lithologies. East of Beşen, the Kilisecik Formation is tectonically overlain by the Late Jurassic–Early Cretaceous Berit Metaophiolite (Kbe). In the southern parts of the study area, the Kilisecik Formation is unconformably overlain by the Middle–Late Miocene Güredin Formation (Tg), composed of conglomerate, sandstone, and claystone. These units belonging to the Berit Nappe are tectonically overlain by the Dedearıçgediđi Nappe to the east of Beşen.

In the mapped area, the lowermost unit of the Dedearıçgediđi Nappe is the Triassic Dedearıçgediđi Formation (TRd), which is composed predominantly of dolomite, recrystallized dolomite, and limestone lithologies. This unit tectonically overlies the Late Jurassic–Early Cretaceous Berit Metaophiolite (Kbe) situated within the Berit Nappe. Overlying the Dedearıçgediđi Formation is the Late Eocene–Early Miocene Alacık Formation (Tal), which begins with basal levels rich in Nummulites fossils and transitions upward into conglomerate, sandstone, and claystone.

The Alacık Formation is unconformably overlain by the Early–Middle Miocene Kilisecik Formation (Tk). Above the Kilisecik Formation, the Middle Miocene Atlık Formation (Ta), composed of limestone rich in algal and coral fossils, crops out around Mount Atlık, while the Middle–Late Miocene Tekir Formation (Tte) overlies the Kilisecik Formation around Sorkun village. The Atlık Formation is conformably overlain by the Tekir Formation (Tte), which consists mainly of conglomerates. The general characteristics of the lithostratigraphic units cropping out within the two nappe slices in the study area are presented below in chronological order.

1.5. Description and Stratigraphic Position of the Dedearıçgediđi Formation

The study area is generally characterized by steep, rugged, and rocky topography, displaying prominent cliffs and a distinctive grey coloration. The unit that forms extensive exposures in the northeastern part of the area consists predominantly of limestone, dolomite, and recrystallized dolomite. Dolomite and recrystallized dolomite facies are common at the basal levels of the unit, and local marbleization is observed near the tectonic contact. Limestone becomes the dominant lithology toward the upper levels. The limestones exhibit a weathered surface in greyish-beige tones, whereas their fresh surfaces display light red-beige to pinkish colours. These limestones are hard, compact, and massive in

appearance. The dolomite and recrystallized dolomite horizons, in contrast, display dark grey weathered surfaces and black fresh surfaces. These levels are massive, fractured, and mechanically competent.

Based on regional correlations and fossil-bearing units above and below it, Baydar and Yergök (1996) proposed that the age of the Dedeardıçgediği Formation ranges from Late Permian to Middle Triassic. However, due to the absence of Permian lithologies or fossil assemblages within the study area, a Triassic age assignment is considered more appropriate. In the field, the Dedeardıçgediği Formation occupies a tectonic position above the Berit Metaophiolite. It is unconformably overlain by the Kilisecik Formation near Elsen village and by the Alacık Formation near Eleman village (Baydar & Yergök, 1996).

1.6. Lithology, Structure, and Tectonic Relations of the Berit Metaophiolite

The term Berit Metaophiolite was first introduced to describe the metamorphosed ophiolitic rocks exposed on Mount Berit (Perinçek & Kozlu, 1984; Robertson, 2006). The unit comprises various sedimentary and metamorphic blocks, including micritic limestone, crystalline limestone, conglomerate, quartzite, and mica schist. Overall, the presence of peridotite, gabbro, diabase, and numerous dike intrusions indicates that the unit represents a relatively complete ophiolitic sequence.

Within the study area, the Berit Metaophiolite displays variable lithologies, primarily diabase, mica schist, and strongly weathered grey–green gneisses. The weathered surfaces typically exhibit green, light green, and reddish-brown tones, whereas the fresh surfaces range from dark green to dark red. The rocks are generally very hard, compact, intensely fractured, and contain calcite-filled veins. At the contact between the Berit Metaophiolite and the overlying Dedeardıçgediği Formation, intense cataclastic deformation is observed, indicating significant tectonic activity along this boundary.

Previous studies have assigned the unit a Late Jurassic–Early Cretaceous age (Perinçek & Kozlu, 1984), although Gözübol and Gürpınar (1980) reported a Late Cretaceous age. In the study area, the Berit Metaophiolite tectonically overlies the Early–Middle Miocene Kilisecik Formation, and in turn, is itself tectonically overlain by the Dedeardıçgediği Formation (Varol et al., 2012).

1.7. Lithostratigraphic Characteristics and Stratigraphic Relationships of the Alacık Formation

The name Alacık Formation was first introduced by Erdoğan (1975) to describe the variegated clastic sediments of Late Eocene age exposed around the Helete district near Göksun. The formation begins with basal levels rich in Nummulites fossils, indicating a shallow-marine depositional environment. Up section, it grades into thin- to medium-bedded conglomerates and sandstones, and finally into claystone-dominated intervals toward the top.

The conglomerates are grey to locally reddish in colour, consisting of angular to sub-angular clasts derived from dolomite, radiolarite, ophiolitic rocks, and limestone. They are generally fine- to medium-grained and display frequent fractures and joints, some of which contain calcite infill. The sandstones are grey, coarse-grained, and composed mainly of angular grains sourced from ophiolitic rocks, dolomite, limestone, and radiolarite. The unit is pervasively fractured and displays well-developed jointing, commonly filled with calcite. The claystone are bluish-grey, thinly bedded, and locally interbedded with sandstone lenses, reflecting lower-energy depositional conditions.

Based on lithostratigraphic relationships and regional correlations, the Alacık Formation is assigned a Late Eocene–Early Miocene age (Darbaş & Gül, 2006). The formation overlies the Dedeağaçlı Formation with an angular unconformity, indicating a significant tectonic phase prior to deposition. It is itself unconformably overlain by the Kilisecik Formation, reflecting another regional deformation and erosion event (Varol et al., 2012).

1.8. Lithological Characteristics, Fossil Content, and Stratigraphic Relations of the Kilisecik Formation

The term Kilisecik Formation was first introduced by Gül (2000) to describe the marl, shale, and sandstone lithologies exposed around the Kilisecik locality. The formation begins with shale at its base and continues upward as an alternating sequence of marl and sandstone. In the Yazlak Hill area, a limestone bed approximately 10 meters thick is observed within the succession. The shales display yellowish-green weathered surfaces and bluish-green fresh surfaces. They are thin- to medium-bedded, fissile, and locally exhibit chaotic structures. The marls also show yellowish-green weathered surfaces, with light to dark green fresh surfaces, and are thin- to medium-bedded with moderate hardness and characteristic concentric fracturing. The sandstones display reddish-yellow weathered surfaces and light green fresh surfaces; they are thin-bedded, intensely fractured, commonly with calcite-filled veins, and of moderate hardness. Slump

structures observed at several levels indicate depositional instability, possibly related to synsedimentary tectonism or slope failure.

Based on the identification of fossils such as *Miogypsina* sp., *Elphidium* sp., *Orbulina* sp., *Lepidocyclina* sp., and *Globorotalia* sp., the formation was assigned an Early–Late Miocene age (Darbaş & Gül, 2017). Additional microfossils including *Globigerinoides trilobus* Reuss, *Globigerinoides ruber* (d'Orbigny), *Globorotalia obesa* Bolli, and *Orbulina universa* (d'Orbigny) suggest an Early–Middle Miocene age (Gül et al., 2005). The Kilisecik Formation overlies the Alacık Formation with an angular unconformity. It is conformably overlain by the Tekir Formation near Sorkun village and by the Atlık Formation around Mount Atlık. East of Beşen village, the Kilisecik Formation is tectonically overlain by the Berit Metaophiolite. The spring waters investigated in this study emerge from within the lithologies of the Kilisecik Formation (Figure 4).



Figure 4 View from the Kilisecik Formation

1.9. Depositional Characteristics, Fossil Assemblages, and Stratigraphic Position of the Atlık Formation

The cream-coloured, thick-bedded to locally massive limestone unit exposed in the study area was previously described as the Atlık Member of the Lice Formation and the Parpiyayla Limestone Member of the Yenice kale Formation (Baydar & Yergök, 1996; Darbaş, 2021). Earlier studies reported that this unit consists predominantly of limestone and exhibits lateral transitions into sandy carbonates or marls, and toward the upper levels into marl–limestone alternations (Yılmaz et al., 1988). For the first time in this study, the distinctive limestone unit

exposed on Mount Atlık, characterized by abundant algal and coral fossils, is formally defined as the Atlık Formation.

The Atlık Formation consists of greyish-beige weathered surfaces and beige fresh surfaces. The limestones are medium- to thick-bedded, locally massive, hard, and mechanically competent, with abundant fractures and calcite-filled veins. The rocks are sparitic in texture and contain rich assemblages of algal and coral fossils, indicating deposition within a reefal environment (Sinanoğlu & Sarkar, 2025). Biostratigraphic studies have yielded a Middle Miocene age based on benthic and planktonic foraminifera such as *Orbulina* sp., *Globigerina* sp., *Globigerinoides* sp., and members of the Textulariidae and Ataxophragmidiidae families, as well as diverse algal fossils. Additional taxa—*Amphistegina* cf. *lessoni* (d'Orbigny), *Operculina complanata* (Defrance), *Pararotalia* sp., *Austrotrillina howchini* (Schlumberger), together with coral and *Ostrea* fossils—support a Burdigalian (Early Miocene) age. Considering all available fossil evidence, the Atlık Formation is assigned an Early–Middle Miocene age in this study.

1.10. Sedimentology, Depositional Environment, and Stratigraphic Framework of the Tekir and Güredin Formations

The conglomerates extensively exposed around Tekir village were first described as the Tekir Formation by Gül (2000). The formation is generally grey, moderately to highly indurated, and medium- to thick-bedded. The clasts are derived from ophiolitic rocks, dolomite, radiolarite, and limestone, and range from 2 to 25 cm in diameter. They are mostly well rounded. The unit contains a sand-rich matrix and is predominantly clast-supported, although matrix-supported intervals also occur. Stratigraphically, the Tekir Formation conformably overlies the Middle Miocene Atlık Limestone, suggesting a Middle–Late Miocene age. Sedimentological features indicate deposition in a progressively shallowing setting, likely along a submarine slope associated with a fan-delta system (Yılmaz et al., 1988).

The term Güredin Formation (Gül, 2000) is used to describe a sequence composed of debris-flow deposits, sandy conglomeratic channel fills, and interbedded sandstone–claystone units. In the study area, the formation begins with conglomerates at its base. These conglomerates consist of clasts derived from ophiolitic rocks, dolomite, limestone, and schist; their colours vary from brown to reddish, with green hues developed in carbonate-enriched zones. The clasts are generally rounded, medium- to coarse-grained, moderately indurated, and exhibit good sorting but poor grading. Up section, the conglomerates

gradually transition into sandstones. The sandstones display reddish weathered surfaces and greenish-grey fresh surfaces. They are commonly medium-bedded, locally thin-bedded, of moderate hardness, and intensely fractured. Sandstone intervals transition both laterally and vertically into claystone units. The claystones have yellowish-green weathered surfaces and green fresh surfaces, are thin-bedded, exhibit low mechanical strength, and locally show fissile texture.

Microfossils recovered from the Güredin Formation include *Globobulimina pupoides* (Cushman), *Globobulimina pyrula* (Cushman), *Haplophragmoides* sp., *Valvulineria complanata*, *Orbulina suturalis* Bornnimmann, *Orbulina universa* (d'Orbigny), *Globigerinoides* sp., *Globigerinoides obliquus* (Bolli), *Hastigerina siphonifera* (d'Orbigny), *Globorotalia continuosa* (Blow), *Globigerina woodi-aperture* (Jenkins), *Globigerina praebulloides* (Blow), *Uvigerina flinti* (Bowerbank), *Uvigerina anberiana* (d'Orbigny), *Eponides* sp., *Cibicides* sp., *Bulimina inflata* (Esper), *Sigmoilopsis celata* (Costa), and *Globigerina falconensis* (Blow). These assemblages support a Middle–Late Miocene age (Kırdan, 2004).

1.11. Structural Framework of the Study Area: Fault Architecture and Fold Geometry

The study area contains numerous steep, near-vertical faults, largely controlled by the high regional topographic gradients. These faults are predominantly observed in the northern parts of the area, particularly around Mount Atlık and east of Beşen village. The major fault systems exhibit NW–SE and NE–SW orientations, with additional E–W trending structures. On the eastern and western flanks of Mount Atlık, the faults form a series of step-like, en echelon developments that cut through the lithologies of the Atlık Formation. A NW–SE trending fault extending from the western slope of Kızılkaya Hill toward Beşen cuts across the Alacık, Kilisecik, and Dedeardıçgediği formations, as well as the Berit Metaophiolite. Another NW–SE oriented fault, considered the continuation of this structure, offsets the Alacık and Dedeardıçgediği formations on the western slope of Mount Atlık. South of Yazlak Hill, two E–W trending faults cut through the Kilisecik Formation and the Berit Metaophiolite.

In the northwestern part of the study area, near the İlica settlement—which is known for its thermal springs—an E–W oriented vertical fault has been identified cutting the Kilisecik Formation. This structure is thought to play a major role in controlling the ascent and discharge of thermal waters to the surface. In general, the E–W faults appear to be the oldest, while the NW–SE and NE–SW oriented faults represent younger, conjugate sets. In addition to these faults, a NE–SW

trending thrust fault has been mapped east of Beşen village, where the Berit Metaophiolite is tectonically emplaced over the Kilisecik Formation, and the Dedeardıçgediği Formation is similarly emplaced over the metaophiolite (Varol et al., 2012). In the southern part of the study area, distinct fold structures are developed within the Güredin Formation. The most prominent of these is a NW–SE trending syncline extending between İncirliçedik Hill and Kızıldamlar Hill west of Boybeyler village. North of this syncline, a NW–SE trending anticline—believed to be its structural continuation—extends from the southwestern slope of Binektaş Hill toward Yücekah Hill farther northwest (Varol et al., 2012).

1.12. Geodynamic Evolution and Depositional History of the Study Area

The study area lies within the suture zone where the Arabian and Anatolian plates collided and welded together. Lithostratigraphic units ranging in age from Triassic to Quaternary are exposed throughout the region. Although the rocks in the wider area experienced intense deformation during the Late Cretaceous, Eocene, and Miocene, the deformation features preserved within the study area are predominantly Miocene and younger. The depositional environments and the geodynamic evolution of the exposed units are interpreted below (Varol et al., 2012).

During the Early Mesozoic, continental rifting initiated the fragmentation of crustal blocks, ultimately leading to the opening of the southern branch of the Neotethys Ocean between the Malatya (Tauride) platform in the north and the Arabian Plate in the south. Extensive carbonate deposition occurred along the passive continental shelf during the oceanization phase. The Dedeardıçgediği Formation is interpreted to represent these thick carbonate sequences that accumulated on the shelf during the Mesozoic oceanic development.

Oceanization ceased during the Late Cretaceous as the southern branch of the Neotethys began to close. The regionally extensive Berit (Göksun) Ophiolite was emplaced during the Late Cretaceous within a north-dipping intra-oceanic subduction zone. The association of mafic to felsic volcanism and volcanoclastic sediments indicates the presence of a newly formed (ensimatic) intra-oceanic volcanic arc. Subduction that initiated in the Late Cretaceous persisted into the Eocene and intensified as the oceanic crust of the southern Neotethys was progressively consumed (Robertson et al., 2006).

As collision between the Arabian and Anatolian plates progressed, a flexural basin developed on the subsiding Arabian Plate beneath the load of the northward-advancing allochthonous units (Önal, 1985; Kozlu, 1997). Initially, high-energy deposits accumulated within this basin, followed by deeper-water

sediments associated with basin deepening. The Alacık Formation records this transition and began to accumulate during the Late–Middle Eocene. However, by the onset of the Aquitanian, the basin underwent significant deformation and contraction. During the Late Eocene to Early Miocene, the region was affected first by large-scale nap emplacement and subsequently by deformation controlled by strike-slip faulting related to crustal thickening. A regional transgression during this period facilitated the deposition of the Kilisecik Formation, composed mainly of deep-marine claystone (Gül et al., 2005).

In the Early Miocene, continued southward advance of the allochthonous thrust sheets resulted in further crustal shortening and localized uplift, promoting the deposition of the reefal limestones of the Atlık Formation in relatively shallow marine settings. As the allochthonous units advanced southwards, the basin narrowed, and under the combined load of thrust sheets and accumulating sediments, the basin deepened again (Önal, 1985). During the Middle–Late Miocene, a renewed compressional regime caused further uplift and deformation. This period is characterized by the tectonic emplacement of the Berit Metaophiolite over the Early–Middle Miocene Kilisecik Formation, followed by the tectonic emplacement of the Dedeağzı Formation over the metaophiolite (Varol et al., 2012). On the submarine slopes formed by nap emplacement, debris-flow deposits of the Middle–Late Miocene Güredin Formation accumulated in shallow areas, whereas claystone-rich sediments of the same formation were deposited in deeper settings (Önal, 1985). From the Middle Miocene to the end of the Late Miocene, progressive basin closure took place. During this stage, the channelized conglomeratic deposits of the Tekir Formation were deposited, marking the final phase of sedimentation in the basin.

2. MATERIALS AND METHODS

Over the course of a year, cold water samples were routinely taken from the springs in the Süleymanlı-Ilica area. The samples underwent additional laboratory examinations after being first measured on-site. To ascertain the sources and characteristics of the waters, the geological, geochemical, isotopic geochemistry, and hydrochemical features of the cold springs from the research region were investigated. All data were assessed collectively and in-depth analysis were carried out. The possible utility of the geothermal water sources was also revealed by the chemical studies.

Samples were taken from various sites throughout the year in order to evaluate the Süleymanlı-Ilica area's cold springs' usability. Polypropylene containers (1000 mL) were used to collect the samples, which were designated as ISS-1

(Halimin Lake spring), ISS-2 (Ziyaret spring), ISS-3 (Süleymanlı water spring), ISS-4 (Hacı İbrahim Uşağı Water Spring). At every test location, these containers were cleaned with natural water to avoid contamination from foreign materials.

Ion Chromatography (IC), ICP-OES, and ICP-MS were used at ACME Analytical Laboratories in Canada to perform heavy metal and anion-cation analyses. Temperature, conductivity (EC), dissolved oxygen (DO), pH, and other factors were measured on-site. More samples were brought to the lab and kept in a refrigerator at +4°C for additional analysis. At the DSI TAKK Department Isotope Laboratory, water samples from the four sources (ISS-1, ISS-2, ISS-3 and ISS-4) were also examined using IRMS and Tritium (^3H) procedures in accordance with IAEA experimental standards. The origins of the waters were ascertained, rock-fluid interactions were evaluated, boiling-evaporation processes resulting from abrupt pressure drops underground were identified, the mixing of waters from various aquifers and/or sources was examined, and general information regarding the age of the groundwater was gathered using oxygen-18 (^{18}O) and deuterium (^2H) isotopes.

3. RESULTS and DISCUSSIONS

3.1. Hydrogeological Cycle, Climatic Conditions, and Precipitation Characteristics of the Study Area

During the hydrogeological cycle—a continuous process involving evaporation, condensation, precipitation, surface runoff, and infiltration—water undergoes substantial physical and chemical changes (Clark & Fritz, 1997). Due to its molecular structure, water is one of the most effective solvents on Earth. The chemical variability of natural waters can be attributed in part to the diverse lithologies they encounter along their flow paths. Consequently, it is expected that groundwater and surface water acquire distinct chemical signatures that reflect the geological environments through which they circulate (Drever, 1997).

Kahramanmaraş is characterized by a Mediterranean climate, with hot and dry summers and mild, rainy winters. According to data obtained from the Turkish State Meteorological Service, the region receives precipitation mainly from mid-autumn to late spring, with an annual average of 60.41 mm. June, July, and August constitute the warmest months (Figure 5). The mean temperature for August is 36.1 °C, while the lowest recorded temperature—-9.6 °C—occurred in January 1968. The highest temperature on record is 47.5 °C, measured in August 1927.

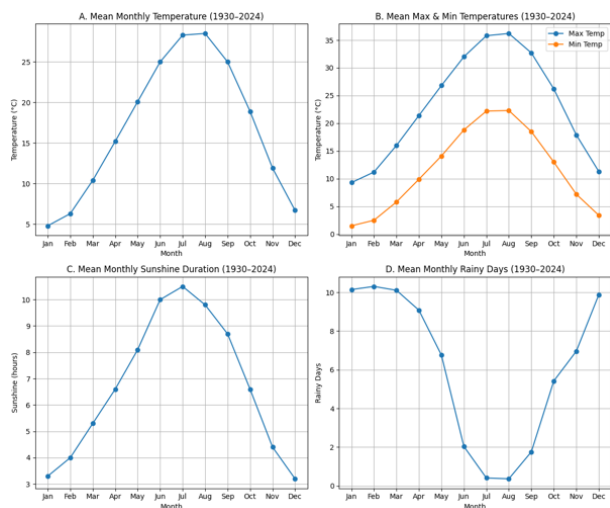


Figure 5. Multi-panel climate summary for the Kahramanmaraş–Süleymanlı–Ilica region based on long-term meteorological records (1930–2024). A: Mean monthly temperature, B: Mean monthly maximum and minimum temperatures, C: Mean monthly sunshine duration, D: Mean monthly number of rainy days. These climatic parameters control evaporation, groundwater recharge, and hydrogeochemical processes in the study area (TSMS, 2025).

3.2. Hydrochemical Characterization and Irrigation Water Quality Assessment of the Süleymanlı–Ilica Springs

Due to these climatic characteristics, the study area possesses a rich surface-water network. In the Süleymanlı–Ilica localities where this study was conducted, the presence of limestone units surrounding the area, combined with their intensely fractured and fissured nature, significantly enhances secondary permeability (Ford & Williams, 2007). This structural configuration plays a crucial role in controlling groundwater flow and is considered to have a major influence on the aquifer quality and the development of the abundant spring waters in the region.

In order to assess the water quality of the Süleymanlı–Ilica springs, a one-year monitoring program was conducted, covering measurements of temperature, pH, electrical conductivity, alkalinity, trace-element concentrations, major anions and cations, and dissolved metals. Throughout the monitoring period, spring temperatures ranged between 10°C and 14°C, indicating that the Süleymanlı–Ilica sources fall within the category of cold-water springs (Table 1). Dissolved oxygen (DO) values, expressed as percent saturation, were found to be relatively similar across all sampled springs, ranging between approximately 20 and 55%.

This corresponds to about 2–6 mg/L at the measured temperatures and indicates oxic conditions without pronounced reducing environments.

Table 1. Summary of the physicochemical characteristics and major ion composition of the Süleymanlı–İlica spring waters (ISS-1, ISS-2, ISS-3, ISS-4). Values are presented as mean ± standard deviation

Geochemical Parameters				
	ISS-1	ISS-2	ISS-3	ISS-4
Temperature (°C)	13.25±0.96	13.06±0.24	13.09±0.49	15.14±1.31
pH	8.03±0.40	8.22±0.11	8.03±0.32	8.19±0.13
EC (µS/cm)	370.25±112.64	325±19.30	462.25±103.25	524.75±105.09
DO (%)	30.75±2.22	33.25±11.00	33.25±15.46	26.50±3.42
Na ⁺ (ppm)	7.18±1.04	3.98±0.57	6.32±2.03	6.32±2.03
K ⁺ (ppm)	1.70±1.34	0.60±0.11	1.27±0.75	0.45±0.03
Ca ⁺² (ppm)	50.51±1.76	65.66±10.80	76.30±8.15	76.92±1.22
Mg ⁺² (ppm)	14.63±7.17	9.68±1.10	13.53±2.33	17.53±0.17
Cl ⁻ (ppm)	3.05±0.39	3.69±1.98	2.60±1.13	2.86±0.23
HCO ₃ ⁻ (ppm)	228.48±69.71	206.50±21.28	256.96±20.11	304.88±20.08
SO ₄ ⁻² (ppm)	21.33±0.76	24.09±1.40	26.93±1.05	13.61±0.30
NO ₃ ⁻ (ppm)	6.93±2.10	10.34±3.28	3.45±0.43	4.85±0.53
F ⁻¹ (ppm)	0.14±0.04	0.20±0.01	0.15±0.02	0.11±0.01
Trace metals (Al, Cd, Cu, Cr, Mn, Ni, Pb etc.) (ppb)	<DL	<DL	<DL	<DL

Note: DL = Instrumental detection limit; EC = electrical conductivity; DO = dissolved oxygen; concentrations of major ions are reported in ppm.

Hydrochemical data obtained from the Süleymanlı–İlica springs demonstrate that all four sources (ISS-1, ISS-2, ISS-3, ISS-4) exhibit a consistent geochemical character that aligns well with national and international drinking water standards, including WHO (2022), EPA, and the EU Drinking Water Directive (2020/2184). The measured temperatures, ranging between approximately 12 and 16°C, are typical of shallow, precipitation-fed spring systems and do not pose any adverse health concern. pH values remain within the slightly alkaline range (7.8–8.2), reflecting the buffering capacity of the carbonate-rich lithologies surrounding the aquifer, particularly limestone and dolomite units. Such alkaline conditions are well within the recommended limits for potable water, enhancing

palatability while reducing the potential for corrosive reactions within distribution systems.

Electrical conductivity values across all sampling points are low, varying between 325 and 525 $\mu\text{S}/\text{cm}$, which corresponds to low total dissolved solids and confirms the absence of salinity-related degradation. These low EC values further indicate minimal anthropogenic influence and a groundwater body predominantly shaped by natural water–rock interaction processes. Major cation and anion concentrations also support this interpretation. Sodium levels, ranging from 4 to 7 ppm, are far below regulatory thresholds and signify a water type free from domestic or agricultural salt contamination. Potassium concentrations are similarly low, reinforcing the lack of agricultural input. Calcium and magnesium concentrations exhibit patterns typical of carbonate aquifers, contributing to moderate water hardness that is not only aesthetically acceptable but also considered beneficial for human health through dietary mineral intake.

Chloride and sulfate concentrations remain well below WHO guideline values, confirming that neither sewage leakage nor evaporite dissolution is influencing the aquifer. Bicarbonate levels are moderate to high, reflecting equilibrium with carbonate minerals and contributing to the stable alkalinity of the waters. Nitrate concentrations across all sources are low, with the highest value (approximately 10 ppm at ISS-2) remaining far below the WHO limit of 50 mg/L. These results indicate that agricultural runoff, livestock activity, or septic contamination is not affecting the study area. From a public health perspective, such low nitrate levels eliminate concerns regarding methemoglobinemia, especially in infants.

Fluoride concentrations, ranging from 0.11 to 0.20 ppm, are considerably lower than the WHO upper limit of 1.5 mg/L and do not present any risk of fluorosis. At the same time, fluoride levels remain below the optimal range for dental caries prevention but fall within acceptable natural background concentrations. Dissolved oxygen values, when corrected to mg/L, correspond to clean, oxygenated spring conditions and indicate that the aquifer is not subject to reducing geochemical environments. This supports the interpretation of active recharge, short to moderate residence time, and minimal organic contamination.

Perhaps the most important finding is the uniformly low concentrations of trace metals across all sampling locations. Elements such as Al, Cd, Cu, Cr, Mn, Ni, Pb, Fe, Zn, and Co were measured at levels near or below analytical detection limits. These results highlight the pristine geochemical nature of the aquifer and confirm the absence of industrial, mining, or geogenic metal contamination. The

extremely low metal concentrations carry substantial public health significance, as chronic exposure to metals—particularly Pb, Cd, and Cr—can cause neurological, renal, and carcinogenic effects; however, no such risk exists for the Süleymanlı–İlica springs.

Collectively, the hydrogeochemical profile of the Süleymanlı–İlica region reflects a clean, carbonate-hosted aquifer system that is largely unaffected by anthropogenic stressors. All measured parameters comply with major international drinking water guidelines. The combination of low salinity, moderate hardness, slightly alkaline pH, low nitrate levels, and negligible metal content demonstrates that the springs represent a high-quality water source suitable for long-term human consumption without posing any chemical health risks. These results additionally suggest that the aquifer benefits from effective natural filtration, minimal pollutant loading, and stable recharge dynamics controlled primarily by regional climatic and geological conditions. Piper diagram results show that the Süleymanlı–İlica springs have a distinctly Ca–HCO₃ type hydrogeochemical facies (Piper, 1944) (Figure 6). This distribution confirms that intense water-rock interaction with carbonate units in the region is the dominant process and that the waters have a geologically similar origin. The close proximity of the springs in the central diamond field is consistent with low salinity, weak evaporitic contribution, and limited anthropogenic influence in all samples.

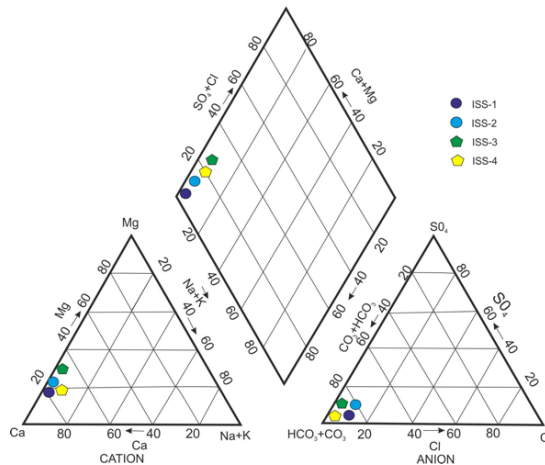


Figure 6 Piper diagram illustrating the major-ion chemistry of the Süleymanlı–İlica spring waters (ISS-1, ISS-2, ISS-3, ISS-4). All samples plot in the Ca–HCO₃ hydrochemical facies, indicating water–rock interaction dominated by dissolution of carbonate lithologies. The clustering of the four sampling points suggests a similar geochemical evolution controlled primarily by local carbonate bedrock and limited anthropogenic influence.

In Gibbs diagrams, all samples fall within the “rock dominance” zone with low TDS values (Gibbs, 1970). The low ratios of $\text{Na}^+ / (\text{Na}^+ + \text{Ca}^{2+})$ and $\text{Cl}^- / (\text{Cl}^- + \text{HCO}_3^-)$ suggest that the chemical composition of the waters is primarily controlled by the dissolution of carbonate rocks and limited silicate weathering. No enrichment trend indicating evaporation control or marine intrusion was observed in any of the samples; this is consistent with the region's inland, continental, and low-evaporation hydrogeological location. According to the Gibbs Type I plot (Figure 7), all samples fall clearly within the rock-dominance zone. Figure 8 further demonstrates that the $\text{Cl}^- / (\text{Cl}^- + \text{HCO}_3^-)$ ratios remain low, supporting a carbonate-dominated geochemical environment.

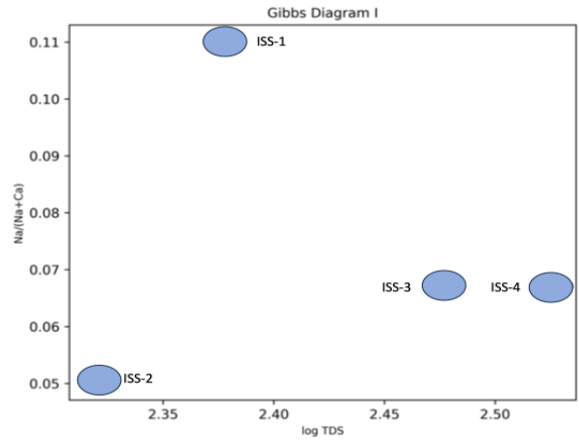


Figure 7 Gibbs diagram (Type I) showing the relationship between $\text{Na}^+ / (\text{Na}^+ + \text{Ca}^{2+})$ ratios and TDS for the Süleymanlı–İlca springs.

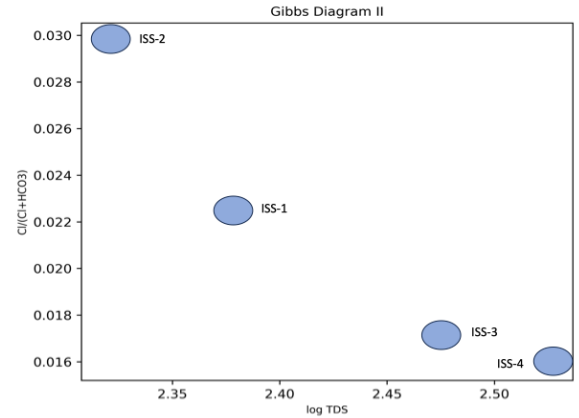


Figure 8. Gibbs diagram (Type II) showing $\text{Cl}^- / (\text{Cl}^- + \text{HCO}_3^-)$ ratios versus TDS.

USSL diagram results show that all Süleymanlı–İlca springs are classified as C1–S1 (U.S. Salinity Laboratory Staff, 1954) (Figure 9). This class is

characterized by the presence of both low electrical conductivity (EC) and low sodium adsorption rate (SAR). Since all waters carry a low risk of salinity and low risk of alkalinity, they can be safely used in both permeable and moderately heavy soils. This indicates that the spring waters do not create any restrictions in terms of agricultural irrigation and that the carbonate-based aquifers in the region offer good quality water with low salinity pressure.

Wilcox diagram results show that all of the Süleymanlı–İlca springs have very low Na% and very low EC values (Wilcox, 1955) (Figure 10). All samples fall within the “Very well–Well” use zone of the diagram, which is consistent with near-perfect water quality for irrigation, negligible risk of sodicity, and no destructive effect on soil structure. Furthermore, none of the samples approach the “suspicious” or “not used” zones; this indicates that the springs are completely safe in terms of salinity and alkalinity risks.

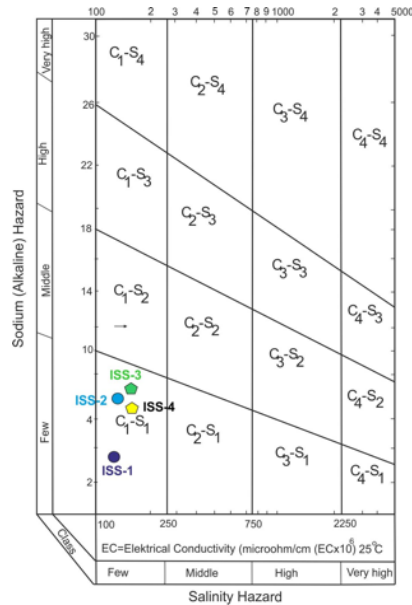


Figure 9 USSS (United States Salinity Laboratory) classification diagram showing the irrigation-water suitability of the Süleymanlı–İlca spring samples (ISS-1, ISS-2, ISS-3, ISS-4). All samples fall within the C1–S1 field, indicating low salinity hazard and low sodium (alkalinity) hazard. This distribution demonstrates that the waters pose no restriction for irrigation and are suitable for most soil types and salt-sensitive crops.

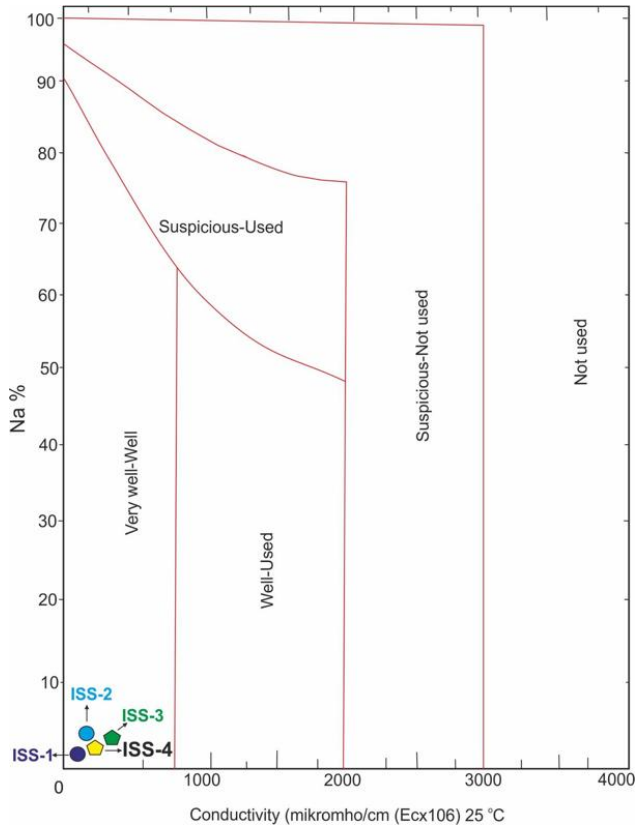


Figure 10 Wilcox diagram illustrating the irrigation-water suitability of the Süleymanlı–İlica spring samples (ISS-1, ISS-2, ISS-3, ISS-4). All waters plot within the “Very well–Well” usable field, characterized by low sodium percentage (Na%) and low electrical conductivity (EC). The distribution shows that the spring waters are of excellent to good quality for irrigation and pose no significant sodicity or salinity hazard.

Chloro-alkaline indices (CAI_1 and CAI_2) were negative in all samples. Negative CAI values indicate that reverse ion exchange processes are dominant between the groundwater and the aquifer matrix, meaning that Ca^{2+} and Mg^{2+} tend to dissolve into the solution, while Na^+ binds to the rock structure (Srinivasamoorthy et al., 2014). This situation observed in the Süleymanlı–İlica waters points to a geochemical regime where carbonate rock dissolution and ion exchange with clay minerals operate together, but which changes the overall composition in favor of Ca–Mg-rich waters rather than Na-rich waters. As shown in Table 2, negative CAI_1 and CAI_2 values suggest active reverse ion exchange in the aquifer. Negative CAI values indicate reverse ion exchange processes where Ca^{2+} and Mg^{2+} replace Na^+ in the aquifer matrix.

Table 2. Chloro-alkaline index (CAI) values for the Süleymanlı–Ilica spring waters.

Spring	CAI1	CAI2
ISS-1	-3.133	-0.063
ISS-2	-0.809	-0.021
ISS-3	-3.189	-0.048
ISS-4	-2.548	-0.038

The springs form a single main cluster, confirming hydrochemical homogeneity within the carbonate aquifer. Hierarchical clustering analysis based on the Ward method using standardized master ion data showed that the samples belonged to a single cluster with very low Euclidean distances. In the dendrogram, ISS-3 and ISS-4 are seen to be slightly closer together due to their higher Ca–HCO₃ contents and EC values, while ISS-1 and ISS-2 form a second subgroup with relatively lower mineralization levels. However, the fact that all samples are located in the same master cluster supports the conclusion that there is a single dominant aquifer type in the Süleymanlı–Ilica system. As seen in the dendrogram (Figure 11), all sources cluster together, supporting the interpretation of a common recharge and geochemical evolution pathway.

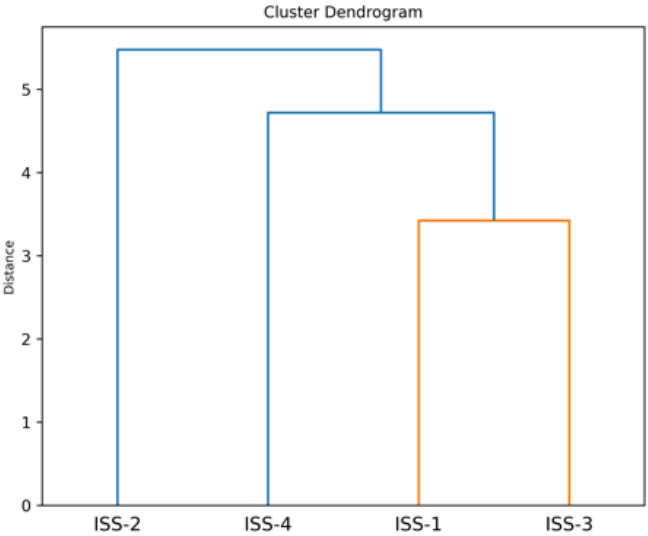


Figure 11 Hierarchical cluster dendrogram based on major ion concentrations.

The Schoeller semi-logarithmic diagram (Figure 12) reveals a remarkably similar ionic pattern for all four springs (ISS-1 to ISS-4), indicating a common hydrogeochemical origin. The waters are characterized by high HCO_3^- and Ca^{2+} concentrations, typical of a Ca–Mg– HCO_3 facies derived from carbonate rock dissolution. Very low Cl^- and $\text{Na}^+ + \text{K}^+$ values suggest minimal salinity and limited cation exchange, supporting the interpretation of fresh, young groundwater with short residence times. Slightly elevated Ca^{2+} and HCO_3^- levels in ISS-4 may indicate a longer water–rock interaction pathway.

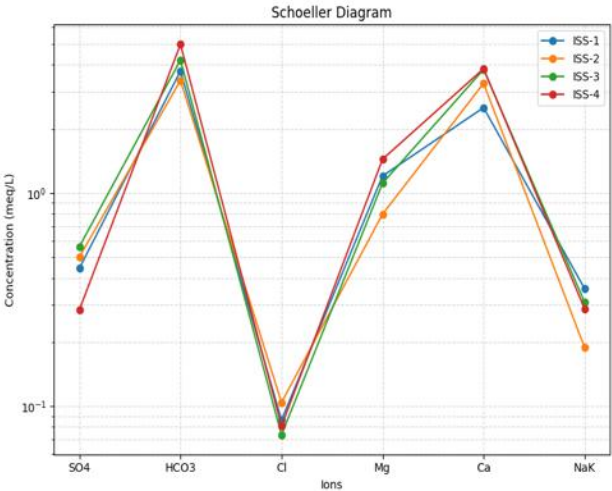


Figure 12 Schoeller diagram of the Süleymanlı–İlca spring waters (ISS-1 to ISS-4). The samples exhibit a characteristic Ca– HCO_3 hydrochemical facies with low Cl^- and Na^+ contents, indicating dominance of carbonate rock dissolution and limited influence of evaporite or anthropogenic inputs. The similar line patterns suggest that all four springs are fed by the same hydrogeological system.

Environmental Isotope Hydrology of the Süleymanlı–İlca Springs

Environmental isotopes provide essential information on groundwater origin, recharge mechanisms, flow dynamics, and residence times (Jeelani et al., 2021). In the Süleymanlı–İlca (Kahramanmaraş) region, stable isotopes of oxygen ($\delta^{18}\text{O}$) and hydrogen ($\delta^2\text{H}$), together with radiogenic tritium (^3H), were analyzed to characterize recharge elevations, meteoric inputs, and water–rock interaction processes (Figure 13). Interpretations were made with reference to the Global Meteoric Water Line (GMWL; Craig, 1961) and the Mediterranean Meteoric Water Line (MMWL), and seasonal sampling during both wet and dry periods enabled evaluation of hydrological variability over one annual cycle.

Isotope data clearly demonstrate that the Süleymanlı–Ilıca springs are of meteoric origin. The distribution of $\delta^{18}\text{O}$ and $\delta^2\text{H}$ values shows that all samples plot very close to both the GMWL and MMWL, indicating negligible evaporation prior to recharge and direct infiltration of precipitation into the aquifer. Tritium concentrations ($\approx 3.7\text{--}5.7$ TU) confirm that the waters are young, reflecting recharge from recent rainfall and short groundwater flow paths. Their position within the “young water” domain in the tritium–Cl diagram further indicates an absence of deep circulation or mixing with older groundwater. Likewise, the tritium–EC relationship demonstrates short residence times, limited mineral dissolution, and minimal geochemical evolution during subsurface transport. These combined results show that the Süleymanlı–Ilıca springs discharge rapidly recharged, shallow-circulation groundwater systems that preserve the isotopic signature of local precipitation.

The isotope compositions exhibit minimal seasonal variation, indicating stable and consistent meteoric recharge conditions (Table 3). For example, $\delta^{18}\text{O}$ at ISS-1 varies only between -9.14‰ (dry season) and -9.22‰ (wet season), while $\delta^2\text{H}$ remains within a narrow range (-51.88‰ to -51.78‰). Similar patterns observed in ISS-2, ISS-3, and ISS-4 imply that all springs respond to the same recharge regime. Their consistent placement between the GMWL and MMWL confirms that recharge occurs rapidly through fractures, faults, and karstic openings characteristic of the surrounding limestone formations.

Table 3. Stable isotope ($\delta^{18}\text{O}$ – $\delta^2\text{H}$), d-excess and tritium values of the Süleymanlı–Ilıca springs.

Spring	Season	$\delta^{18}\text{O}$ (‰ SMOW)	$\delta^2\text{H}$ (‰ SMOW)	d-excess (‰)	Tritium (TU)
ISS-1	Dry	–9.14	–51.88	21.24	5.53
ISS-1	Wet	–9.22	–51.78	21.98	3.69
ISS-2	Dry	–9.00	–52.09	19.91	5.66
ISS-2	Wet	–9.07	–52.31	20.25	3.75
ISS-3	Dry	–8.92	–51.05	20.31	4.42
ISS-3	Wet	–9.15	–52.14	21.06	4.09
ISS-4	Dry	–8.75	–50.02	19.98	5.35
ISS-4	Wet	–8.84	–50.73	19.99	4.33

Deuterium excess (d-excess), calculated as $d = \delta^2\text{H} - 8 \cdot \delta^{18}\text{O}$ (Dansgaard, 1964), ranges from 19.91‰ (ISS-2) to 21.98‰ (ISS-1). These relatively high values reflect a continental moisture source and further confirm that recharge occurs without significant evaporative enrichment. Tritium, with a half-life of 12.43 years (Clark & Fritz, 1997), is highly effective for estimating groundwater age. The measured tritium values (3.69–5.66 TU) correspond to recharge ages

generally <50 years, confirming young groundwater with short residence times. In the $\delta^{18}\text{O}$ – ^3H and $\delta^2\text{H}$ – ^3H diagrams (Figure 13b,d), all samples fall within the young-water domain, and their low chloride contents combined with high tritium values (Figure 13c) indicate minimal mixing with deeper, older groundwater. Among the four springs, ISS-1 exhibits slightly lower tritium values, suggesting a marginally longer circulation pathway, though still within the young groundwater range.

Collectively, the isotope data depict a hydrogeological system recharged directly by meteoric precipitation infiltrating through fractured and karstified carbonate rocks around the study area. The close similarity of isotope signatures among all springs indicates hydraulic connectivity, rapid subsurface flow, and minimal isotopic modification during groundwater transit. Tritium and EC data confirm short residence times and weak water–rock interaction. Geologically, the aquifer is associated with the Early–Middle Miocene Kilisecik Formation—comprising shale, claystone, limestone, and sandstone—and structurally influenced by the Berit Nappe. High-permeability fractured limestone units enhance aquifer recharge and promote rapid discharge of low-mineralization cold springs.

In summary, the Süleymanlı–İlica spring system is fed by young, rapidly recharged meteoric waters circulating through structurally controlled carbonate aquifers. The springs show minimal evaporative enrichment, limited geochemical evolution, and efficient flow pathways, consistent with regional geological conditions and supported by both hydrochemical and isotopic evidence.

Deuterium excess (d-excess), calculated as $d = \delta^2\text{H} - 8 \cdot \delta^{18}\text{O}$ (Dansgaard, 1964), ranges from 19.91‰ (ISS-2) to 21.98‰ (ISS-1). These elevated d-excess values suggest a continental moisture source and support the interpretation that recharge occurs without significant evaporative enrichment. Tritium (^3H), with a half-life of 12.43 years (Clark & Fritz, 1997), was employed to estimate groundwater age. Measured Tritium concentrations (3.69–5.66 TU) indicate that the waters represent young groundwater with short residence times (< 50 years). In the $\delta^{18}\text{O}$ – ^3H and $\delta^2\text{H}$ – ^3H diagrams (Figure 13b,d), all samples fall within the young-water domain. The Cl – ^3H relationship (Figure 13c) reveals low salinity and high tritium values, implying limited mixing with deeper or older groundwater. Among the springs, ISS-1 shows marginally lower tritium values, suggesting a slightly longer but still young circulation pathway.

The combined isotope systematics illustrate a hydrogeological system recharged directly by meteoric precipitation infiltrating through fractured

limestone and karstic structures surrounding the study area. The isotopic uniformity among all springs denotes a hydraulically connected system characterized by rapid subsurface flow and minimal isotopic modification during transit. Tritium and electrical conductivity data further confirm short residence times and weak water–rock interaction. Geologically, the aquifer is associated with the Early–Middle Miocene Kilisecik Formation, consisting of shale, claystone, limestone, and sandstone, and structurally influenced by the Berit Nappe. High-permeability fractured limestone units enhance groundwater storage and rapid discharge.

Overall, the environmental isotope data demonstrate that the Süleymanlı–Ilıca spring system is fed by young, rapidly recharged meteoric waters circulating through structurally controlled carbonate aquifers. The springs exhibit minimal evaporative enrichment, limited geochemical evolution, and rapid flow pathways, consistent with hydrochemical evidence and regional geological structure.

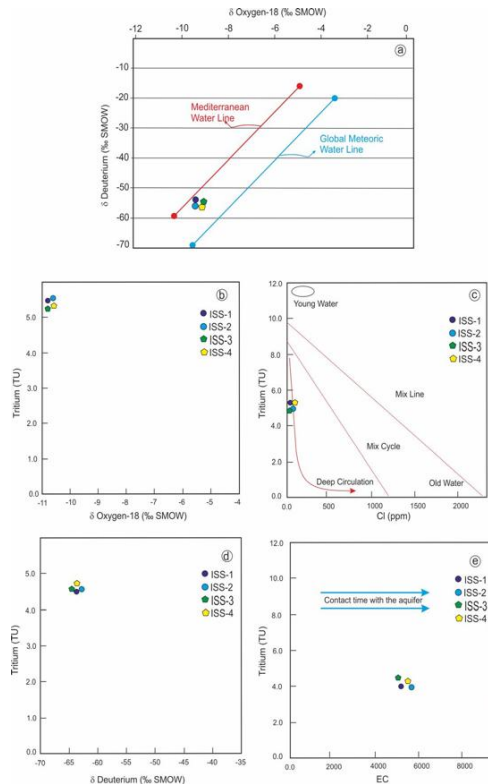


Figure 13 Stable isotope ($\delta^{18}\text{O}$ – $\delta^2\text{H}$) and tritium (^3H) systematics of the Süleymanlı–Ilıca spring waters.

CONCLUSION

The hydrogeochemical and environmental isotope investigations conducted on the Süleymanlı–Ilica spring system demonstrate that the aquifer is a young, rapidly recharged, carbonate-dominated groundwater body with minimal geochemical evolution and exceptionally good water quality. Major-ion chemistry consistently identifies a Ca–Mg–HCO₃ facies, controlled primarily by dissolution of limestone and dolomite units belonging to the Kilisecik Formation and structurally influenced by the Berit Nappe. Low electrical conductivity (EC), low salinity, and near-complete absence of trace metals (<DL) collectively indicate a pristine aquifer with virtually no anthropogenic impact.

Hydrochemical indices—including Gibbs plots, USSS and Wilcox irrigation classifications, chloro-alkaline indices (CAI₁, CAI₂), and Schoeller patterns—confirm that water chemistry is governed mainly by water–rock interaction in carbonate lithologies, with negligible evaporative concentration or cation-exchange pathways that would enrich Na⁺ or Cl[–]. All springs fall within the C1–S1 and “Very well–Well” irrigation suitability classes, indicating excellent usability for agricultural purposes.

Stable isotope compositions ($\delta^{18}\text{O}$, $\delta^2\text{H}$) plot along the Global Meteoric Water Line (GMWL) and Mediterranean Water Line (MWL), revealing a clear meteoric origin and absence of evaporative modification prior to recharge. Tritium concentrations (3.69–5.66 TU) demonstrate that the springs represent young groundwater with residence times <50 years, recharged directly by modern precipitation infiltrating through fractured and karstified carbonate units. The uniformity of isotopic signatures among the four springs confirms hydraulic connectivity and a shared recharge mechanism.

Overall, the geochemical and isotopic evidence suggests that the Süleymanlı–Ilica spring system is a dynamically recharged, structurally controlled, high-permeability aquifer with short groundwater flow paths and minimal water–rock interaction. These findings are consistent with the geological framework dominated by Miocene carbonate formations and tectonically fractured units. The exceptionally low mineralization, low nitrate concentrations, absence of heavy metals, and favorable pH conditions confirm that the spring waters are of excellent drinking-water quality and pose no geochemical risk to public health. All measured parameters comply with international drinking-water standards, and no health-based exceedances were identified, confirming that the springs pose no toxicological risk. The study contributes both to the regional hydrogeological understanding of the Kahramanmaraş Basin and to broader conceptual models of young meteoric aquifers in fractured carbonate terrains.

REFERENCES

- Baydar, O., & Yergök, A. F. (1996). Güneydoğu Anadolu–Kenar Kıvrım Kuşağı–Amanos Dağları kuzeyi ve Doğu Torosların jeolojisi. Jeoloji Etütleri Dairesi, Ankara.
- Clark, I., & Fritz, P. (1997). *Environmental isotopes in hydrogeology*. CRC Press.
- Craig, H. (1961). Isotopic variations in meteoric waters. *Science*, 133(3465), 1833–1834. <https://doi.org/10.1126/science.133.3465.1833>
- Dao, P.U., Heuzard, A.G., Le, T.X.H., Zhao, J., Yin, R., Shang, C., Fan, C. (2024). The impacts of climate change on groundwater quality: A review. *Science of The Total Environment*, 912, 169241. <https://doi.org/10.1016/j.scitotenv.2023.169241>.
- Darbaş, G., & Gül, M. (2006). Alacık Formasyonuna (Kahramanmaraş Havzası) ait kömürlü seviyelerin paleontolojisi. *KSÜ Fen ve Mühendislik Dergisi*, 9(2), 71–81.
- Darbaş, G., & Gül, M. (2017). New observations of planktonic foraminifera from the Burdigalian-Langhian deposits (Kilisecik Formation) of Kahramanmaraş/Turkey. *Journal of African Earth Sciences*, 134, 457-465. <https://doi.org/10.1016/j.jafrearsci.2017.07.010>.
- Darbaş, G., Bayrak, F., Gürel, A. (2021). Usability value of the Yenicekale formation exposure around the Pazarcık (Kahramanmaraş) as a cement raw material. *Bull. Min. Res. Exp.* 166, 53-69.
- Deng, J., Yang, G., Yan, X., Du, J., Tang, Q., Yu, C., Pu, S. (2024). Quality evaluation and health risk assessment of karst groundwater in Southwest China. *Science of The Total Environment*, 946, 174371. <https://doi.org/10.1016/j.scitotenv.2024.174371>.
- Drever, J. I. (1997). *The Geochemistry of Natural Waters* (3rd ed.). Prentice Hall.
- Erdoğan, T. (1975). VI. Bölge Gölbaşı dolayının jeolojisi (TPAO Rapor No. 929). Ankara: Türkiye Petrolleri Anonim Ortaklığı. (Yayımlanmamış rapor).
- European Union. (2020). Directive (EU) 2020/2184 of the European Parliament and of the Council of 16 December 2020 on the quality of water intended for human consumption. *Official Journal of the European Union*, L 435, 1–62.
- Ford, D. C., & Williams, P. (2007). *Karst Hydrogeology and Geomorphology* (2nd ed.). John Wiley & Sons. <https://doi.org/10.1002/9781118684986>
- Gibbs, R.J. (1970). Mechanisms controlling world water chemistry. *Science*, 170(3962), 1088–1090.

- Gleick, P.H. (1998). Water in Crisis: Paths to Sustainable Water Use. *Ecological Applications*, 8(3), 571-579. [https://doi.org/10.1890/1051-0761\(1998\)008\[0571:WICPTS\]2.0.CO;2](https://doi.org/10.1890/1051-0761(1998)008[0571:WICPTS]2.0.CO;2)
- Gözübol, A.M., & Gürpınar, O. (1980). Kahramanmaraş kuzeyinin jeolojisi ve tektonik evrimi. 5. Türkiye Petrol Kongresi Bildirileri içinde (ss. 21–29). Ankara.
- Gül, M.A. (2000). Kahramanmaraş yöresinin jeolojisi (Doktora tezi). Hacettepe Üniversitesi, Fen Bilimleri Enstitüsü, Ankara.
- Gül, M., Darbaş, G., & Gürbüz, K. (2005). Alacık Formasyonunun (En geç Orta Eosen–Erken Miyosen) K. Maraş Havzası içindeki tektono-stratigrafik konumu. *İstanbul Üniversitesi Mühendislik Fakültesi Yerbilimleri Dergisi*, 18(2), 183–197.
- Jeelani, G., Lone, S.A., Nisa, A. U., Deshpande, R.D., & Padhya, V. (2021). Use of stable water isotopes to identify and estimate the sources of groundwater recharge in an alluvial aquifer of Upper Jhelum Basin (UJB), western Himalayas. *Hydrological Sciences Journal*, 66(16), 2330–2339. <https://doi.org/10.1080/02626667.2021.1985126>
- Kırdan, M. S. (2004). Döngele, Fatmalı ve Pınarbaşı (Kahramanmaraş) civarı Tersiyer istifinin mikropaleontolojik özellikleri ve ortamsal yorumu (Yüksek lisans tezi). Çukurova Üniversitesi, Fen Bilimleri Enstitüsü.
- Kozlu, H. (1997). Doğu Akdeniz bölgesinde yer alan Neojen basenlerinin (İskenderun, Misis–Andırın) tektono-stratigrafi birimleri ve bunların tektonik gelişimi (Doktora tezi). Çukurova Üniversitesi, Fen Bilimleri Enstitüsü, Adana, Türkiye.
- Mohamed, M.M., Parimalarenganayaki, S., Khan, Q., Murad, A. (2021). Review on the use of environmental isotopes for groundwater recharge and evaporation studies in the GCC countries. *Groundwater for Sustainable Development*, 12, 100546. <https://doi.org/10.1016/j.gsd.2021.100546>.
- Mutua, S., Ghysels, G., Anibas, C., Obando, J., Verbeiren, B., Van Griensven, A., Vaessens, A., & Huysmans, M. (2020). Understanding and conceptualization of the hydrogeology and groundwater flow dynamics of the Nyando River Basin in Western Kenya. *Journal of Hydrology: Regional Studies*, 32, 100766. <https://doi.org/10.1016/j.ejrh.2020.100766>
- Önalın, M. (1985). Tersiyer istifinin sedimenter özellikleri ve çökelme ortamları. *İstanbul Üniversitesi Mühendislik Fakültesi Yerbilimleri Dergisi*, 5(1–2), 39–78.
- Perinçek, D., & Kozlu, H. (1984). Stratigraphy and structural relations of the units in the Afşın–Elbistan–Doğanşehir Region, Eastern Taurus. In *Proceedings of*

- the International Symposium on the Geology of the Taurus Belt (1983) (pp. 181–198). Mineral Research and Exploration Institute (MTA), Geological Section.
- Piper, A.M. (1944). A graphic procedure in the geochemical interpretation of water analyses. *Transactions of the American Geophysical Union*, 25(6), 914–928.
- Robertson, A. H. F., Ustaömer, T., Parlak, O., Ünlügenç, U. C., Taşlı, K., & İnan, N. (2006). The Berit transect of the Tauride thrust belt, S Turkey: Late Cretaceous–Early Cenozoic accretionary/collisional processes related to closure of the Southern Neotethys. *Journal of Asian Earth Sciences*, 27(1), 108–145.
- Sinanoğlu, D. & Sarkar, S. (2025). Early-Middle Miocene coralline algal assemblages from the Siirt province, southeastern Türkiye: microtaphofacies analysis and palaeoenvironmental implications. *Lethaia*, 58(2), 1-10.
- Srinivasamoorthy, K., Chidambaram, S., Sarma, V.S., Vasanthavigar, M., Vijayaraghavan, K., Rajivgandhi, R., Anandhan, P., Manivannan, R. (2009). Hydrogeochemical Characterisation of Groundwater in Salem District of Tamilnadu, India. *Research Journal of Environmental and Earth Sciences*, 1(2), 22-33, 2009
- TSMS, Turkish State Meteorological Service (2025). Available at: www.dmi.gov.tr.
- Turkish Standards Institution (TSE). (2005). TS 266: İnsani tüketim amaçlı sular – İçme ve kullanım suyu. Türk Standardları Enstitüsü, Ankara.
- U.S. Salinity Laboratory Staff. (1954). *Diagnosis and Improvement of Saline and Alkali Soils* (USDA Handbook No. 60). U.S. Government Printing Office.
- Varol, L., Kop, A., & Darbaş, G. (2012). Koçlar–Sarigözel (Kahramanmaraş kuzeyi) arasında kalan bölgenin jeolojik özellikleri, Doğu Toroslar. *Kahramanmaraş Sütçü İmam Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi*, 15(1), 43-56.
- Wilcox, L. V. (1955). *Classification and use of irrigation waters*. USDA Circular 969. Washington, D.C.: United States Department of Agriculture.
- World Health Organization. (2022). *Guidelines for drinking-water quality: Fourth edition incorporating the first and second addenda*. Geneva, Switzerland: World Health Organization.
- Yılmaz, Y., Gürpınar, O., & Yiğitbaş, E. (1988). Amanos Dağları ve Maraş dolaylarında Miyosen havzalarının tektonik evrimi. *Türkiye Petrolleri Anonim Ortaklığı Bülteni*, 1(1), 52–72.

CHAPTER 17

Material Handling Concepts

Yusuf Ziya Koç¹

¹ Manager of Ceva Logistics, İstanbul, Türkiye, ORCID ID: /0009-0007-3139-4064

1. Material Handling Concepts

Material handling represents one of the foundational pillars of contemporary industrial operations, serving as a critical link between production stages, storage environments, and distribution networks. Encompassing the coordinated processes of moving, storing, controlling, and protecting materials, it plays an essential role in ensuring that goods flow efficiently and safely within manufacturing plants, warehousing facilities, and broader logistics systems. As global supply chains become increasingly complex and the demand for higher productivity intensifies, the design and optimization of material handling systems have gained strategic importance for organizations seeking to improve operational performance and reduce costs.

This study provides a comprehensive examination of the conceptual framework, structural components, and operational mechanisms that define modern material handling systems. Particular emphasis is placed on understanding the functional principles that guide system design, the technological and mechanical elements that enable material movement, and the ways in which these systems are integrated into large-scale production and supply chain environments. By exploring both theoretical constructs and practical considerations, the article seeks to offer a multidimensional perspective on how material handling supports workflow continuity, enhances process reliability, and contributes to overall organizational efficiency.

To reinforce these theoretical foundations with practical insight, the article concludes with a case study illustrating the operation of a material handling system in a real industrial setting. This case study highlights the interplay between system architecture, equipment selection, and process performance, demonstrating how design decisions directly affect operational outcomes such as throughput, flexibility, and resource utilization. Through this applied example, readers gain a clearer understanding of how material handling principles translate into real-world practice and why effective system planning is vital for achieving sustainable industrial efficiency.

1.1. Basic Mechanism of the warehouse System

In order to understand the mechanisms of a material handling system, various systematic approaches are used to categorize warehouse and handling activities. These classifications are typically based on:

- the functions performed by the warehouse,
- the philosophies guiding the allocation of materials within the facility,

- the fixtures and hardware employed in storage and handling operations,
- the methods by which materials are identified and addressed, and
- the types of storage containment systems utilized.

It is important to note that material handling systems—like all engineered products—can differ significantly in terms of durability, quality, and design characteristics. However, a detailed examination of these product-specific variations lies beyond the scope of this article.

1.2. According to the functions performed by a warehouse

The functions performed in all goods warehousing types - receive, store, pick and ship are identical, so that basic functions listed below:

- Receiving the goods from a source
- Storing the goods
- Picking the goods when they are required
- Shipping the goods to the appropriate customer

The only true distinction among the four are the source from which the goods are received and the user to whom the goods are shipped.

Table 1. Warehouse Types

Source	User	Description
Outside	Inside	Raw material warehouse receives the goods from an outside source, stores the goods, picks the goods and ships the goods to an inside user.
Inside	Outside	Finished goods warehouse receives the goods from an inside source, stores the goods, picks the goods and ships the goods to an outside user.
Outside	Outside	Distribution warehouse receives the goods from an outside source, stores the goods, picks the goods and ships the goods to an outside user. This category also includes logistics service providers (LSPs).
Inside	Inside	In-process Inventory warehouse receives the goods from an inside source, stores the goods, picks the goods and ships the goods to an inside user.

The distinctions among these various types of warehouses are largely determined by differences in their sources, management structures, and intended end users. At a fundamental level, any operation whose core activities include receiving, storing, picking, and shipping can be classified as a warehousing function—regardless of its specific position or role within the broader logistics network of the organization [2].

1.3. According to the philosophies of materials allocated in a warehouse.

A number of warehouse layout philosophies exist to serve as guidelines for the development of an effective warehouse layout. Each warehouse layout alternative should be evaluated against the specific criteria established for each of these warehouse layout philosophies.

1.3.1. Popularity Philosophy: In a typical warehouse environment, it is frequently observed that approximately 85% of total product throughput is generated by only 15% of the stored items. An additional 10% of throughput is attributed to roughly 30% of the items, while the remaining 5% arises from the remaining 55% of the inventory. This distribution demonstrates a characteristic pattern: warehouses generally contain a small proportion of highly active items (A items), a moderate proportion of medium-activity items (B items), and a large proportion of low-activity items (C items).

Within the popularity-based warehouse layout philosophy, the design and organization of the facility should therefore be oriented primarily around the relatively small subset of highly active A items, as these account for the majority of material movement. This principle underscores that items with the highest throughput should be placed in locations that enable efficient handling and minimize travel distances. Accordingly, high-turnover items should be stored in readily accessible areas, using storage methods that promote optimal space utilization.

For example, in bulk storage systems, high-throughput items are best accommodated in deep storage blocks, where their rapid rate of movement reduces the likelihood of honeycombing and supports high-density space utilization. In contrast, placing low-throughput items in such deep bulk areas can result in substantial honeycombing losses, as these slow-moving items impede the use of the surrounding space until they are retrieved.

Thus, achieving optimal warehouse efficiency requires a deliberate alignment between item turnover rates, storage locations, and storage methods, ensuring that

the physical layout of the facility supports the operational realities of item demand and movement [2].

Table 2. Popularity philosophy

Item Type			% in total items	% in attribution
Highly	active	items	20	80
(often called A items)				
Moderately	active	items	30	15
(often called B items)				
Infrequently	active	items	50	5
(often called C items)				

1.3.2. Similarity Philosophy: If items that are frequently received or shipped together, they should be stored in close proximity within the warehouse in order to facilitate operational efficiency and minimize material handling time. This principle is grounded in the concept of functional association, which emphasizes that items exhibiting complementary demand patterns or common logistical flows should be spatially grouped. For instance, in the case of a retail automotive parts distributor, it is highly unlikely that a customer purchasing a spark plug wrench will simultaneously require an exhaust system tailpipe. Conversely, a customer who buys a spark plug wrench is much more likely to also need a condenser, ignition points, and spark plugs. Since these items are typically sold and shipped together, they should be stored within the same designated storage zone. Similarly, components such as tailpipes, mufflers, brackets, and gaskets—which belong to the same functional subsystem—should also be stored in close proximity to streamline order fulfillment and inventory control processes. In some cases, certain items are routinely received together from the same supplier or production line. Grouping these items within the same area not only simplifies receiving operations but also reduces internal transportation and handling effort. Furthermore, similar types of materials generally require comparable storage and handling methods; hence, consolidating them in a shared area enhances both space utilization and operational efficiency. However, an important exception to this principle arises when items are so similar in appearance that their proximity could lead to order picking or shipping errors. For example, two-way and four-way electrical switches may look identical while serving entirely different functions. Storing such items adjacent to each other increases the risk of mispicking, mislabeling, or incorrect shipment, thereby compromising accuracy and customer satisfaction.[2].

1.3.3. Size Philosophy: The size philosophy suggests that heavy, hard-to-handle goods should be stored close to their point of use. The cost of handling these items is usually much greater than that of handling other items. That is an incentive to minimize the distance over which they are handled. In addition, if the ceiling height in the warehouse varies from one to another, the heavy items should be stored in the areas with a low ceiling and the lightweight, easy-to-handle items should be stored in the areas with a high ceiling. Available cubic space in the warehouse should be used in the most effective way while meeting restrictions on floor loading capacities. Lightweight material can be stored at greater heights within typical floor loading capacities than heavy materials can. The size philosophy also asserts that the size of the storage location should fit the size of the materials to be stored.[2].

1.3.4. Oddly Shaped Philosophy: Oddly shaped and crushable items, subject to stacking limitations, will *dictate* special storage methods and layout configurations to effectively use available cubic space. Hazardous materials such as explosive, corrosives and highly flammable chemicals must be stored in accordance with government regulations. Items of high value or items commonly subject to pilferage may require increased security measures such as isolated storage with restricted access. The warehouse layout must be adapted to provide the needed protection. The compatibility of items stored close together must also be examined. Contact between certain individually harmless materials can result in extremely hazardous reactions and/or significant product damage. Specific steps must be taken to separate incompatible materials. Oftentimes, the easiest way to accomplish this objective is through the warehouse layout.[2]

1.3.5. Space Utilization Philosophy: This philosophy can be separated into four areas: conservation of space, limitations on use of space, accessibility of material and orderliness.

1.4. The conservation of space principle asserts that the maximum amount of material should be concentrated within the storage area, the total cubic space available should be effectively used and the potential honeycombing within the storage area should be minimized. Unfortunately, these objectives often conflict. Increased concentration of material will usually cause increased honeycombing allowances. Therefore, determining the proper level of space conservation is a matter of making trade-offs among the objectives that maximize use of space.

1.5. Limitations on use of space must be identified early in the layout planning process. Space requirements for building support columns, trusses, sprinkler system components, heating system components, fire extinguishers and hoses and

emergency exits will affect the suitability of certain storage and handling methods and layout configurations. Floor loading capacities will restrict storage heights and densities.

1.6. The warehouse layout should meet specified objectives for material accessibility. Main travel aisles should be straight and should lead to doors in order to improve maneuverability and reduce travel times. Aisles should be wide enough to permit efficient operations, but they should not waste space. Aisles widths should be tailored to the type of handling equipment using the aisle and the amount of traffic expected.

1.7. The Orderliness Principle The orderliness principle emphasizes that effective warehouse management fundamentally begins with a mindset of maintaining organization and discipline. Proper housekeeping is a prerequisite for efficient material handling and spatial utilization. Aisles should be clearly demarcated using aisle tape or floor markings; otherwise, materials may encroach upon these spaces, thereby reducing accessibility and operational efficiency. Empty or unutilized spaces within a storage area must be identified and corrected promptly, as such inefficiencies directly affect storage capacity. For instance, if a storage section is designed to accommodate five pallets, but improper placement causes one pallet to infringe upon the space designated for an adjacent pallet, an empty or unusable gap will emerge. Consequently, the section will be capable of holding only four pallets instead of five. The lost storage capacity cannot be recovered until the entire area is cleared and reorganized.[2]

1.8. According to the fixture and hardware used in a warehouse.

A wide range of well-established and efficient storage systems are available for warehouse operations; however, only a limited number of properly designed stockrooms or warehouses can function effectively by relying on a single system. The primary objective in developing an efficient stockroom lies in determining which storage hardware should be utilized and the extent to which each system type should be implemented. To ensure the selection of optimum storage equipment, it is essential to identify the functional capabilities required for the intended operations.

Faced with the wide selection of storage hardware currently available, individuals responsible for stockroom design should recognize the advantages and disadvantages of each. Some of the basic storage hardware and major points regarding their usage, are listed below.

- Shelving maximizes space for small-size items and small-quantity items and requires trays to contain material on shelf.
- Mezzanine shelving increases equipment cost, provides ability to manually pick on two levels and requires means to move material up to and down from mezzanine.
- High Rise Shelving – no mezzanine eliminates possibility of manually picking second level, reduces employee fatigue and order picker trucks.
- Floor storage provides best ratio of aisle to storage cube; requires uniform material in a single row and well design loads to avoid tipping and damage to bottom load and allows use of stacking frames for fragile materials.
- Conventional pallet rack provides ability to use cube space, protects individual loads; provides access to individual loads; allows hand stacking of large items and reduces aisles with side load fork trucks.
- Two deep pallet rack provides better ratio of aisle to storage cube than conventional pallet rack, requires same material in both location for maximum labor and equipment utilization.
- Cantilever rack maximizes storage space for long irregular shaped items and for materials that vary in length and increases rack cost.
- Flow rack provides FIFO material movement; increases equipment cost, allows increased order pick productivity and requires uniform material.
- Automated storage and retrieval systems (AS/RS) increases total equipment cost; eliminates labor cost; reduces lighting, heat and insurance costs and allows use of rack supported building which reduces depreciation period.

1.9. According To How Materials Addressed

The material storage system is important from the viewpoints of space utilization, ease of training and error potential. There are three basic material storage location systems: Fixed (sometimes referred as "dedicated" or "assigned"), random (sometimes referred as "floating") and combination.

1.9.1. Fixed Location Storage: Each individual stock keeping unit will always be stored in a specific storage location. No other stock keeping unit may be stored in that location, even though that location may be empty. If fixed location storage

is used, a given stock keeping unit must be assigned sufficient space to store the maximum amount of the stock keeping unit that will ever be on hand at any one time. Fixed location storage trades efficiency in use of space for easy accessibility to material. This system makes it easy for store personnel both to store and to find specific materials, with relatively simple type of addressing system. It also reduces paperwork and sources of error. Fixed storage's main disadvantages are an inability to maximize space utilization (empty spots remain empty) and a lack of flexibility when expansion of space is required for an item.[1]

1.9.2. Random Location Storage: Any stock keeping unit may be assigned to any available storage location. A stock keeping unit in location A one month might be stored in location B the following month and a different stock keeping unit stored in location A. For random location storage the quantity of items on hand at any time will be the average amount of each stock keeping unit. In other words, when the inventory level of one item is above average, another item will likely have an inventory level that is below average; the sum of the two will be close to the average. Random location storage trades accessibility to material for efficiency in use of space. Advantages of random storage are that it provides FIFO (First In First Out) material rotation and that there is no limit to its flexibility in expanding or contracting space for an item. The biggest advantage to random storage is maximum space utilization, which is especially important with expensive equipment, such as automated storage retrieval systems. The disadvantages of random storage are: the paperwork system and employee input must be accurate or there will be more errors, more information must be recorded on every transaction and there can be a lag in information for order pickers when new material comes in and located.[1]

1.9.3. Hybrid Location Storage: It is a combination of the two mentioned above. A grocery store is an excellent example of hybrid location storage. Fixed location storage is used in the front room of a grocery store where the consumers shop. Pickles are assigned a fixed location and only pickles will be stored in that location. Pickles will not be found in any other location in the front room of the grocery store. In the back room, or storeroom, of the grocery store, however, the excess, or overstock, merchandise is usually stored randomly. Pickles may be found in one location one week and a different location the next week. Because hybrid location storage is based on a mixture of fixed location storage, its planned inventory level falls between the fixed location quantity and random location quantity. At what point between the fixed location and random location quantity the planned inventory level falls is dependent on the percentage of inventory to be assigned fixed locations. Hybrid location storage trades materials handling

efficiency for middle of the road efficiency in use of space and accessibility to material.

Compared to the use of space by random location storage for the same material, fixed location storage generally require 65 to 85 % more space. With the escalating costs of money, land and construction, few firms can afford to build a fixed location storage warehouse, which would be 75 % larger than that required for random location storage. The expense of developing and maintaining an effective material locator system for random location storage - when compared with these costs - is easily justified. Consequently, one should always carefully evaluate random location storage before deciding to use fixed location storage. Rarely will the gains in accessibility to material made by fixed location storage be enough to offset its high space costs. Occasionally, however, efficient use of space is not a critical factor, so fixed location storage is preferred. For example, when the items to be stored are extremely small and/or extremely valuable, accessibility to them and accountability for them may be all-important. Few jeweler care about the use of space when they are storing diamond rings.

1.10. According To the Storage Containment System

The storage containment system can be either closed or open. In either case, there are trade-offs.

1.10.1. Closed containment system: In a closed containment system, materials are physically controlled in a warehouse. The warehouse should be strictly controlled. The closed system provides the strong security and the greatest degree of data integrity but it does require fixed warehouses and additional material handling.[1].

1.10.2. Open containment system: Open containment system sometimes referred to as four wall inventory, there is no warehouse. The big advantage to the open system is that material is stored close to the point of use. There is no restriction, in terms of fencing or doors, to obtaining material. In addition, there is a minimum amount of materials handling in moving material to the work area and in worker access. The major disadvantage of an open system is that there is virtually no control of material.

2. Case Study

The categories presented are discussed using a real industrial example: Hyper Market. There are two major blocks in the building. The first one is retailer room where customers are walking and buying goods. Second one is retailer room

where all goods are stored. Now, according to the criteria that we mentioned above, we analyze this company:

According to the functions performed by a warehouse: Storeroom receives the goods from an outside source (i.e. suppliers), stores the goods, picks the goods and ships the goods to an inside user (i.e. retailer room). So that, it is Raw materials Storeroom. Retailer room receives the goods from an inside source (i.e. storeroom), stores the goods, picks the goods and ships the goods to an outside user (i.e. customers). So that, it is Finished Goods Warehouse.

According to the philosophies of materials allocated in a warehouse: The storeroom suggest popularity philosophy that the storeroom should be planned around the small number of high turnover items that constitute the great majority of the activity in the warehouse. The retail room suggest similarity philosophy that the retail room should be planned around items that are commonly received and/or shipped together should be stored together. For example, tea and sugar; tissue and soaps and so on are stored in the same location.

According to the fixture and hardware used in a warehouse: Conventional pallet rack is used in storeroom. All corridors are arranged for forklifts in storeroom. Standard shelving in customer height is available in retailer room. All corridors are enough comfortable for arranged customer to walk and material handling is manipulated for two wheelers and pallet jacks.

According to how materials addressed: Fixed location storage is used in the retail room. Materials are assigned a fixed location and only pickles will be stored in that location. Materials will not be found in any other location in the retail room. For example, all frozen foods, cleaning tissues, detergent agents, and so on are stored in same shelves. In the storeroom, however, all goods is usually stored randomly. So that, it is random location storage. According to the storage containment system: Since the storeroom should be strictly controlled, closed containment system is applied. On the contrary, open containment system is applied in retail room, except some valuable materials such as electronic devices. Customers are free to move among the shelves, pick and put their shopping cars.

The result of the analysis is summarized in Table 3.

Table 3. The result of the analysis

Criteria	Stock Room	Retail Room
According to the functions performed by a warehouse	Raw materials Storeroom	Finished Goods Warehouse
According to the philosophies of materials allocated in a warehouse.	Popularity	Similarity
According to the fixture and hardware used in a warehouse.	Conventional pallet	Standard shelving
According to how materials addressed.	Random	Fixed
According to the storage containment system.	Closed	Open

3. Referances

- [1] MAGED E.L. and AMOS J.M.; 1995; "Total Material Management"; Chapman & Hall; 2nd Edition
- [2] Chapter 69 Storage and Warehousing; JEERY D.SMITH; Handbook of Industrial Engineering Edited by GAVRIEL GALVENDY – Institute of Industrial Engineering; 2nd Edition
- [3] KOÇ Y.Z.; 1998; "Spare Part Warehousing"; Marmara University, Master

CHAPTER 18

Renewable Energy Applications in University Campuses: Reducing Carbon Emissions Based On Solar Energy Systems (SES) at Mersin University

**Hüdaverdi Arslan¹ & Sena Nur Demirel²
Anıl Erdağ Nomer³**

¹ Asst. Prof. Dr., Mersin University, Faculty of Engineering, Department of Environmental, ORCID:0000-0002-3053-6944

² MSc Student., Uludağ University, Faculty of Engineering, Department of Environmental, ORCID:0000-0003-1956-8846

³ PhD Candidate., Mersin University, Department of Nanotechnology and Advanced Materials, ORCID:0000-0002-8997-6936

INTRODUCTION

In the 18th and early 19th centuries, the development of the Industrial Revolution created a need for energy sources, which increased interest in fossil fuels. However, the increase in fossil fuel use has created global carbon emissions and led to climate change. With the emergence of climate change, efforts to reduce carbon emissions have begun. (Arena & Brooks, 2025). The Intergovernmental Panel on Climate Change (IPCC) defines climate change as “a change in climate conditions that results in long-term changes in climate characteristics.” (Eggleston, Buendia, Miwa, Ngara, & Tanabe, 2006). In addition, the United Nations Framework Convention on Climate Change (UNFCCC) has emphasized that the primary cause of climate change is human activity (UN, 1992).

Both definitions of climate change indicate that it is based on direct or indirect human influence. The primary reason for climate change being human-based is carbon emissions. Greenhouse gases released into the atmosphere by humans create carbon emissions and give rise to the concept of carbon footprint that is currently on our agenda (Sreng & Gümrükçüoğlu Yiğit, 2017). With the introduction of the concept of carbon footprint into our lives, the shift towards renewable energy has increased, and the goal is to reduce fossil fuel consumption through the use of renewable energy. In line with these goals, many countries are engaged in strategic efforts to increase their renewable energy capacity within the framework of carbon neutrality.

By utilizing clean energy sources such as solar and wind energy, which are among the foremost renewable energy sources, it enables the reduction of currently used fossil fuels and the achievement of sustainability by creating low carbon emissions (Alagoz & Alghawi, 2023). Solar energy systems are among the technologies with the least environmental impact when compared to fossil fuel-based energy production methods. These systems, which do not directly produce greenhouse gas emissions during electricity generation, play a crucial role in reducing carbon dioxide emissions. Studies have shown that the use of solar energy results in a significantly lower carbon footprint throughout its life cycle compared to fossil fuels (Tsoutsos, Frantzeskakis, & Gekas, 2005).

As in countries around the world, the use of clean energy sources is at the forefront of sustainability efforts across Turkey due to their low carbon emissions. In this context, it is evident that the energy production potential of renewable energy sources such as solar, wind, biomass, and geothermal is increasing day by day. A large portion of the budget allocated to strengthen Turkey's energy supply potential and reduce external dependence covers energy investments (Keleş & Bilgen, 2012). Higher education institutions play a significant role in advancing sustainability and reducing carbon emissions through campus-based clean energy initiatives (Lozano et al., 2015). In line with carbon-neutral mobility, in addition to sector-specific clean energy initiatives, universities are also undertaking numerous projects within their own institutions, both in pursuit of sustainability and to reduce carbon emissions (Yavuz, Kara, & Yanıktepe, 2023).

Universities leading the way in sustainability and carbon neutrality goals are meeting part of their energy needs from renewable sources through solar energy systems installed on campus, while also reducing their emissions into the atmosphere and taking the lead in positively impacting their carbon footprint (Ramos et al., 2015). Such initiatives not only enable universities to fulfill their environmental responsibilities but also contribute to setting an example for sustainable energy policies within society (Singh, Meena, & Khandelwal, 2023). Solar energy systems installed on university campuses offer significant potential for sustainable energy production and carbon emission reduction. Comparing electricity production from on-campus clean energy with consumption and calculating emission reductions in terms of carbon dioxide equivalent (ton CO_{2e}) provides a quantitative assessment of the environmental impact of clean energy applications (Peng, Lu, & Yang, 2013).

In this context, it highlights the strategic importance of solar energy in helping universities achieve their carbon-neutral campus goals. Solar energy systems offer not only environmental benefits but also economic and strategic advantages. By significantly reducing the high electricity bills generated by renewable energy

technologies, photovoltaic systems have become a competitive energy source. Investments in solar energy systems strengthen energy supply security by reducing dependence on energy imports and are seen to contribute to economic sustainability in the longer term (Bazalian et al., 2013). Thus, universities are among the most economical investments in achieving sustainability goals.

SCOPE AND METHODOLOGY

Definition of Scope

This study focuses on Mersin University as a university that pioneers sustainability goals and carbon-neutral mobility. Mersin University began operations on November 10, 1992, and since its establishment, it has grown to include 19 faculties, 6 colleges, 11 vocational schools, 5 institutes, and 40 research centers. It has a total of 9 campuses: Çiftlikköy, Yenişehir, Erdemli, Silifke, Tece, Mut, Anamur, Gülnar, and Tarsus. This study will focus on the Çiftlikköy Campus, where solar energy systems are installed, and the campus has an approximate area of 3.628.292,42 m² (Mersin Üniversitesi, 2024).



Figure 1. Mersin University Çiftlikköy Campus

Campus-Scale Solar Energy Systems

Five different solar energy systems have been installed at five different locations on the Mersin University Çiftlikköy Campus with the aim of sustainable energy production and reducing carbon emissions. These installations are divided into two main areas within the campus. The first area is the Hospital SES area, where the Mersin University Hospital is located and which has the greatest energy needs. Three locations are associated with this area: the Large Parking Lot SES, the Heating Center SES, and the Faculty Parking Lot SES. The other area is the University SES area, which was created to meet the energy needs of the rest of the campus. Two locations are associated with this area: the Faculty of Architecture SES and the Faculty of Medicine Dean's Office SES area.

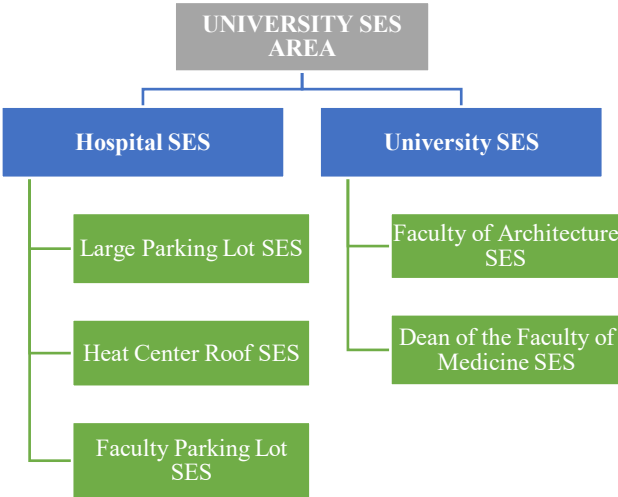


Figure 2. Mersin University Council SES Areas

Carbon Emission Calculation Methodology

Thanks to the solar energy systems (SES) installed on campus, the amount of energy produced on campus will be compared with the amount of energy consumed on campus in 2025, and the carbon emissions they generate will be examined in terms of carbon dioxide equivalent (ton CO_{2e}). Based on the data obtained, the amount of emissions generated by the electricity consumed before the installation of solar energy systems will be determined, as well as the percentage reduction in emissions achieved by the amount of electricity generated after the solar energy systems were put into operation. A quantitative comparison of the current situations will be made. As a calculation methodology, the Intergovernmental Panel on Climate Change (IPCC) guidelines do not provide a

fixed emission factor for emissions generated from electricity consumption. It is recommended that country-specific grid emission factors be used to calculate emissions. (Jamsranjav, 2023).

Carbon intensity from electricity generation varies significantly depending on countries' energy production. Based on the data examined, Turkey's average emission factor associated with electricity generation is approximately 0,442 kg CO₂ kWh⁻¹. This value is quite close to the global average of approximately 0,436 kg CO₂kWh⁻¹, indicating that Turkey's electricity generation system is heavily reliant on fossil fuels (T.C. Enerji ve Tabii Kaynaklar Bakanlığı, 2023).

In countries where coal is more intensively used in electricity generation, it is noteworthy that emission factor values are significantly higher. In this context, when looking at the carbon intensity of electricity consumption in China, it ranges from approximately 0,580–0,650 kg CO₂kWh⁻¹, while in India, this value ranges from 0,700–0,800 kg CO₂kWh⁻¹. These high values stem from coal being the primary energy source for electricity generation in both countries. Lower emission factors are observed in countries where fossil fuel use is gradually being reduced or where different sources are balanced in the production mix.

In the United States, the emission factor associated with electricity consumption is approximately 0,350–0,450 kg CO₂kWh⁻¹ reflecting a consumption structure that uses both coal and natural gas. In Germany, the consumption-based emission factor is approximately 0,300–0,400 kg CO₂kWh⁻¹ indicating that coal use continues despite the increased use of renewable energy sources. In countries with low carbon formation, the emission factor associated with electricity consumption is observed to be at low levels. In France, the emission factor is approximately 0,050–0,100 kg CO₂kWh⁻¹, which stems from the country's nuclear energy-heavy electricity consumption structure. Similarly, in Sweden, the emission factor corresponding to electricity consumption is quite low, at approximately 0,020–0,050 kg CO₂kWh⁻¹, reflecting the predominance of hydroelectric and nuclear energy. (International Energy Agency (IEA), 2023).

In the UK, the emission factor ranges from approximately 0,200–0,250 kg CO₂kWh⁻¹, indicating an energy transition process based on natural gas consumption alongside the use of renewable energy sources. Although this value is lower than that of coal-heavy systems, it is still considered high compared to countries with completely low-carbon energy systems. Overall, countries' energy policies are directly related to their fossil fuel dependency and the share of renewable and nuclear energy sources in their production mix and consumption. In this context, Turkey's emission factor, which is close to the global average, highlights the need to increase renewable energy investments and raise the share of low-carbon energy sources (Our World In Data, 2024).

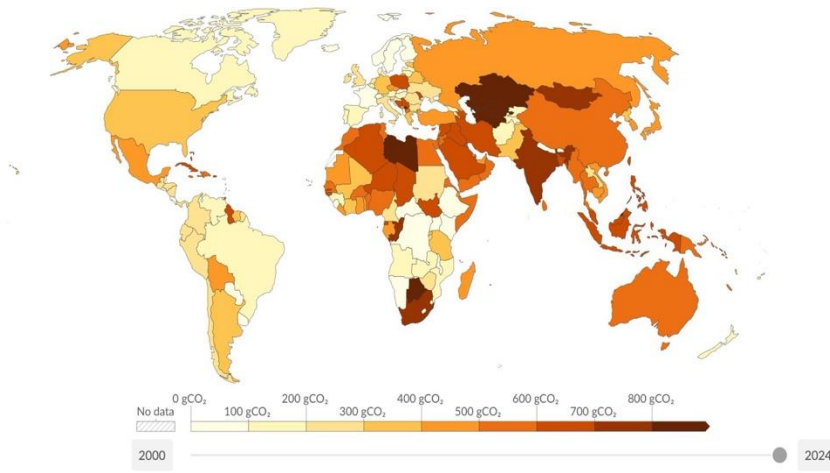


Figure 3. Countries' Emission Factor Levels (Our World In Data, 2024).

Table 1. Carbon Emission Factors by Country

Country / Region	Emission Factors (kg CO ₂ kWh ⁻¹)
Türkiye	~0,442 kg CO ₂ kWh ⁻¹
World Average	~0,436 kg CO ₂ kWh ⁻¹
China	~0,580–0,650 kg CO ₂ kWh ⁻¹
United States	~0,350–0,450 kg CO ₂ kWh ⁻¹
India	~0,700–0,800 kg CO ₂ kWh ⁻¹
Germany	~0,300–0,400 kg CO ₂ kWh ⁻¹
France	~0,050–0,100 kg CO ₂ kWh ⁻¹
Sweden	~0,020–0,050 kg CO ₂ kWh ⁻¹
United Kingdom	~0,200–0,250 kg CO ₂ kWh ⁻¹

Within the Turkish electricity sector, CO₂ emissions per unit of electricity generated are estimated to be approximately 0,5-0,8 kg CO₂kWh⁻¹ (Arı & Köksal, 2011). The emission factor to be used in the calculation methods for determining emissions from electricity consumption in Turkey, including emissions from solar and wind energy use, is specified as 0,6261 tons CO_{2e} MWh⁻¹ in the Turkey National Electricity Grid Information Form published by the Ministry of Energy and Natural Resources (T.C. Enerji ve Tabii Kaynaklar Bakanlığı, 2024). Since data for December 2025 was not available for the calculations, calculations were made based on 11 months of data.

The emission value specified by the Ministry of Energy and Natural Resources of the Republic of Turkey will be used in the calculation method. The calculation method used is the location-based calculation method included in Scope 2 reporting in the GHG Protocol (GreenHouse Gas Protocol, 2015). To find the tonne of carbon dioxide equivalent (tons CO_{2e}) emitted for electricity consumption, the amount of energy currently consumed in kWh will be multiplied by the emission factor. The calculation methodology is described below.

- Emissions calculation in kilograms of CO_{2e}:

$E_{\text{emission}} = \text{Amount of emissions from consumption (kg CO}_{2e}\text{)}$

$E_{\text{KWh}} = \text{Amount of electricity consumed (KWh)}$

$E_{\text{factor}} (\text{kg CO}_{2e} \text{ KWh}^{-1}) = \text{Country-specific emission factor (kg CO}_{2e} \text{ kWh}^{-1})$

$$E_{\text{emission}} (\text{kg CO}_{2e}) = E_{\text{KWh}} \times E_{\text{factor}} (\text{kg CO}_{2e} \text{ KWh}^{-1}) \quad (1)$$

- Emissions calculation in tons of CO_{2e} e:

$E_{\text{emission}} (\text{tons CO}_{2e}) = \text{Amount of emissions resulting from consumption (tons CO}_{2e}\text{)}$

$$E_{\text{emission}} (\text{tons V}) = E_{\text{emission}} (\text{kg CO}_{2e}) / 1000 \quad (2)$$

The emission factor value of 0,6261 ton CO_{2e} MWh⁻¹ will be calculated in kg CO_{2e} kWh⁻¹ and used as 0,6261 kg CO_{2e} kWh⁻¹

Based on the data obtained from the calculation outputs and the comparison to be made, the aim is to clearly demonstrate the importance of clean energy use within universities and to show that it provides significant strategic and environmental benefits in terms of reducing carbon emissions released into the atmosphere, achieving carbon-neutral campus goals, and strengthening energy supply security.

UNIVERSITY SES APPLICATION

Photovoltaic (PV) solar energy systems are systems composed of semiconductor materials that convert sunlight into electrical energy. The basic principle of these systems is that when sunlight falls directly on photovoltaic cells, electrons are activated and electrical energy is produced as direct current (DC). The direct current produced is converted into alternating current (AC) via inverters and made suitable for the grid or local consumption (Messenger & Ventre, 2010).

PV systems hold a significant position among renewable energy technologies due to their low operating costs, modular structures, long service life, and the fact that they do not generate direct greenhouse gas emissions from the moment they are commissioned. With these characteristics, PV systems play an effective role in reducing carbon emissions and combating climate change (Tsoutsos, Frantzeskaki, & Gekas, 2005).

A total of 2,950 photovoltaic panels (PV) were used in the solar energy systems established within the university. Of these panels, 2,405 are located in the Hospital SES area and 545 are located in the University SES area. The photovoltaic panels (PV) used in the installation are Plura brand, measuring $2279 \times 1134 \times 40$ mm and having a nominal power of 550 W. The characteristics of the photovoltaic panels (PV) used are given in Figure 4

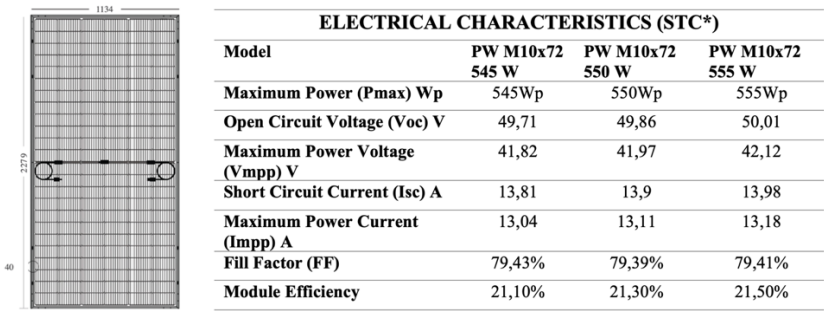


Figure 4. Plura brand photovoltaic (PV) panel features (Plurawatt , 2025)

Regular maintenance and operational activities are required to ensure the long-term performance of photovoltaic solar energy systems. It is recommended that the surface cleaning of panels be performed every 3–6 months, inverters be checked every six months, and electrical connections be inspected annually. It

has been reported that energy production losses can reach 5–20% in systems where regular maintenance is not performed (NREL, 2018).

Hospital Ses Area

Mersin University Hospital is among healthcare facilities with high energy consumption, and therefore offers significant potential for renewable energy applications. In this context, the solar energy systems (SES) established within Mersin University Hospital, located in the Çiftlikköy Campus, aim to meet a portion of the hospital's electricity needs from renewable sources and, consequently, reduce carbon emissions. The hospital, with a capacity of 860 beds, began operations on May 20, 2014, on an area of approximately 50,000 m², and is evaluated within the scope of sustainable energy studies conducted across the campus. (Mersin University Hospital, 2014).



Figure 5. Mersin University Hospital

A solar energy system has been installed at three locations within the hospital's SES area. These locations are designated as the Large Parking Lot SES Area, the Heating Center SES Area, and the Faculty Parking Lot SES Area



Figure 6. Hospital SES Locations

The total electricity consumption of Mersin University Hospital in 2025, the 12-month electricity consumption data provided by the hospital management, and the resulting emission quantities are evaluated in Table 2. Healthcare facilities are among the structures with high energy consumption due to the requirement to provide uninterrupted service, the use of advanced medical equipment, air conditioning, ventilation, and heating systems, and 24-hour work schedules (González, Calcedo, & Salgado, 2018). In this context, Mersin University Hospital stands out as one of the units with the highest energy needs when considering the campus as a whole.

Table 2. Mersin University Hospital 2025 Electricity Consumption and CO₂e Emission Distribution

Period	Amount of Electricity Consumed (kWh)	Emission Quantity (tons CO ₂ e)
January	1.213.890,30	760,02
February	1.444.891,77	904,65
March	838.687,50	525,10
April	846.058,50	529,72
May	1.144.829,70	716,78
June	1.845.018,00	1.155,17
July	2.516.308,20	1.575,46
August	2.748.816,00	1.721,03
September	2.072.252,70	1.297,44
October	1.223.056,80	765,76
November	1.589.380,95	995,11
December	0,00	0,00
General Total	17.483.190,42	10.946,23

When the current distributions are examined, Mersin University Hospital will consume a total of 17.483.190,42 of electrical energy in 2025. This consumption has been calculated under the Location-based scope of the GHG protocol's Scope 2 reporting, revealing that it will generate a total of 0.946,23 tons CO_{2e} in 2025.

Table 3. Large Parking Lot SES 2025 Electricity Production and Corresponding CO_{2e} Emission Quantity

LARGE PARKING LOT SES		
Period	Amount of Electricity Generated (kWh)	Corresponding Emission Quantity (tons CO_{2e})
January	0,00	0,00
February	11.677,58	7,31
March	112.430,13	70,39
April	172.484,73	107,99
May	256.735,99	160,74
June	267.380,08	167,41
July	202.488,08	126,78
August	101.699,11	63,67
September	87.768,01	54,95
October	77.746,73	48,68
November	74.178,51	46,44
December	0,00	0,00
General Total	1.364.588,95	854,37

The large parking lot SES area will generate a total of 1.364.588,95 kWh of electrical energy by 2025. This production corresponds to a total of 854,37 tons CO_{2e} emissions in 2025, based on the GHG Protocol Scope 2 reporting under the Location-based emissions calculation.

Table 4. Heat Center Roof SES 2025 Electricity Production and Corresponding CO2e Emission Amount in Tons

HEAT CENTER ROOF SES		
Period	Amount of Electricity Generated (kWh)	Corresponding Emission Quantity (tons CO ₂ e)
January	0,00	0,00
February	7.406,74	4,64
March	52.336,89	32,77
April	37.152,35	23,26
May	73.765,58	46,18
June	77.039,45	48,23
July	72.474,63	45,38
August	62.954,64	39,42
September	51.126,88	32,01
October	39.397,38	24,67
November	23.192,76	14,52
December	0,00	0,00
General Total	496.847,30	311,08

The heating center roof within the SES area will generate a total of 496.847,30 KWh of electrical energy by 2025. This production corresponds to a total of 311,08 tons CO₂e emissions in 2025, according to the GHG Protocol Scope 2 reporting under the Location-based emissions calculation.

Table 5. Faculty Parking Lot SES 2025 Electricity Production and Corresponding CO2e Emission Amount in Tons

FACULTY PARKİNG LOT SES		
Period	Amount of Electricity Generated (kWh)	Corresponding Emission Quantity (tons CO _{2e})
January	0,00	0,00
February	3.177,15	1,99
March	65.997,64	41,32
April	57.690,34	36,12
May	103.220,91	64,63
June	105.861,36	66,28
July	93.587,60	58,60
August	81.013,66	50,72
September	81.210,89	50,85
October	67.538,65	42,29
November	37.859,49	23,70
December	0,00	0,00
	697.157,69	436,49

The faculty parking lot within the SES area will generate a total of 697,157.69 kWh of electrical energy by 2025. This production corresponds to a total of 436,49 tons CO_{2e} emissions in 2025, based on the GHG Protocol's Scope 2 reporting under the Location-based emissions calculation.

Looking at the three locations within the hospital SES area, a total of 2,558,593.94 kWh of electricity was generated in 2025. The corresponding emissions for the amount of electricity generated were calculated as 1.601,94 tons CO_{2e}. Based on these calculations, an annual recovery rate of 14.63% was achieved.

University Ses Area

Solar energy systems (SES) have been installed at two main locations within the Mersin University Çiftlikköy Campus to meet the campus's energy needs. The first of the installed areas is the Architecture Faculty SES area, located in a central position, while the other location is the Faculty of Medicine Dean's Office SES area, which meets the energy needs of the area independent of Mersin University Hospital.



Figure 7. University SES Locations

The total electricity consumption of Mersin University in 2025, the 12-month electricity consumption data provided by the university administration, and the resulting emission amounts are evaluated in Table 6.

Table 6. Mersin University 2025 Electricity Consumption and ton CO₂e Emission Distribution

Period	Amount of Electricity Consumed (kWh)	Emission Quantity (tons CO ₂ e)
January	197.887,73	123,90
February	258.764,63	162,01
March	183.060,68	114,61
April	162.280,13	101,60
May	213966,9	133,96
June	282.035,25	176,58
July	295.213,28	184,83
August	237.144,21	148,48
September	303.945,08	190,30
October	237.144,21	148,48
November	215.585,64	134,98
December	0,00	0,00
General Total	2.587.027,72	1.619,74

When the current distributions are examined, Mersin University as a whole will consume a total of 2,587,027.72 kWh of electrical energy in 2025. This consumption has been calculated within the Location-based scope of Scope 2

reporting in the GHG protocol, revealing that it will generate a total of 1.619,74 tons CO_{2e} emissions in 2025.

The carbon emissions generated have been targeted for reduction by utilizing clean energy sources. In this context, solar energy systems (SES) have been installed in the university's architecture and medical faculties, and the annual electricity production potential and corresponding emission amounts have been calculated in tons of CO_{2e} and are presented in Table 7 and Table 8. The calculations were made taking into account the maintenance and repair times of photovoltaic (PV) solar energy systems.

Table 7. Architecture Faculty SES 2025 Electricity Production and Corresponding ton CO_{2e} Emission Amount

FACULTY OF ARCHITECTURE SES		
Period	Amount of Electricity Generated (kWh)	Corresponding Emission Quantity (tons CO _{2e})
January	465,16	0,29
February	9.263,26	5,80
March	18.689,26	11,70
April	15.939,27	9,98
May	32.457,28	20,32
June	33.730,99	21,12
July	30.173,60	18,89
August	25.771,89	16,14
September	24.973,76	15,64
October	20.749,06	12,99
November	13.233,74	8,29
December	0,00	0,00
General Total	225.447,27	141,15

The Faculty of Architecture will generate a total of 225,447.27 kWh of electrical energy within the SES area by 2025. This production corresponds to a total of 141,15 ton CO_{2e} emissions in 2025, based on the GHG Protocol Scope 2 reporting under the Location-based emissions calculation.

Table 8. Faculty of Medicine Dean's Office SES 2025 Electricity Production and Corresponding CO2e Emission Quantity

DEAN OF THE FACULTY OF MEDICINE SES		
Period	Amount of Electricity Generated (kWh)	Corresponding Emission Quantity (tons CO _{2e})
January	210,64	0,13
February	8.490,32	5,32
March	0,00	0,00
April	0,00	0,00
May	0,00	0,00
June	15.234,31	9,54
July	15.668,52	9,81
August	13.899,28	8,70
September	12.862,54	8,05
October	10.843,69	6,79
November	6.738,42	4,22
December	0,00	0,00
General Total	83.947,72	52,56

The Faculty of Medicine Dean's Office was unable to take measurements in March, April, and May 2025 due to repair work within the SES area. Within this scope, it provides a total of 83,947.72 kWh of electricity production in 2025. This production corresponds to a total of 52,56 ton CO_{2e} emissions in 2025, according to the GHG Protocol Scope 2 reporting under the Location-based emissions calculation. Looking at the two locations in the University SES area, a total of 309,394.99 kWh of electricity was generated in 2025. The emission amount corresponding to the amount of electricity generated was calculated as 193,71 ton CO_{2e}. Based on these calculations, an annual recovery rate of 11.96% was achieved.

FINDINGS and EVALUATION

Findings obtained based on the electricity production data of the solar energy systems (SES) installed at Mersin University's Çiftlikköy Campus and the electricity consumption data at the university and hospital levels are presented. Carbon emissions calculated based on annual electricity consumption amounts for 2025 and emission reduction amounts achieved as a result of SES applications were quantitatively evaluated in terms of carbon dioxide equivalent tons (tons CO_{2e}). The findings obtained were examined comparatively on a field basis, and

the contribution of solar energy systems to the reduction of carbon emissions as a result of their installation was revealed. In addition, the same data is examined comparatively with estimates made assuming that solar energy systems (SES) will operate at full capacity in 2026.

Based on the data obtained, when Mersin University Çiftlikköy Campus and Mersin University Hospital are evaluated together, it was determined that 87% of the total electricity consumption was from the hospital, while the consumption of the university campus accounted for 13%. A comparison of the consumption amounts of Mersin University Çiftlikköy Campus and Mersin University Hospital is presented in Figure 8.

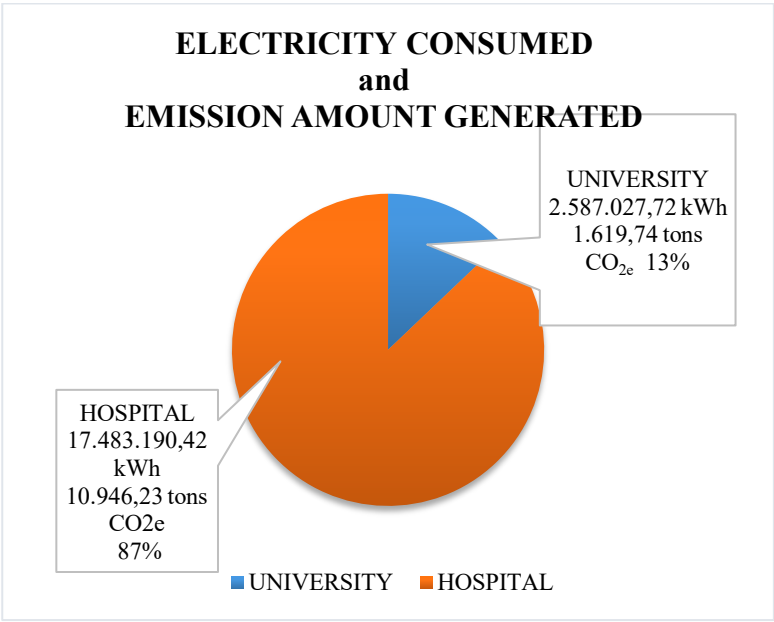


Figure 8. Hospital and University Consumption Data for 2025

Looking at the campus as a whole, the total electricity consumed was 20.070.218,74 kWh, resulting in emissions of 12.565,96 ton CO_{2e}.

Production data for the solar energy systems (SES) established at Mersin University Hospital and within the university, along with the corresponding emission quantities, are presented in Figure 9 and Figure 10 to reduce emissions.

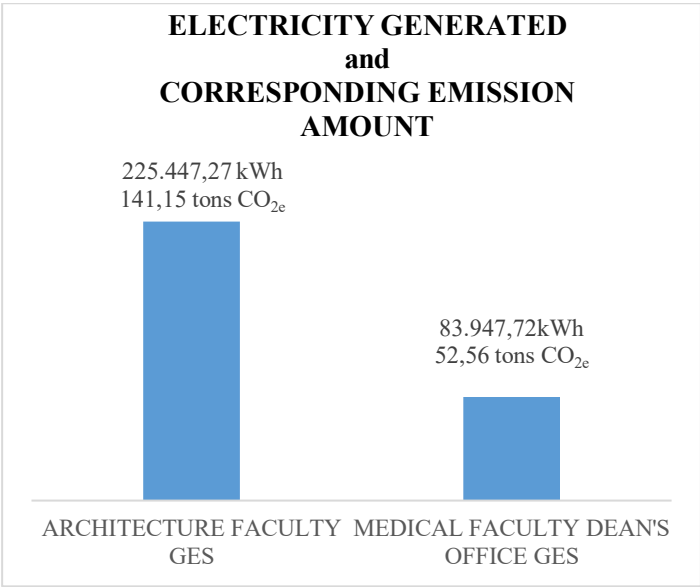


Figure 9. University SES Field 2025 Production Data

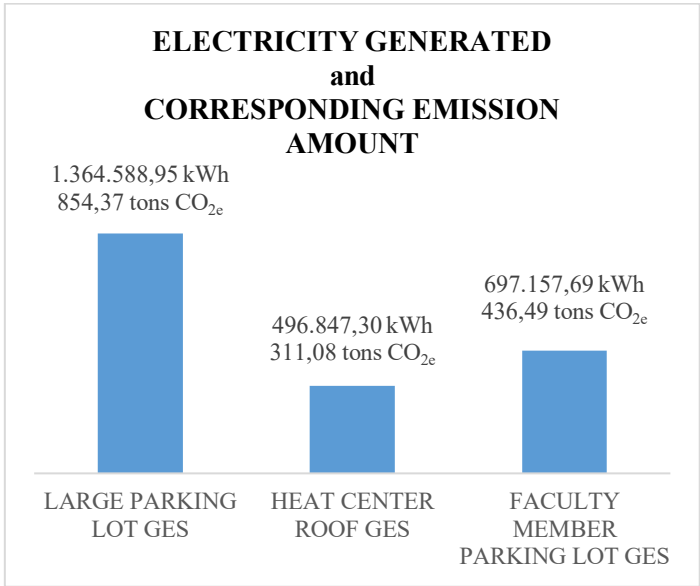


Figure 10. Mersin University Hospital SES 2025 Production Data

Looking at the total production data, the solar energy systems installed for the Mersin University Hospital SES area generated a total of 2,558. 593.94 kWh of electricity was generated, corresponding to 1.601,94 tons CO_{2e} emissions, while the university's SES area generated 309,394.99 kWh of electricity, corresponding to 193,71 tons CO_{2e} emissions. Based on these data, the total emissions corresponding to the 2,867,988.93 kWh of electricity generated by the solar energy systems amounted to 1.795,65 ton CO_{2e}.

Table 9. Reduction Rate in Post-Production Consumption (%)

Period	Consumption Amount (kWh)	Production Amount (kWh)	Reduction Rate in Post-Production Consumption (%)
Mersin University Çiftlikköy Campus	20.070.218,14	2.867.988,93	14,29

Table 10. Post-Production Consumption Emission Reduction Rate (%)

Period	Emissions from Consumption (tons CO _{2e})	Emissions from Production (tons CO _{2e})	Reduction Rate from Post-Production Consumption (%)
Mersin University Çiftlikköy Campus	12.565,96	1.795,65	14,29

In line with the results, as indicated in Table 9 and Table 10, thanks to the solar energy systems established to reduce consumption and emissions resulting from consumption, a total reduction of 14.29%. This highlights the role of solar energy systems installed on university campuses in promoting sustainable energy production and reducing carbon emissions. From a university perspective, this not only addresses environmental emission responsibilities but also contributes to setting an example for sustainable environmental and energy policies within society.

2026 Emission Estimates And Targets

Based on the electricity production data of the solar energy systems (SES) installed at Mersin University's Çiftlikköy Campus and the electricity consumption data at the university and hospital levels, it is targeted to increase the reduction rate in post-production consumption in 2026, assuming that renewable energy systems operate at full capacity without any downtime, repairs, or maintenance. As a result of this target, the estimated production quantities of the solar energy systems under full capacity operating conditions have been calculated, and carbon emission reduction targets for 2026 have been established within the scope of the estimated calculations..

Calculation results were generated assuming a 5% increase in electricity consumption across the university and hospital and an emission factor of 0,6261 kg CO₂ kWh⁻¹, with estimated values provided in Table 11.

Table 11. 2026 Consumption and Emission Estimates for Mersin University and Mersin University Hospital

Period	UNIVERSITY		HOSPITAL	
	Amount of Electricity Consumed (kWh)	Amount of Emissions (tons CO ₂ e)	Amount of Electricity Consumed (kWh)	Amount of Emissions (tons CO ₂ e)
January	197.887,73	123,90	1.213.890,30	760,02
February	258.764,63	162,01	1.444.891,77	904,65
March	183.060,68	114,61	838.687,50	525,10
April	162.280,13	101,60	846.058,50	529,72
May	213.966,90	133,96	1.144.829,70	716,78
June	282.035,25	176,58	1.845.018,00	1.155,17
July	295.213,28	184,83	2.516.308,20	1.575,46
August	237.144,21	148,48	2.748.816,00	1.721,03
September	303.945,08	190,30	2.072.252,70	1.297,44
October	237.144,21	148,48	1.223.056,80	765,76

November	215.585,64	134,98	1.589.380,95	995,11
December	246.761,65	154,50	1.778.850,32	1.113,74
General Total	2.833.789,37	1.774,24	19.262.040,74	12.059,96

Based on estimated values, it was determined that the total electricity consumed across the campus in 2026 will be 22,090,830.11 kWh, resulting in 13.834,20 tons CO_{2e}.

The estimated production data for 2026 for the solar energy systems (SES) installed across the campus and the corresponding emission values are shown in Table 12 and

Table 13 and the total sums are shown in Figure 11 and Figure 12.

Table 12. Estimated SES Production Data in 2026 (Amount of Electricity Generated (kWh))

Period	UNIVERSITY SES		HOSPITAL SES		
	ARCHITECTURE FACULTY SES	MEDICAL FACULTY DEAN'S OFFICE SES	LARGE PARKING LOT SES	HEAT CENTER ROOF SES	ACULTY MEMBER PARKING LOT SES
January	2.465,16	2.210,64	7.000,00	4.400,00	1.900,00
February	9.263,26	8.490,32	11.677,58	7.406,74	3.177,15
March	18.689,26	14.200,00	121.430,13	52.336,89	65.997,64
April	15.939,27	14.500,00	172.484,73	57.152,35	77.690,34
May	32.457,28	15.800,00	256.735,99	73.765,58	103.220,91
June	33.730,99	18.234,31	267.380,08	77.039,45	105.861,36
July	30.173,60	25.668,52	202.488,08	72.474,63	93.587,60
August	35.771,89	33.899,28	236.699,11	86.954,64	101.013,66
September	24.973,76	22.862,54	98.768,01	51.126,88	81.210,89
October	20.749,06	10.843,69	77.746,73	39.397,38	67.538,65
November	13.233,74	6.738,42	74.178,51	23.192,76	37.859,49
December	5.465,16	4.210,64	72.000,00	17.400,00	22.900,00
General Total	242.912,43	177.658,36	1.598.588,95	562.647,30	761.957,69

Table 13. Emissions Quantities Corresponding to Estimated SES Production in 2026 (tons CO₂e)

Dönem	UNIVERSITY SES		HOSPITAL SES		
	ARCHITECTURE FACULTY SES	MEDICAL FACULTY DEAN'S OFFICE SES	LARGE PARKING LOT SES	HEAT CENTER ROOF SES	ACULTY MEMBER PARKING LOT SES
January	1,54	1,38	4,38	2,75	1,19
February	5,80	5,32	7,31	4,64	1,99
March	11,70	8,89	76,03	32,77	41,32
April	9,98	9,08	107,99	35,78	48,64
May	20,32	9,89	160,74	46,18	64,63
June	21,12	11,42	167,41	48,23	66,28
July	18,89	16,07	126,78	45,38	58,60
August	22,40	21,22	148,20	54,44	63,24
September	15,64	14,31	61,84	32,01	50,85
October	12,99	6,79	48,68	24,67	42,29
November	8,29	4,22	46,44	14,52	23,70
December	3,42	2,64	45,08	10,89	14,34
General Total	152,09	111,23	1.000,88	352,27	477,06

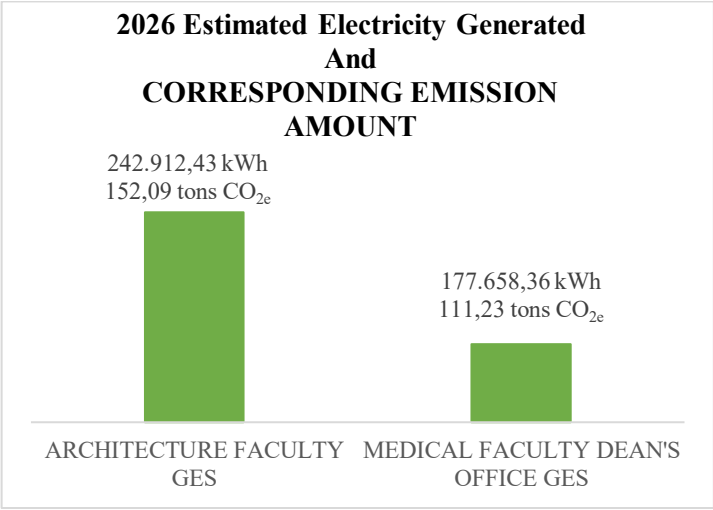


Figure 11. University SES Area 2026 Production Forecasts

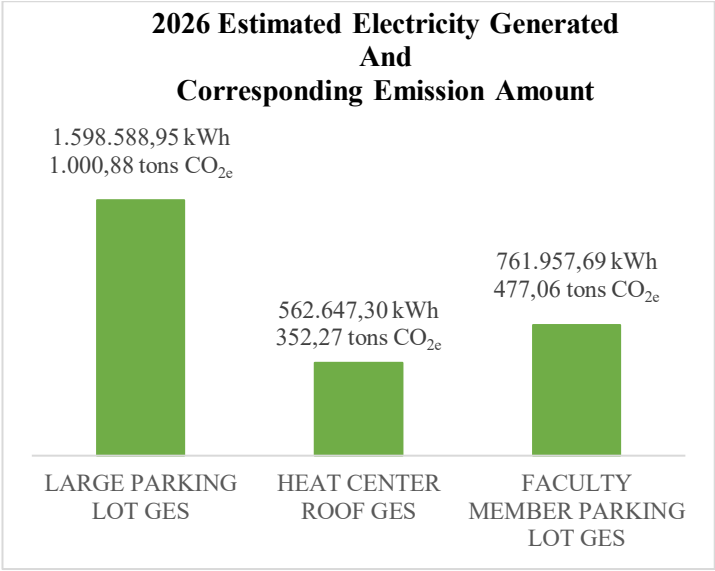


Figure 12. Mersin University Hospital SES 2026 Production Forecasts

Looking at the estimated production data for 2026, the solar energy systems installed for the Mersin University Hospital SES area generated a total of 2,923.193.94 kWh of electricity was generated, corresponding to 1.830,21 tons CO_{2e} emissions, while the university's solar energy field generated 420,570.79 kWh of electricity, corresponding to 263,32 tons CO_{2e} emissions. Based on these data, the total emissions corresponding to the 3,343,764.73 kWh of electricity generated by solar energy systems amounted 2.093,53 ton CO_{2e}.

Table 14. Estimated Reduction Rate in Post-Production Consumption (%)

Period	Consumption Amount (kWh)	Production Amount (kWh)	Reduction Rate in Post-Production Consumption (%)
Mersin University Çiftlikköy Campus	22.090.830,11	3.343.764,73	15,13

Table 15. Estimated Post-Production Consumption Emission Reduction Rate (%)

Period	Emissions from Consumption (tons CO ₂ e)	Emissions from Production (tons CO ₂ e)	Reduction Rate from Post-Production Consumption (%)
Mersin University Çiftlikköy Campus	13.834,20	2.093,53	15,13

Based on the results of the estimated calculations, as indicated in

Table 14 and Table 15, it has been determined that the solar energy systems established to reduce consumption levels and emissions resulting from consumption have achieved a total reduction of 15.13% in consumption and emissions across the campus for the year 2026.

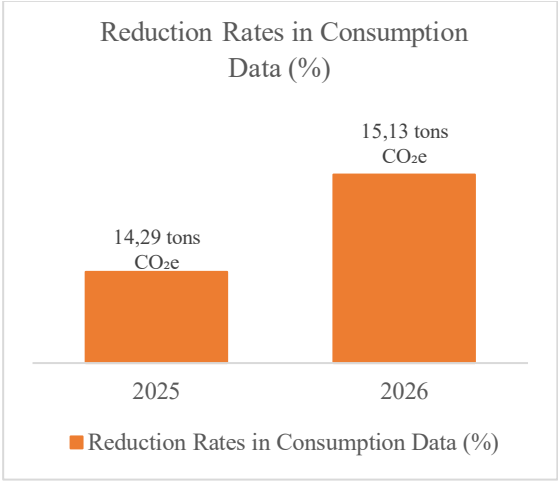


Figure 13. Comparison of 2025 Emission Data and 2026 Estimated Emission Data

When comparing the emission reduction data for 2025 with the estimated emission reduction data for 2026, it has been observed that the system's full capacity operation results in an increase in the emission reduction rate. This situation highlights the importance of solar energy systems installed on university campuses, sustainable energy production, and the reduction of carbon emissions. When viewed within the context of universities, it contributes to the budgetary efforts and initiatives of universities to reduce greenhouse gases and increase electricity production from renewable energy sources.

ACKNOWLEDGEMENTS

We would like to thank the administration of Mersin University for their support and cooperation in providing data on electricity consumption and solar energy systems at Mersin University Çiftlikköy Campus and Mersin University Hospital during the course of this study. The data shared has made a significant contribution to the scientific soundness and reliability of the study.

Furthermore, the practices carried out within the university in the field of sustainable environment and renewable energy, along with the institutional support provided, have been instrumental in the completion of this study. We would like to thank all relevant units for their contributions and support to scientific research.

REFERENCES

- Alagoz, E., & Alghawi, Y. (2023). The future of fossil fuels: Challenges and opportunities in a low-carbon world. *International Journal of Earth Sciences Knowledge and Applications*, 5(3), 381–388.
- Ari, İ., & Köksal, M. A. (2011). Carbon dioxide emission from the Turkish electricity sector and its mitigation options. *Energy Policy*, 39(10), 6120–6135. <https://doi.org/10.1016/j.enpol.2011.06.067>
- Arena, F., & Brooks, R. (2025). The role of fossil fuels in historical and current CO₂ levels. ResearchGate. https://www.researchgate.net/publication/391848227_The_Role_of_Fossil_Fuels_in_Historical_and_Current_CO_Levels
- Bazalian, M., Onyeji, I., Liebreich, M., MacGill, I., Chase, J., Shah, J., ... Zhengrong, S. (2013). Re-considering the economics of photovoltaic power. *Renewable Energy*, 53, 329–338. <https://doi.org/10.1016/j.renene.2012.11.029>
- Eggleston, S., Buendia, L., Miwa, K., Ngara, T., & Tanabe, K. (2006). 2006 IPCC guidelines for national greenhouse gas inventories. Intergovernmental Panel on Climate Change.
- González, A. G., Calcedo, J. G., & Salgado, D. R. (2018). Evaluation of energy consumption in German hospitals: Benchmarking in the public sector. *Energies*, 11(9), 2279. <https://doi.org/10.3390/en11092279>
- Greenhouse Gas Protocol. (2015). GHG Protocol scope 2 guidance: An amendment to the GHG Protocol corporate standard. <https://ghgprotocol.org/sites/default/files/2023-03/Scope%202%20Guidance.pdf> International Energy Agency. (2023). Emissions factors & electricity data. <https://www.iea.org/data-and-statistics/data-product/emissions-factors-2024>
- Jamsranjav, B. (2023). IPCC emission factor database (EFDB). Intergovernmental Panel on Climate Change.
- Keleş, S., & Bilgen, S. (2012). Renewable energy sources in Turkey for climate change mitigation and energy sustainability. *Renewable and Sustainable Energy Reviews*, 16, 5199–5206. <https://doi.org/10.1016/j.rser.2012.04.013>
- Lozano, R., Ceulemans, K., Almeida, M. A., Huisingh, D., Lozano, F. J., Waas, T., ... Hugé, J. (2015). A review of commitment and implementation of sustainable development in higher education: Results from a worldwide survey. *Journal of Cleaner Production*, 108, 1–18. <https://doi.org/10.1016/j.jclepro.2014.09.048>
- Mersin Üniversitesi. (2024). Mersin Üniversitesi tarihçesi. <https://www.mersin.edu.tr/universitemiz/tarihce>

- Mersin Üniversitesi Hastanesi. (2014). Mersin Üniversitesi Hastanesi hakkında. <https://hastane.mersin.edu.tr/tr/kurumsal/>
- Messenger, R. A., & Ventre, J. (2010). Photovoltaic systems engineering (3rd ed.). CRC Press.
- National Renewable Energy Laboratory. (2018). Best practices for operation and maintenance of photovoltaic and energy storage systems (3rd ed.). <https://www.nrel.gov>
- Our World in Data. (2024). Carbon intensity of electricity generation. <https://ourworldindata.org/grapher/carbon-intensity-electricity>
- Parida, B., Iniyar, S., & Goic, R. (2011). A review of solar photovoltaic technologies. *Renewable and Sustainable Energy Reviews*, 15(3), 1625–1636. <https://doi.org/10.1016/j.rser.2010.11.032>
- Peng, J., Lu, L., & Yang, H. (2013). Review on life cycle assessment of energy payback and greenhouse gas emission of solar photovoltaic systems. *Renewable and Sustainable Energy Reviews*, 19, 255–274. <https://doi.org/10.1016/j.rser.2012.11.035>
- Peng, J., Lu, L., & Yang, H. (2013). Review on life cycle assessment of energy payback and greenhouse gas emission of solar photovoltaic systems. *Renewable and Sustainable Energy Reviews*, 19, 255–274. <https://doi.org/10.1016/j.rser.2012.11.035>
- Plurawatt. (2025). PW M10X72 545–555W photovoltaic module. <https://plurawatt.com/tr/m10x72-545-555w>
- Ramos, T. B., Caeiro, S., van Hoof, B., Lozano, R., Huisingh, D., & Ceulemans, K. (2015). Experiences from the implementation of sustainable development in higher education institutions: Environmental management for sustainable universities. *Journal of Cleaner Production*, 106, 3–10. <https://doi.org/10.1016/j.jclepro.2014.01.094>
- Singh, A., Meena, H., & Khandelwal, C. (2023). Sustainability assessment of higher education institutions: A systematic literature review. *Engineering Proceedings*, 37(23), 1–7. <https://doi.org/10.3390/engproc2023037023>
- Sreng, R., & Gümrükçüoğlu Yiğit, M. (2017). Carbon footprint studies on Esentepe Campus of Sakarya University, Turkey in 2015. *Sakarya Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 1095–1099.
- T.C. Enerji ve Tabii Kaynaklar Bakanlığı. (2023). Türkiye elektrik üretimi ve elektrik tüketim noktası emisyon faktörleri. <https://enerji.gov.tr/evced-cevre-ve-iklim-elektrik-uretim-tuketim-emisyon-faktorleri>

- T.C. Enerji ve Tabii Kaynaklar Bakanlığı. (2024). Türkiye ulusal elektrik şebekesi emisyon faktörü. <https://enerji.gov.tr/evced-cevre-ve-iklim-turkiye-ulusal-elektrik-sebekesi-emisyon-faktoru>
- Tsoutsos, T., Frantzeskaki, N., & Gekas, V. (2005). Environmental impacts from the solar energy technologies. *Energy Policy*, 33(3), 289–296. [https://doi.org/10.1016/S0301-4215\(03\)00241-6](https://doi.org/10.1016/S0301-4215(03)00241-6)
- Tsoutsos, T., Frantzeskaki, N., & Gekas, V. (2005). Environmental impacts from the solar energy technologies. *Energy Policy*, 33(3), 289–296. [https://doi.org/10.1016/S0301-4215\(03\)00241-6](https://doi.org/10.1016/S0301-4215(03)00241-6)
- United Nations. (1992). United Nations framework convention on climate change. United Nations.
- Yavuz, A. B., Kara, O., & Yanıktepe, B. (2023). Karbon ayak izi tespiti: Osmaniye Korkut Ata Üniversitesi örneği. *Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi*, 12(4), 1523–1530.

CHAPTER 19

Applicability of Multicriteria Decision-Making Approaches in Underground Dam Planning

Ahmet Emir Köse¹ & Salim Serkan Nas² & Yusuf Kaya³

¹ PhD. Student, Department of Civil Engineering, Gumuhane University, Gumushane 29100, Turkiye, ORCID: 0000-0003-3890-5721

² Prof. Dr., Department of Civil Engineering, Gumuhane University, Gumushane 29100, Turkiye, ORCID: 0000-0001-9054-4674

³ Asst. Prof. Dr., Department of Civil Engineering, Gumuhane University, Gumushane 29100, Turkiye, ORCID: 0000-0002-5985-353X

NTRODUCTION

Currently, dynamics such as population growth, urbanization, industrialization and climate change are key factors that directly affect the quantity and quality of water resources. Globally, only a small portion of freshwater resources is available for human use. It is reported that 1.4 billion people lack access to sufficient drinking water, while 2.3 billion people do not have access to safe water. The decline in annual per capita water availability in Turkiye from 1.652 m³ in 2000 to 1.346 m³ in 2020 clearly demonstrates that the country is experiencing increasing water stress. Therefore, ensuring water security in both quantitative and qualitative terms, especially in arid and semi-arid regions, requires the implementation of integrated basin management approaches (Kolay and Ozturk, 2021).

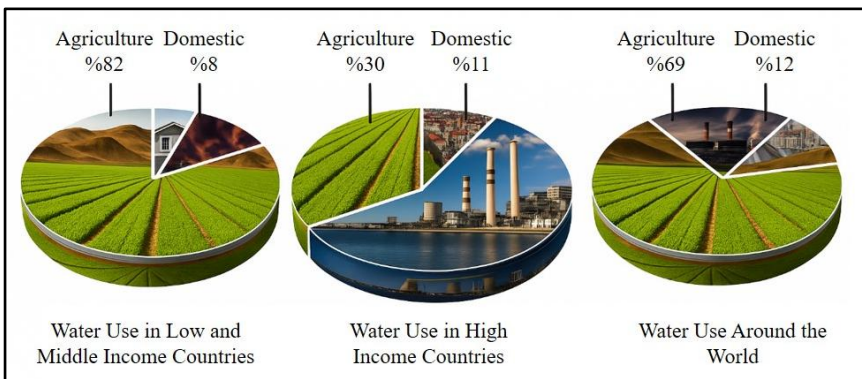


Figure 1. Water usage areas in the world (Atci, 2025).

Surface dams, which have been used for many years to address the increasing need for water storage, represent an important solution. However, they have significant limitations, including high evaporation losses, deterioration of water quality due to external pollutants, inundation of large agricultural areas and various ecological impacts. Given that evaporation losses have a substantial influence on the annual water budget, particularly in semi-arid regions, the need for alternative and more sustainable storage techniques has increased. In this context, underground dams have emerged as an effective alternative. These structures, which store water within the aquifer, eliminate evaporation losses, preserve natural ground filtration and have a significantly lower environmental impact compared to surface reservoirs. Underground dams have been widely implemented since the 1970s in many regions, including Japan, Brazil, India, South Korea, China and several African countries. In Turkiye, they have begun to gain importance since the 2000s, with notable examples such as Yahsihan (2003), Malibogazi (2004), Sancar (2007), Iskilip (2010) and Baskil (2011) (Apaydin, 2022a).



Figure 2. Serinyol underground dam (Hatay)(Soz Newspaper, 2022).

The planning of underground dams depends not only on hydrogeological properties but also on many factors such as lithology, permeability, soil type, topographic slope, precipitation and flow regime, aquifer type, land use and distance to residential areas. Most of these parameters involve uncertainty and may conflict with one another. For example, while alluviums with high conductivity offer advantages in terms of water storage, the same environment may present disadvantages when faults and fractures increase the risk of leakage. Therefore, the site selection and design of underground dams constitute a multidimensional, interdisciplinary process that is too complex to be addressed solely through the intuitive judgments of decision makers (Tuzlu, 2025).

Multi-Criteria Decision Making (MCDM) methods stand out as an important scientific tool for managing this complex structure. The Analytical Hierarchy Process (AHP), developed by Saaty in the 1970s, enables the weighting of both qualitative and quantitative criteria through pairwise comparisons and Malczewski (1999) demonstrates that integrating AHP with Geographic Information Systems (GIS) provides a powerful decision-support mechanism for land suitability analyses. The Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) method, on the other hand, ranks alternatives based on the principle of proximity to the ideal solution; when used in conjunction with AHP, it yields a consistent and quantifiable analytical structure. The use of these methods in water resources and hydrogeology has increased rapidly in recent years (Arca and Celik, 2025).

UNDERGROUND DAMS

Underground dams are defined as engineering structures that intercept groundwater flow in aquifers and store water within natural or artificial aquifer systems. They are typically constructed in areas where valleys narrow and alluvial aquifers exist, by placing an impermeable barrier down to the base of the aquifer. These barriers can be built using various materials such as clay, concrete, plastic or block rock and the choice of material depends on geological and economic conditions. In the literature, these structures are categorized under various classifications, such as submerged, burial or raised dams and they are also referred to by different names depending on their body material and location (Kaya, 2020).

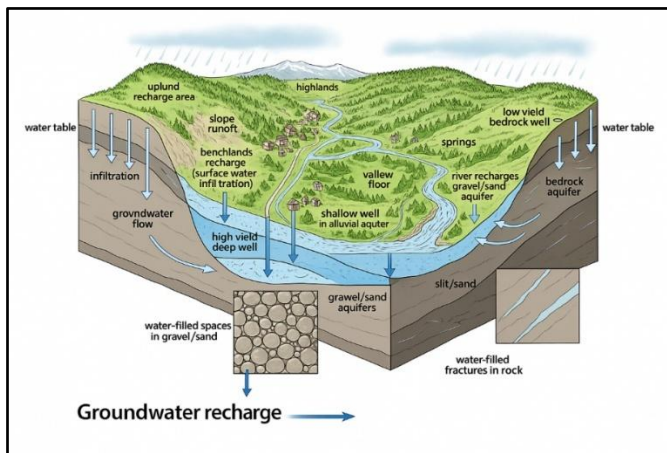


Figure 3. Groundwater Flow and Aquifer Structure (Hydrogeological Section) (Platin, 2018).

In arid and semi-arid regions, underground dams stand out as an important solution for the sustainable management of water resources. Among artificial recharge methods, underground dams are reported to offer effective applications for developing small-volume water resources, increasing groundwater levels and improving water quality in response to risks such as low rainfall, climate change, increased water demand and groundwater depletion (Tuzlu, 2025). Therefore, the implementation of underground dams has accelerated, particularly in regions at high risk of water scarcity, such as Africa, the Middle East and Turkiye. The development of numerous underground dam projects in Turkiye within the scope of action plans implemented by the State Hydraulic Works (SHW) demonstrates that this method is considered an economical, rapid and efficient artificial recharge technique (Apaydin, 2022b).

The successful operation of underground dams depends on appropriate site selection and accurate determination of the aquifer's hydrogeological

characteristics. Criteria such as aquifer thickness, hydraulic conductivity, recharge conditions and sidewall permeability are decisive factors. Compared to surface dams, determining reservoir volume in underground dams is more challenging and requires expertise due to the absence of visible aquifer boundaries (Tuzlu, 2025). Because of these difficulties, studies in rock aquifers or karst areas are more complex and the amount of water that can be stored in these environments can be higher. Indeed, the Miyakojima Dam in Japan, which stores 20 million m³ of water and supplies 50 000 m³ of water per day, demonstrates this capacity (Dogan, 2025).

Types of Underground Dams

Groundwater structures are generally classified into two main groups: underground dams and sand storage dams. Underground dams are structures constructed below the ground surface to prevent groundwater flow within a natural aquifer. Sand storage dams, on the other hand, are systems that store water within sedimentary material accumulated behind the dam. Underground dams are often used to raise water levels or to prevent saltwater intrusion into freshwater aquifers. Petersen (2010) stated that sand storage dams are frequently damaged by floods, alterations in flow conditions and structural deformations and therefore exhibit lower performance than underground dams. He also noted that sand storage dams have higher construction costs and recommended underground dams as the preferred option (Rajabi et al., 2025).

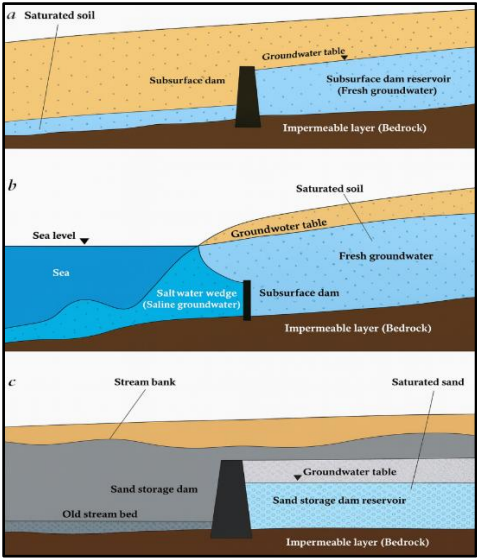


Figure 4. The longitudinal sections of (a) Subsurface dam for water supply; (b) Subsurface dam for preventing seawater intrusion; (c) Sand-storage dam (Rajabi vd., 2025).

Underground dams are divided into two distinct classes based on the type of soil in which they are constructed: dams built in alluvial aquifers and dams constructed on karst rocks. The construction of underground dams on alluvial deposits is more common, particularly in arid and semi-arid regions. Underground dams in karst rocks, however, are used less frequently due to geological complexity, high costs and uncertainties regarding storage capacity. These uncertainties make it difficult to predict the behavior of both natural groundwater flow and artificial recharge. Therefore, the planning of underground dams in karst environments is not common. Consequently, the construction of underground dams requires detailed geological and hydrogeological studies (Rajabi et al., 2025).

Advantages and Disadvantages of Underground Dams

While underground dams have certain technical, economic and environmental advantages compared to surface dams, they also present some limitations. The prominent advantages and disadvantages are listed below (Kolay and Ozturk, 2021).

Advantages of Underground Dams

- Because underground dams are buried structures, they can be constructed with steeper slope angles and thinner sections, requiring lower excavation and backfill costs compared to surface dams.
- Since water is stored underground, there is no loss of surface land, which significantly reduces expropriation costs.
- As there is no open water body on the surface, they do not pose a security risk and allow the area to continue its previous land use following dam construction.
- They enable water supply in areas where geological conditions are unsuitable for surface dam construction.
- They provide protection against saltwater intrusion by preventing seawater from entering freshwater aquifers.
- When water pumped from the dam is used upstream, it can return to the aquifer, creating a cyclical water movement.
- Due to their buried structure, there is no risk of dam collapse or sudden water discharge, thereby minimizing flood hazards.
- Because no open reservoir exists on the surface, the economic life of the structure is not reduced by sedimentation.

- Since storage occurs underground, there is no water loss due to evaporation and the system is not directly affected by environmental pollution.

Disadvantages of Underground Dams

- Due to the heterogeneous structure of aquifers, it is difficult to precisely determine hydrogeological parameters (storage coefficient, permeability etc.); nevertheless, these parameters are often generalized to represent the entire aquifer.
- The amount of storable water may remain at very low levels in some areas.
- Downstream aquifers may be negatively affected by dam impoundment.
- Water levels in the downstream area may decrease, leading to ground subsidence.
- Lower downstream water levels may increase the risk of seawater intrusion.
- Groundwater retention may increase the liquefaction potential in surrounding aquifers.
- Changes in upstream and downstream water levels may lead to alterations in the vegetation, soil fauna and flora of the region.
- Increased pore pressure resulting from stored water may increase uplift pressure on infrastructure in the area.
- Downstream water delivery can be costly due to pumping expenses or the excavations required for gravity-based conveyance. Since the structure is entirely underground, detecting water leakage is difficult, which can increase operation and maintenance costs.

Advantages and Disadvantages of Underground Dams Compared to Above-Ground Dams

In above-ground dams, the inability to use the land beneath the reservoir for other purposes leads to significant social and economic challenges, such as land loss, expropriation costs, ownership disputes and forced relocation. In underground dams, such issues do not arise because a large surface reservoir is not formed. Furthermore, the construction process for underground dams is shorter, easier to implement and, in many cases, less expensive than that of

surface dams. Another major advantage of these dams is their significantly lower water loss due to evaporation (Kaya, 2020).



Figure 5. In above-ground dams, mostly productive lands or settlements remain under the lake and become unusable (Corum-Obruk dam reservoir) (Kaya, 2020).

However, the planning and site selection processes for underground dams are more complex than for surface dams. While determining reservoir capacity in surface dams can be relatively easy through mapping and streamflow measurements, estimating the amount of water that can be stored within the porous geological unit and accurately assessing aquifer recharge conditions in underground dams are considerably more challenging. Therefore, detailed hydrogeological studies are essential for underground dams. These studies include geological and hydrogeological mapping, geophysical investigations, well drilling, field and laboratory testing and the application of various calculation methods (Kaya, 2020).

Size and Capacities of Underground Dams

Underground dams offer significant advantages over surface dams, such as eliminating evaporation losses, preventing surface land loss and providing a safer storage environment against external factors. However, the most significant disadvantage of these systems is their limited storage capacity. In surface dams, the reservoir volume is entirely filled with water, whereas in underground dams, water is stored only within porous grains or within fracture-fissure systems of the rock mass. Therefore, even if the total storage volume appears large, the actual storage capacity is determined by pore size and porosity (Apaydin, 2022a).

Nevertheless, regular recharge from upstream or surface water sources significantly mitigates the negative effects of low storage capacity. In this context, the evaluation of underground dams requires consideration not only of the storage capacity dependent on the rigid geological skeleton but also of

seasonal and hydrodynamic recharge conditions. This necessitates advanced expertise in project design and site selection and requires comprehensive hydrogeological investigations. These investigations include detailed geological and hydrogeological mapping, geophysical measurements, well drilling, laboratory and field tests and the application of different calculation methods (Apaydin, 2022a).

Surface dams constructed in Türkiye exhibit a wide range of capacities for irrigation and drinking water supply; storage volumes vary from millions to billions of cubic meters and dam heights typically range from 15 to 100 meters. In contrast, underground dams constructed before the action plan had dam heights ranging from 13 to 31.5 meters, are generally earthfill structures and include a compacted clay layer as an impermeable core. Annual storage capacities range from 0.7 to 6 hm³ (Apaydin, 2022a).

Obtaining Water from Underground Dams

In underground dams, the choice of whether to extract water by gravity or by pumping depends on topographic and aquifer characteristics. In areas with sufficient slopes, water can be brought to the surface by gravity; however, in areas with low slopes or thick aquifers, pumping becomes necessary. While gravity-driven systems have higher construction costs, they require no energy expenditure during operation, resulting in lower long-term operating costs. Therefore, during underground dam design, it is crucial to evaluate water extraction options through a detailed feasibility study that considers topographic, geological and economic conditions (Kaya, 2020).

In underground dams constructed in alluvial aquifers or coastal aquifers, water is extracted through caisson wells or boreholes. In gravity-driven systems, the water intake structure is placed at the elevation determined during dam construction and the depth of the intake directly affects extraction capacity. The depth of the intake structure is a key factor in determining the amount of groundwater that can be extracted from the porous medium. In cases where gravity extraction is not possible, large-diameter wells are drilled and pumping is used. In aquifers with low hydraulic conductivity, gravity alone is insufficient, making pumping a more appropriate solution (Kaya, 2020).

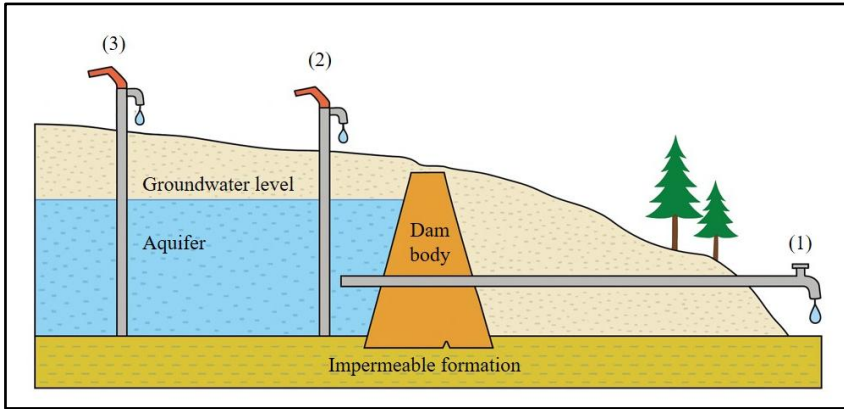


Figure 6. Options for obtaining water from underground dams; 1: Gravity, 2: Pumping from caisson wells, 3: Pumping from boreholes (Kaya, 2020).

In Japan, one of the countries with the highest use of underground dams worldwide, the vast majority of systems operate through pumping. Underground dams are widely used, particularly in coastal aquifers, to store freshwater that would otherwise discharge into the sea and to prevent saltwater intrusion. The Nare Underground Dam in Burkina Faso, Africa, where pumps are powered by solar energy, is a notable example of energy-efficient operation. In Brazil, approximately 500 small-scale underground dams were constructed during the drought of the 1990s and these systems generally serve small-scale drinking water and irrigation needs (Apaydin, 2022c).

The majority of underground dams constructed in Türkiye in the 2000s were designed to operate by gravity. The Kirikkale-Yahsihan, Ankara-Kalecik-Malibogazi and Corum-Iskilip underground dams are examples of this. In contrast, the Ankara-Elmadag-Kargali Underground Dam is one of the first to operate through pumping and incorporates both surface and underground reservoirs. While most new dams constructed under the Underground Dam Action Plan are designed for gravity operation, some projects employ different dam types and water intake methods. For example, the Bartın-Bahcecik Underground Dam collects water through drainage pipes and transfers it downstream by gravity, while the Izmir-Kiraz-Suludere Underground Dam plans to pump water from wells into the reservoir. With the increasing number of underground dams in Türkiye, new dam types and diverse water intake solutions are being implemented (Apaydin, 2022c).

Underground Dams in the World

A review of the literature on underground dams reveals that these structures have been implemented in many countries, particularly for the purpose of providing drinking water. Underground dams have been constructed as primary

sources of drinking water in Kenya, Tanzania, Somalia, Burkina Faso, India, China and Korea. The distance between settlements and water resources in Africa has been a significant factor contributing to the widespread use of underground dams in these countries. Korea is reported to have six underground dams. In India, where rainfall is limited except during monsoon seasons, underground dams are considered a critical water storage solution. Approximately 500 underground dams were constructed in Brazil to meet drinking water and irrigation needs during the drought of the 1990s (Dogan, 2025).

In Asia, especially in Japan, underground dams are extensively used in coastal aquifers. These dams were constructed to prevent seawater intrusion and protect freshwater reserves. The Fukuzato Underground Dam on Okinawa Island is notable for its storage capacity of 20 million cubic meters. Various underground dams are also located on islands such as Waite, Tengakuma, Tunegami, Kabasima, Minafuku and Kaki (Dogan, 2025).

While underground dam applications are limited in Europe, these dams have been used primarily for storage purposes in Germany, France and Italy. Underground dams were constructed in Austria and Greece for groundwater development and in the former Yugoslavia for preventing seawater intrusion. Managed Aquifer Recharge (MAR) techniques are more commonly used in European countries, aiming to raise water levels and meet agricultural water demands through underground storage of surface water. However, underground dams are not common in Scandinavian countries. In these countries, efforts to sustainably manage groundwater and protect reserves are prioritized. MAR applications are used to store surface water underground and to ensure the long-term protection of water resources (Dogan, 2025).

Underground dams in Turkiye

Underground dam projects in Turkiye have a long history and have been developed by various institutions at different scales. The General Directorate of Soil Conservation and Irrigation and the General Directorate of Rural Services have constructed numerous small-scale underground dams in recent years. According to 2023 data from the SHW, a total of 127 underground dams have been completed nationwide over the last two decades. In Anatolia and many other regions of Turkiye, particularly along tributaries of major valleys where alluvium narrows, geomorphological conditions offer suitable locations for the construction of underground dams (Dogan, 2025).

The first underground dam in Turkiye was constructed in Cesme to prevent seawater from mixing with underground freshwater reserves. In the following years, new underground dams were built to meet drinking water and irrigation needs. For example, in 2003, the Yahsihan Underground Dam-constructed by the

5th Regional Directorate of the State Hydraulic Works and the Kirikkale Provincial Directorate of Rural Services-was built to supply drinking water. The structure stands 14 meters above ground level. The Malibogazi Underground Dam in Ankara-Kalecik was constructed by the SHW in 2004 to provide irrigation services. Similarly, an underground dam was built for irrigation purposes in Seyitgazi, Eskisehir, in 2007 and another dam was constructed by the SHW in Corum in 2010 to supply drinking water. An underground dam used for irrigation in the Baskil district of Elazig was completed in 2011. Additionally, the Kirikkale Provincial Special Administration constructed two underground dams for drinking water in the village of Asagi Olunlu in the Bahsili district (Dogan, 2025).

New projects have recently been brought to the agenda. According to press reports, the region's first underground dam project is being prepared in Malatya-Puturge by the 9th Regional Directorate of the SHW for 2025. All these examples demonstrate that underground dams in Turkiye are becoming an increasingly important water management tool for both drinking water supply and irrigation activities (Dogan, 2025).

Site selection for underground dams

Site selection is one of the most critical factors directly impacting the success of underground dam projects. The site selection methodology for underground dams differs from that for surface dams in several ways. Porous and permeable alluvial sand-gravel deposits, which are generally not preferred for surface dam construction, are considered a favorable feature for the feasibility of underground dams. However, the thickness of the aquifer in which water will be stored is a decisive criterion; if the aquifer is too thick, construction costs increase, whereas insufficient thickness results in inadequate storage (Apaydin, 2022b).

In this context, aquifers with sufficient thickness and horizontal extent, suitable recharge properties, high storage coefficient and hydraulic conductivity, an impermeable base unit and a narrow cross-section suitable for dam construction constitute the fundamental requirements for underground dams. Furthermore, access to high-quality groundwater and affordable construction materials are also important considerations. Pre-construction studies for underground dams are not limited to geological and hydrogeological mapping, geophysical investigations, topographic analyses, exploratory drilling and aquifer tests (Apaydin, 2022b).

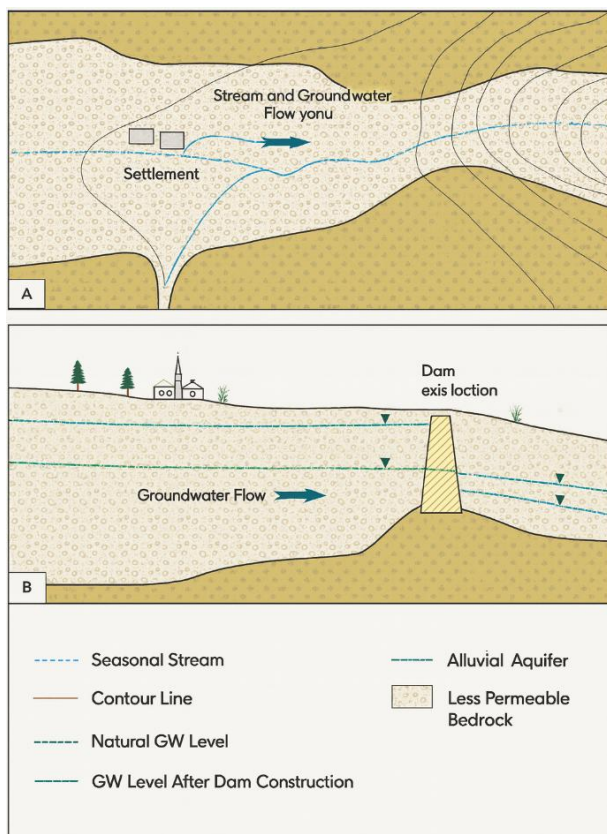


Figure 7. Figure 4. Ideal conditions for building an underground dam in valley alluviums (A: Map, B: Section) (Kolay and Ozturk, 2021).

Broader-scale assessments encompassing the entire aquifer and, if necessary, the entire basin are necessary. During the site selection process, it is essential to examine potential changes in the downstream area of the dam. The potential for swamp formation due to rising groundwater levels should be considered and legal arrangements for protecting water rights should be planned in advance. In some underground dams constructed in Türkiye in the 2000s (e.g., Malibogazi and Iskilip), the dams were positioned a few meters below the natural ground and covered with coarse-grained permeable material, preventing surface water accumulation. However, in areas with low slopes, the possibility of water levels reaching the surface and creating swamps should be taken into account and the project should be designed accordingly. For the sustainable operation of the dam basin, water withdrawals from wells-particularly those drilled in the upstream area-should be restricted and the area should be legally designated as a protected zone and announced in the Official Gazette. The same announcement should also include provisions for pollutant control (Apaydin, 2022b).

Conditions Required for Underground Dam Construction

The points to consider in underground dam construction are listed as follows (Kaya, 2020):

- Alluvium thickness is generally preferred not to exceed 15-20 m; however, dams can be constructed in thicker alluviums if excavation difficulties can be overcome or if plastic concrete curtain or adjacent pile techniques can be applied. When suitable topographic conditions are met, increasing the thickness also increases storage volume. Underground dams can also be constructed in rock aquifers; for example, dams constructed using the injection method in karst limestone in Japan exist.
- The aquifer is expected to be coarse-grained or contain fractures, fissures and dissolution voids; both the storage coefficient (S) and hydraulic conductivity (K) are expected to be high.
- An impermeable base suitable for excavation or injection must exist beneath the alluviums or rock aquifer.
- The aquifer must extend over a wide area or receive adequate recharge from surface runoff and adjacent formations during rainy periods. To reduce excavation and backfill costs, it is preferable to select a site in a narrowing aquifer strait.
- A sufficiently large underground reservoir must be formed behind the dam. Storage volume can be expanded when the slope is low, but this is not mandatory. On steep slopes, the storage volume below the crest may decrease; however, if the aquifer basin is wide and adequately recharged, a dam can still be constructed.

UNDERGROUND DAMS AND MULTI-CRITERIA DECISION MAKING METHODS

Underground dams stand out as strategic water management structures that enable water storage without evaporation loss, preserve water quality through natural soil filtration and reduce environmental impacts. Planning these structures requires the simultaneous analysis of numerous parameters, including geological features, soil type, topographic slope, precipitation regime, groundwater levels, fault lines, land-use patterns and aquifer characteristics. Multi-criteria decision-making methods-particularly AHP, which integrates qualitative and quantitative data and TOPSIS, which is based on the principle of proximity to the ideal solution-are used as important tools to systematically address this multidimensional structure. The AHP method, developed by Saaty in the 1970s, digitizes decision-makers' subjective assessments through consistency checks;

Malczewski (1999) emphasizes that integrating AHP with GIS yields robust results in spatial suitability analyses. Recent examples of the use of these methods in hydrogeology and water resources management include the study by Vashist and Singh (2024), which improved the accuracy of spatial decisions using AHP-based flood risk modeling in the Krishna River Basin. Integrating the criteria weights determined by AHP with TOPSIS enables the identification of the alternative closest to the ideal solution in underground dam site selection and contributes to the reliable evaluation of multi-parameter spatial data (Arca and Celik, 2025).

AHP Method

The AHP method is a multi-criteria decision-making approach developed by Saaty in the 1970s that allows for the combined evaluation of both quantitative and qualitative criteria in complex decision-making problems (Supciller and Capraz, 2011). The widespread use of AHP in the literature has been influenced by its easily applicable structure for decision-makers (Ho, 2008). The method consists of stages such as structuring the decision problem as a hierarchical model, making pairwise comparisons between criteria, calculating relative importance values and performing consistency analysis (Saaty, 2008; Ak Oguz and Koksall, 2018; Ilgaz, 2018).

Step 1: Creating the Hierarchical Structure

The purpose of the decision problem is positioned at the top level, forming a hierarchical structure. The middle level contains the criteria to be used in the decision process and the lowest level includes the alternatives to be evaluated (Saaty, 2008). Figure 8 illustrates this hierarchical structure of the AHP method.

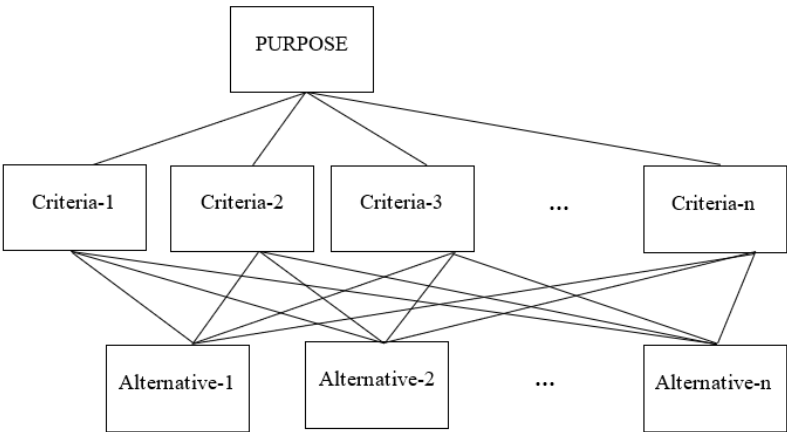


Figure 8. Hierarchical structure of the AHP method (Supciller and Capraz, 2011).

Step 2: Creating Pairwise Comparison Matrices (A) and Determining the Priorities

After the purpose, criteria and any sub-criteria are determined, an $n \times n$ pairwise comparison matrix is created to reveal the relative importance of these elements, as shown in Equation 1 (Saaty, 1990). At this stage, the decision maker compares the relevant criteria or alternatives in a pairwise manner, indicating which one is more important or dominant and to what extent.

$$A = \begin{bmatrix} 1 & a_{21} & a_{31} & \dots & a_{n1} \\ 1/a_{21} & 1 & a_{32} & \dots & a_{n2} \\ 1/a_{31} & 1/a_{32} & 1 & \dots & a_{n1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1/a_{n1} & 1/a_{n2} & 1/a_{n3} & \dots & 1 \end{bmatrix}_{n \times n} \quad (1)$$

The relative importance of each criterion in the pairwise comparison matrix, in terms of its contribution to the objective, as well as the superiority of alternatives against these criteria, is determined through pairwise comparisons based on the judgments of decision makers. In this process, the importance scale developed by Saaty and presented in Table 1 is used to ensure a consistent assessment of superiority levels (Saaty, 1990).

Table 1. Importance scale of the AHP method (Saaty, 1990).

Numerical Value	Definition
1	The items are of equal importance or indifferent between them.
3	Item 1 is slightly more important or slightly more preferred than Item 2.
5	Item 1 is more important or preferred than Item 2.
7	Item 1 is much more important or preferred than item 2.
9	Item 1 is extremely important or extremely preferred over item 2.
2,4,6,8	Intermediate values

Step 3: Calculating the Eigenvector (Relative Importance Vector)

After creating the pairwise comparison matrices, the eigenvector that reflects the relative importance of each element compared to the others is calculated. In this context, the relative weights of the criteria or alternatives are obtained by

determining the eigenvector of the matrix in $n \times 1$ dimension (Supciller and Capraz, 2011).

$$i=1,2,3,\dots,m \text{ and } j=1,2,3,\dots,n;$$

$$b_{ij} = \frac{a_{ij}}{\sum_{i=1}^n a_{ij}} \quad w_i = \frac{\sum_{j=1}^n b_{ij}}{n} \quad (2)$$

To determine the percentage importance distributions of the criteria, column vectors of the form $W = [w_i]_{n \times 1}$ must be calculated. This column vector is obtained by taking the arithmetic average of the elements in each row of the matrix consisting of the b_{ij} values defined in Equation 2 (Supciller and Capraz, 2011).

Step 4: Evaluating the Consistency of the Eigenvector

A consistency ratio (CR) is calculated for each pairwise comparison matrix and this ratio is expected to be at most 0,10. A CR value above 0,10 indicates that the decision maker is making inconsistent judgments in pairwise comparisons; in such cases, these judgments should be reviewed and improved. To calculate the consistency ratio, the largest eigenvalue ($\lambda_{\max}$) of matrix A must first be determined (Equation 4) (Supciller and Capraz, 2011).

$$i=1,2,3,\dots,m \text{ and } j=1,2,3,\dots,n;$$

$$D = [a_{ij}]_{n \times n} \times [w_i]_{n \times 1} = [d_i]_{n \times 1} \quad (3)$$

$$\lambda = \frac{\sum_{i=1}^n \frac{d_i}{w_i}}{n} \quad (4)$$

Another quantity required to calculate the consistency ratio is the randomness index (RI). RI values, which are constant coefficients determined according to the matrix dimension n , are presented in Table 2. In light of this information, the consistency ratio (CR) is calculated using the formula given in Equation 5 (Supciller and Capraz, 2011).

$$CR = \frac{\lambda - n}{(n-1) \times RI} \quad (5)$$

Table 2. Randomness index data (Guner, 2005).

n	1	2	3	4	5	6	7	8	9	10
RI	0	0	0,58	0,9	1,12	1,24	1,32	1,41	1,45	1,49

Step 5: Obtaining the General Result of the Hierarchical Structure

The previous four steps are applied to all levels of the hierarchical structure to complete the necessary calculations. In this step, the $(m \times 1)$ dominance column vectors obtained for each of the n criteria in the hierarchy are combined to form an $(m \times n)$ dimensional DW decision matrix. Multiplying this matrix by the W dominance vector, which represents the relative importance of the criteria, yields the R result vector, representing the final evaluation (Equation 7) (Supciller and Capraz, 2011).

$$i=1,2,3,\dots,n \text{ and } j=1,2,3,\dots,n;$$

$$DW = [w_{ij}]_{m \times n} \quad (6)$$

$$R = DW \times W \quad (7)$$

TOPSIS Method

Hwang and Yoon developed the TOPSIS method as a multi-criteria decision-making approach in the 1980s (Shyjith et al., 2008). This method is based on the principle that an alternative's distance from the positive ideal solution should be as small as possible, while its distance from the negative ideal solution should be as large as possible (Monjezi et al., 2012).

The TOPSIS method follows the steps below:

Step 1: Determining Objectives and Defining Evaluation Criteria

In the first stage of the decision-making process, the main objective of the problem is clearly stated and the evaluation criteria to be used in line with this objective are defined (Yurdakul and Ic, 2003).

Step 2: Creating the Decision Matrix (D)

At this stage, the alternatives ($a_1, \dots, a_n$) are listed in the rows of the matrix, while the values ($y_{1k}, \dots, y_{nk}$), which reflect the performance of each alternative with respect to the relevant criteria, are placed in the columns (Yurdakul and Ic, 2003). Thus, the structure of the decision matrix is created (Equation 8).

$$D = \begin{bmatrix} y_{11} & y_{12} & \dots & y_{1k} \\ y_{21} & y_{22} & \dots & y_{2k} \\ \dots & \dots & \dots & \dots \\ y_{n1} & y_{n2} & \dots & y_{nk} \end{bmatrix} \quad (8)$$

Step 3: Creating the Normalized Decision Matrix (R)

The criterion values in the decision matrix are normalized by dividing each value by the square root of the sum of the squares of all values corresponding to that criterion (Yurdakul and Ic, 2003). This normalization process is performed using Equation 9 and the resulting normalized decision matrix R is presented in Equation 10.

$$r_{ij} = \frac{y_{ij}}{\sqrt{\sum_{i=1}^n y_{ij}^2}} \quad i = 1, 2, \dots, n \quad j = 1, 2, \dots, k \quad (9)$$

$$R = \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1k} \\ r_{21} & r_{22} & \dots & r_{2k} \\ \dots & \dots & \dots & \dots \\ r_{n1} & r_{n2} & \dots & r_{nk} \end{bmatrix} \quad (10)$$

Step 4: Creating the Weighted Normalized Decision Matrix (V)

Using the w_j values representing the weights of each criterion, the elements in the normalized decision matrix are rescaled according to the importance levels of the corresponding criteria. In this way, the weighted normalized decision matrix V is obtained by reflecting the criterion weights onto the normalized values (Monjezi et al., 2012).

$$W = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1k} \\ w_{21} & w_{22} & \dots & w_{2k} \\ \dots & \dots & \dots & \dots \\ w_{n1} & w_{n2} & \dots & w_{nk} \end{bmatrix} \quad (11)$$

Each column element in the normalized decision matrix R is scaled by multiplying it with the corresponding weight value w_{ij} specified in Equation 11. Following this process, the weighted normalized matrix V is obtained, as shown in Equation 12 (Monjezi et al., 2012).

$$V = \begin{bmatrix} v_{11} & v_{12} & \dots & v_{1k} \\ v_{21} & v_{22} & \dots & v_{2k} \\ \dots & \dots & \dots & \dots \\ v_{n1} & v_{n2} & \dots & v_{nk} \end{bmatrix} \quad (12)$$

Step 5: Determining Ideal (A^*) and Negative Ideal (A^-) Solutions

The positive ideal solution consists of the highest performance values for the criteria in the weighted normalized decision matrix, while the negative ideal solution contains the lowest performance values for the same criteria (Shyjith et al., 2008). These two solution sets are calculated using Equations 13 and 14.

$$A^* = \left\{ (\max_i v_{ij} | j \in I), (\min_i v_{ij} | j \in J) \right\} \quad (13)$$

$$A^- = \left\{ (\min_i v_{ij} | j \in I), (\max_i v_{ij} | j \in J) \right\} \quad (14)$$

In the relevant formulas, I represents the benefit (maximization) criteria and J represents the cost (minimization) criteria. Accordingly, the positive ideal solution values obtained using Equation 13 are expressed as $A^* = \{v_1^*, v_2^*, \dots, v_k^*\}$, while the negative ideal solution values obtained from Equation 14 are expressed as $A^- = \{v_1^-, v_2^-, \dots, v_k^-\}$ (Monjezi et al., 2012).

Step 6: Calculating Separation (Distance) Measurements

At this stage, separation (distance) values are calculated to determine the extent to which the alternatives diverge from each other and from the ideal solution. The distance of each alternative from the positive ideal solution is determined using the formula given in Equation 15 (Monjezi et al., 2012).

$$S_i^* = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^*)^2} \quad (15)$$

Similarly, the distance of each alternative to the negative ideal solution is calculated using the formula given in Equation 16 (Monjezi et al., 2012).

$$S_i^- = \sqrt{\sum_{j=1}^n (v_{ij} - v_j^-)^2} \quad (16)$$

Step 7: Calculating Relative Closeness to the Ideal Solution

At this stage, the relative closeness value (C_i^*) is calculated using equation number 17 to determine how close each alternative is to the ideal solution (Monjezi et al., 2012).

$$C_i^* = \frac{S_i^-}{S_i^- + S_i^+} \quad 0 \leq C_i^* \leq 1 \quad (17)$$

Step 8: Ranking the Alternatives According to Their Relative Closeness to the Ideal Solution

In this final stage, the alternatives are ranked based on the calculated relative proximity values (C_i^*) and the alternative with the maximum C_i^* value is determined as the most suitable option (Monjezi et al., 2012).

Intended AHP-TOPSIS Integrated Approach

This integrated method, created by combining AHP and TOPSIS techniques, consists of three basic stages (Supciller and Capraz, 2011):

- (1) Defining the selection criteria to be used in the evaluation,
- (2) Structuring and weighting these criteria in a hierarchical model,
- (3) Evaluating and ranking the alternatives in light of the determined criteria.

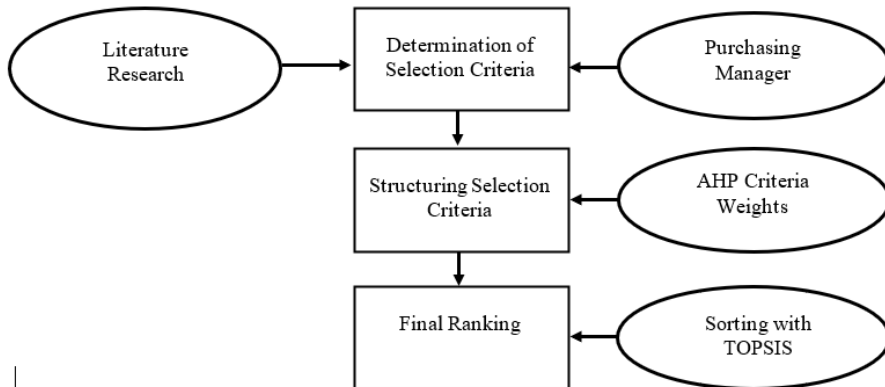


Figure 9. Intended AHP-TOPSIS Method Stages (Supciller and Capraz, 2011).

RESEARCH AND FINDINGS

In the literature, the majority of underground dam site selection studies focus on integrating GIS and remote sensing techniques with multi-criteria decision-making methods (Fathi et al., 2019; Hamlat et al., 2024). In these studies, physical and hydrogeological parameters such as precipitation, geology, slope, land use, drainage density and groundwater recharge are weighted using AHP or similar methods to produce suitability maps, thereby spatially defining suitable areas for dam construction (Ozegin et al., 2023; Arca and Celik, 2025). In addition, some studies compare the site selection process by applying different MCDM methods such as TOPSIS, Evaluation Based on Distance from Average Solution (EDAS) and (Elimination and Choice Translating Reality English) ELECTRE, highlighting the detailed analytical capabilities offered by these techniques (Kharazi et al., 2019; Dortaj et al., 2020).

The literature also reveals that underground dams have multifaceted functions, including sustainable water management, aquifer recharge, water quality protection and saltwater intrusion control. Their increasing strategic importance—particularly in the context of coastal aquifers, arid regions and the impacts of climate change—is emphasized (Epcin, 2023; Sen, 2022; Yifru et al., 2021). The rising water demand in recent years has led underground dams to be increasingly considered a technically and administratively effective solution alternative, a trend clearly demonstrated in literature reviews (Rajabi et al., 2025).

Fathi et al. (2019) used GIS, remote sensing and the AHP to assess criteria such as precipitation, groundwater recharge, geology and slope within a multi-criteria framework to identify suitable sites for underground dams in the Karaj Plain. Criteria weights were determined using AHP and a suitability map was created through weighted overlay analysis. The results showed that areas suitable for underground dam construction in the plain were clearly defined spatially.

Kharazi et al. (2019) developed a four-stage underground dam site selection method based on GIS and AHP, systematically analyzing the elimination of unsuitable areas, the identification of watercourses and axis selection. In the final stage, suitable areas were compared using the TOPSIS and the EDAS methods and priority dam sites were identified. Their findings demonstrated that the developed method provides high accuracy and consistency in underground dam site selection processes.

Dortaj et al. (2020) evaluated geology, hydrology and socioeconomic criteria for underground dam site selection using the ELECTRE I-II-III methods. They reported that the ELECTRE approach offers more detailed analysis capabilities than other methods in multi-criteria decision-making contexts.

Kaynarca et al. (2020) analyzed data from 1.067 wells in the Kirkgoz Basin to create groundwater level maps. They also developed a comprehensive baseline dataset by incorporating land use classes and elevation models derived from Sentinel-2 imagery.

Sezer Guney (2020) stated that underground dams are important structures within sustainable water management, providing reliable water supply particularly in arid regions. He also noted that these dams are an effective method for recharging aquifers and reducing water pollution.

Degerli and Turhan (2021) compiled experimental and numerical studies on saltwater intrusion in coastal aquifers and examined the hydrogeological dynamics of this process in detail. The literature indicates that uncontrolled groundwater use can lead to similar pollution problems in many coastal basins in Turkiye. Their findings revealed that underground dams offer an important alternative solution for both water supply and the prevention of saltwater intrusion.

Yifru et al. (2021) stated that increasing water demand in arid regions causes significant economic challenges, particularly in developing countries. For this reason, they emphasized that sand dams represent an effective solution for water storage when appropriate site selection and engineering design are applied.

Sen (2022) noted that climate change seriously threatens water resources in arid and semi-arid regions and stated that underground dams offer significant advantages in conserving water and mitigating climate risks under such conditions.

Epcin (2023) investigated how underground dams affect saltwater intrusion in coastal aquifers using a three-chamber Plexiglas test setup. It was demonstrated that dams at different heights and locations significantly slow the velocity and acceleration behavior of saltwater and disrupt flow dynamics. The results revealed that dam height and distance from the coast are critical factors in controlling saltwater intrusion.

Ozegin et al. (2023) produced a groundwater potential map by weighting criteria such as geomorphology, drainage and land use using AHP. They classified the study area into four classes, ranging from very low to high potential.

Aras et al. (2024) applied a three-stage decision support system for underground dam site selection in Bursa, evaluating various buffer zones, water demand and geological criteria together. At the end of the study, they identified six suitable underground dam sites.

Cavdar (2024) stated that increasing water losses in surface dams due to climate change and urbanization exacerbate water supply problems. They

emphasized that underground dams are an economical and effective alternative that reduces these losses.

Hamlat et al. (2024) used GIS, remote sensing and AHP methods to determine underground dam site selection in the M'zi Wadi Basin. By weighting numerous hydrological and geological criteria, they determined that 1.27% of the study area was suitable and confirmed these areas with geophysical studies.

Arca and Celik (2025) in their study, identified areas suitable for underground dam construction in the Diyarbakir Plain by integrating GIS-based AHP and TOPSIS methods. Weighting and suitability analyses based on eleven parameters showed that alluvial and low-slope areas had particularly high suitability scores. The results demonstrated that this approach provides an effective decision support tool for sustainable water management in semi-arid regions.

Rajabi et al. (2025) conducted a literature review examining the functions and appropriate site selection criteria for underground dams in the context of increasing water demand in arid regions. The findings indicate that studies focusing on site selection have increased significantly in recent years.

CONCLUSIONS

Underground dams are strategically important structures for effectively managing increasing water demand and climate uncertainty. Especially in semi-arid and arid regions with high evaporation, storing water directly within the aquifer offers a significant advantage by preventing evaporation losses, preserving water quality through natural filtration and largely eliminating problems such as land loss, sedimentation and ecological damage commonly associated with surface dams. Recent projects in various basins in Turkiye demonstrate that underground dams are a reliable alternative for both drinking water supply and irrigation.

However, the underground dam planning process is inherently multidimensional and complex. Numerous interdependent parameters-such as hydraulic conductivity, storage coefficient, aquifer thickness, topographic slope, lithology of geological units, faulting status, recharge potential, land-use patterns and downstream impacts-make the decision-making process unmanageable using traditional methods. Therefore, MCDM methods, thanks to their scientific and systematic analytical capabilities, play a decisive role in underground dam site selection. The use of AHP and TOPSIS, in particular, allows for the integrated evaluation of qualitative and quantitative data, transforming complex criteria sets into a numerical model and making decision-making processes objective and repeatable.

The AHP method stands out in determining the importance levels of criteria from different disciplines in underground dam planning. Its hierarchical structure enables the weighting of hydrogeological, geotechnical, topographic and environmental factors through pairwise comparisons, while the calculation of a consistency ratio increases the reliability of decision-makers' assessments. The criterion weights obtained through this method contribute significantly to structuring the decision problem on scientific foundations and minimizing uncertainties in expert judgments. In multi-parameter problems such as underground dam site selection, this approach demonstrates strong compatibility with both GIS-based analyses and other spatial assessment techniques.

The TOPSIS method, on the other hand, enables the ranking of alternatives according to the principle of "closeness to the ideal solution" by incorporating the weights obtained from AHP into the decision matrix. This model allows for the simultaneous evaluation of the strengths and weaknesses of potential underground dam sites, generating an integrated suitability score based on each site's distance from the positive ideal solution and its proximity to the negative ideal solution. Thus, a comprehensive result is achieved through the simultaneous evaluation of all criteria rather than reliance on a single parameter. A growing number of recent studies demonstrate that the AHP-TOPSIS integrated model provides high accuracy and consistency in underground dam planning.

Consequently, underground dams stand out as a powerful alternative within modern water management policies in terms of both environmental sustainability and economic viability. However, these structures can achieve long-term success only through appropriate site selection, robust hydrogeological analyses and a planning process supported by multi-criteria assessment models. In this context, the integration of MCDM methods such as AHP and TOPSIS into underground dam projects transforms complex hydrogeological parameters into a manageable structure, providing decision-makers with a reliable, numerical and comparable assessment framework. This scientifically supported approach forms the basis for more effective, accurate and sustainable utilization of underground dams against increasing water security pressures in the future.

REFERENCES

- Ak Oguz, M. and Koksall, M. (2018). Supplier selection using AHP and TOPSIS methods. *İstanbul Ticaret University Journal of Science*, 18(34), 69-89.
- Apaydin, A. (2022a). Underground dams against drought: An evaluation within the framework of the ‘Underground Dam Action Plan.’ *2nd International Cappadocia Scientific Research Congress*, Nevsehir, 971-977.
- Apaydin, A. (2022b). Feasibility of underground dams in Türkiye and evaluation of the ‘Underground Dam Action Plan’. *Konya Journal of Engineering Sciences*, 10(1), 130-146.
- Apaydin, A. (2022c). Water intake methods in underground dams and new applications in Türkiye. *International Black Sea Modern Scientific Research Congress*, Rize, 57-63.
- Aras, E., Boz, B., Vaheddoost, B. and Yılmaz, D. (2024). A case study for determination of the best underground dam sites, *Groundwater*, 62(5), 804-816.
- Arca, M. and Celik, R. (2025). Underground dam site selection in the Diyarbakır Plain using a multi-criteria decision-making model and GIS integration. *Dicle University Journal of Science Institute*, 14(1), 77-104.
- Atci, E. B. (n.d.). The state of water resources worldwide. Artemis Aritim. Retrieved November 10, 2025, from <https://www.artemisaritim.com/dunya-genelinde-su-kaynaklarinin-durumu>.
- Cavdar, P. S. (2024). A tool of sustainable control of groundwater resources: Underground dams. *Archives of Advanced Engineering Science*, 1-7.
- Degerli, S. and Turhan, E. (2021). The effect of saltwater intrusion in coastal aquifers and underground dams: A literature review. *Academic Perspective Procedia*, 4(2), 185-195.
- Dogan, G. (2025). Investigation of effective criteria weights in underground dam site selection using the DEMATEL method. Master’s thesis, Mugla Sıtkı Kocman University Institute of Social Sciences, Mugla, 41 p.
- Dortaj, A., Maghsoudy, S., Ardejani, F. D. and Eskandari, Z. (2020). A hybrid multi-criteria decision-making method for site selection of subsurface dams in a semi-arid region of Iran. *Groundwater for Sustainable Development*, 10, 100284.
- Epcin, A. (2023). Investigation of the effect of underground dams on the freshwater-saltwater balance in coastal areas. Master’s thesis, Bursa Technical University Graduate Education Institute, Bursa, 39 p.

- Fathi, A., Lee, T. and Mohebzadeh, H. (2019). Allocating underground dam sites using remote sensing and GIS: A case study on the southwestern plain of Tehran Province, Iran. *Journal of the Indian Society of Remote Sensing*, 47(6), 989-1002.
- Guner, H. (2005). Fuzzy AHP and its application to the supplier selection problem for a business. Master's thesis, Pamukkale University Institute of Science, Denizli, 133 p.
- Hamlat, A., Hamdi, K., Kissari, D. E., Kadri, C. B., Guidoum, A. and Sekkoum, M. (2024). GIS-based multi-criteria decision analysis for groundwater dam site selection in an arid and semi-arid region of Algeria. *Groundwater for Sustainable Development*, 26, 101231.
- Ho, W. (2008). Integrated Analytic Hierarchy Process and its applications: A literature review. *European Journal of Operational Research*, 186(1), 211-228.
- Ilgaz, A. (2018). Evaluation of personnel selection criteria in the logistics sector using AHP and TOPSIS methods. *Suleyman Demirel University Journal of the Institute of Social Sciences*, 1(32), 586-605.
- Kaya, S. (2020). Site selection application for underground dams using the direct current resistivity method. Master's thesis, Ankara University Institute of Science, Ankara, 49 p.
- Kaynarca, M., Demir, N. and San, B. T. (2020). An approach for modeling groundwater resources using remote sensing and GIS techniques: *Kirkgoz Basin (Antalya)*. *Geomatics*, 5(3), 241-245.
- Kharazi, P., Yazdani, M. R. and Khazealpour, P. (2019). Suitable identification of underground dam locations using decision-making methods in a semi-arid region of the Iranian Semnan Plain. *Groundwater for Sustainable Development*, 9, 100240.
- Kolay, E. and Ozturk, B. (2021). Underground dams in water supply: A case study. *Electronic Letters on Science and Engineering*, 17(2), 76-88.
- Malczewski, J. (1999). GIS and multicriteria decision analysis. John Wiley and Sons. ISBN: 978-0-471-32944-2.
- Monjezi, M., Dehghani, H., Singh, T. N., Sayadi, A. R. and Gholinejad, A. (2012). Application of TOPSIS method for selecting the most appropriate blast design. *Arabian Journal of Geosciences*, 5(1), 95-101.
- Ozegin, K. O., Ilugbo, S. O. and Ogunseye, T. T. (2023). Groundwater exploration in a landscape with heterogeneous geology: An application of geospatial and analytical hierarchical process (AHP) techniques in the Edo North region, Nigeria. *Groundwater for Sustainable Development*, 20, 100871.

- Petersen, E. N. (2010). Join the wave with Erik Nissen-Petersen. *Are sand dams the solution for water harvesting*, 1.
- Platin. (March 28, 2018). Groundwater. <https://platinsondaj.com/izmir-sondaj-yeralti-suyu/>. (Accessed: 10.11.2025).
- Rajabi, A. M., Alizadehnia, S. and Bidar, A. S. (2025). A review of studies on underground dam site selection. *Physics and Chemistry of the Earth, Parts A/B/C*, 103995.
- Saaty, T. L. (1990). How to make a decision: The analytic hierarchy process. *European Journal of Operational Research*, 48(1), 9-26.
- Saaty, T. L. (2008). Decision making with the analytic hierarchy process. *International Journal of Services Sciences*, 1(1), 83-98.
- Sezer Guney, B. (2020). Underground dams within the framework of the basin management approach. *Su Kaynakları*, 5(1), 7-12.
- Shyjith, K., Ilangkumaran, M. and Kumanan, S. (2008). Multi-criteria decision-making approach to evaluate optimum maintenance strategy in the textile industry. *Journal of Quality in Maintenance Engineering*, 14(4), 375-386.
- Soz Gazetesi. (March 28, 2022). Serinyol Underground Dam opened right on time. <https://sozgazetesi.com.tr/2022/03/28/serinyol-yeralti-baraji-tam-gununde-acildi/>.
- Supciller, A. and Capraz, O. (2011). Supplier selection application based on the AHP-TOPSIS method. *Istanbul University Econometrics and Statistics e-Journal*, (13), 1-22.
- Sen, Z. (2022). Climate change and Türkiye. *Cevre Sehir ve İklim Dergisi*, 1(1), 1-19.
- Tuzlu, F. (2025). Determination of suitable areas for underground dams using remote sensing and GIS techniques. Doctoral dissertation, Suleyman Demirel University Institute of Science, Isparta, 202 p.
- Vashist, K. and Singh, K. K. (2024). Flood hazard mapping using GIS-based AHP approach for Krishna River Basin. *Hydrological Processes*, 38(6), e15212.
- Yifru, B. A., Chung, I. M., Kim, M. G. and Chang, S. W. (2021). Assessing the effect of land use/land cover and climate change on water yield and groundwater recharge in the East African Rift Valley using an integrated model. *Journal of Hydrology: Regional Studies*, 37, 100926.
- Yurdakul, M. and Ic, Y. (2003). A case study using the TOPSIS method for performance measurement and analysis of Turkish automotive firms. *Gazi Universitesi Muhendislik Mimarlık Fakultesi Dergisi*, 18(1), 1-18.

CHAPTER 20

Position Control of a DC Motor via PID Controller Optimized by Musk Ox Algorithm

Said Müftü¹ & Erdem Dilmen²

¹ Res. Asst., Department of Mechatronics Engineering, Pamukkale University, Denizli, Turkey, ORCID: 0000-0001-5621-7805

² Asst. Prof. Dr., Department of Mechatronics Engineering, Pamukkale University, Denizli, Turkey, ORCID: 0000-0003-3092-0927

1. Introduction

Proportional–Integral–Derivative (PID) control is one of the most widely used control strategies in engineering due to its simplicity, robustness, and effectiveness in regulating dynamic systems. The importance of PID control lies in its ability to minimize the error between a desired reference signal and the actual system output by combining proportional action for immediate response, integral action for eliminating steady-state error and derivative action for improving stability and transient performance. Due to this balanced structure, PID controllers can be implemented without requiring an exact mathematical model of the system, which makes them highly practical for real world engineering applications. PID control has been extensively applied in mechanical engineering for motion control of robotic manipulators and servo systems (Astrom, 1995). In electrical and electronics engineering, it plays a critical role in power systems such as voltage and frequency regulation, motor speed control and converter control (Ogata, 2010). In chemical engineering, PID control is commonly used for temperature, pressure and flow control in process industries where maintaining stable operating conditions is essential for safety (Seborg et al., 2016). In addition, PID controllers are widely adopted in aerospace engineering for flight control systems and attitude stabilization of unmanned aerial vehicles due to their reliability and ease of tuning (Bequette, 2003). Despite the emergence of advanced control methods such as model predictive control (Camacho and Bordons, 1999) and adaptive control (Åström and Wittenmark, 2013), PID control remains dominant in industry. Ongoing research continues to focus on improving PID performance through intelligent tuning techniques, such as fuzzy logic (Passino et al., 1998) and optimization algorithms, demonstrating that PID control is still a prominent part of modern engineering practice.

DC (direct current) motors have a wide range of applications that require a drive element (Brown, 2002; Matsui and Shigyo, 2002). DC motors have two important control objectives: position and speed. The importance of speed (Şahal and Ortatepe, 2025; Şahin et al., 2021) and position (Müftü and Gökçe, 2024) has been emphasized in the literature for different purposes.

Position control of DC motors takes place in a variety of applications in the industry. It offers advantages over other motor types with important factors such as acceleration time and sensitivity to reference values. In the literature, several studies can be reviewed concerning control of DC motors. Fuzzy logic has been used in conjunction with PID controllers for position control of DC motors (Kaplan et al., 2020). In (Çakar et al., 2020) PID controllers are compared with fuzzy controllers while backstepping technique is used for DC motor control in (Doruk and Zuglem, 2018). Also, reader can refer to (Yaren and Kizir, 2022) for model predictive control of DC motor.

Optimization problems are encountered in many fields within today's scientific world. Many methods have been considered as solution approaches (Lange, 2013; Pedregal, 2004). Metaheuristic algorithms are currently seen with new nature-inspired structures (Yang, 2011a, 2011b) and they can be classified as evolutionary, swarm, physics, and nature-based approaches (Fadhil, 2025).

Musk Ox (MO) has demonstrated its success in benchmark functions among metaheuristic algorithms. In addition to benchmark functions, it has proven successful in areas such as three-bar truss design, pressure vessel design, stress/strain analysis, speed reducer design, vehicle side impact design, and piston rod design (Yuan et al, 2025).

Best to the authors' knowledge, this study proposes the first application of MO algorithm to determine the optimal parameters of PID position controller of a DC motor. First, transfer function of the encoder-equipped DC motor (Faulhaber 2342I012CR) is obtained. Then, it is used to simulate the DC motor in Matlab Simulink. The values provided by MO were used as K_p , K_i and K_d parameters of the PID controller employed. Also, performance by MO optimized PID control is compared with the performances when PSO (particle swarm optimization) (Clerc and Kennedy, 2002) and ABC (artificial bee colony) (Karaboga, 2005) algorithms are used. It should be noticed that, PSO and ABC have been widely studied in optimal PID design so far and proven successful in this task. In parallel with that, PID controllers optimized by PSO and ABC produced satisfactory results for a step reference at 45 degrees and for a complex reference consisting of sinusoidal and square waves. However, PID controller with its parameters optimized by MO yielded better performance than the case of PID controllers with optimal parameters provided by ABC and PSO. Thus, the idea of adopting MO optimization to design optimal PID controller is proven useful as an alternative to the state-of-the-art algorithms within the DC motor control framework.

Chapter is organized as follows. Sect. 2 gives a brief explanation of PID control technique. DC motor transfer function is introduced in Sect. 3 and optimization algorithms adopted for PID controller design are presented in Sect. 4. Then, based on the optimally determined controller parameters, numerical experiments associated with position control of DC motor by different PID controllers are conducted in Sect. 5. Finally, study is concluded in Sect. 6.

2. PID Control

The PID control method is one of the most widely used feedback control strategies in industrial practice. Its popularity stems from its simple structure, practical implementability and satisfactory performance for a broad class of plants (Ogata, 2010). A PID controller computes the control action from the

tracking error $e(t)$ between the reference input $r(t)$ and the measured output $y(t)$. The time domain PID control law is as shown in (1).

$$u(t) = K_p e(t) + K_i \int_0^t e(\tau) d\tau + K_d \frac{de(t)}{dt} \quad (1)$$

Here, K_p , K_i and K_d denote the proportional, integral and derivative gains, respectively. The proportional action increases responsiveness while the integral action reduces steady-state error and the derivative action improves transient behavior by anticipating error variation. Fig. 1. illustrates a standard negative-feedback PID control loop. The summing junction forms the error signal $e(t) = r(t) - y(t)$, which is processed by the PID controller to generate the control input $u(t)$ applied to the plant $G(s)$. $H(s)$ (Often unity) denotes the feedback gain.

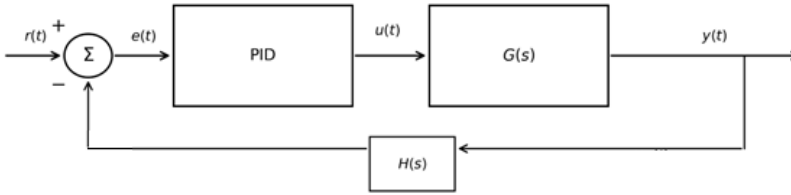


Fig. 1. Standard negative feedback PID control loop.

In the Laplace domain, the PID controller transfer function can be written as in (2).

$$G_{PID}(s) = K_p + \frac{K_i}{s} + K_d s \quad (2)$$

This form is convenient for robustness analysis, allowing designers to relate controller parameters to stability margins and bandwidth (Franklin et al., 2002).

Controller performance depends strongly on tuning. The classical Ziegler–Nichols (ZN) tuning method is an empirical procedure that provides a rapid initial parameter set. It is based on the ultimate gain K_u and the ultimate period T_u observed when the closed-loop system oscillates with constant amplitude (Astrom, 1995).

The ZN ultimate gain tuning procedure is typically applied as follows:

- Set integral and derivative actions to zero, i.e., $K_i = 0$ and $K_d = 0$ (pure P control).

- Increase K_p gradually until sustained (marginal) oscillations occur in the closed loop.
- Record the gain at this point as the ultimate gain K_u .
- Measure the oscillation period as the ultimate period T_u .
- Compute PID parameters using the ZN rules (Table 1), then refine if needed.

In practice, safety constraints and actuator limits should be considered when searching for K_u , and the procedure should be performed carefully to avoid damaging equipment. The obtained settings often yield relatively aggressive responses (higher overshoot), so the method is commonly used as a starting point followed by fine-tuning against performance requirements.

Table 1. Ziegler–Nichols ultimate gain tuning rules for a PID controller.

Controller Type	K_p	K_i	K_d
PID	$0.6 K_u$	$1.2 K_u / T_u$	$0.075 K_u \cdot T_u$

Note that, in this study MO algorithm is adopted as the parameter tuning strategy for the PID controller.

3. DC Motor Transfer Function

This study uses an encoder-equipped DC motor with a reducer, offered by the Faulhaber brand under code 2342012CR. According to the product information shared by the brand, specifications of the DC motor are listed in Table 2.

Table 2. Faulhaber 2342I012cr motor characteristics values.

Feature	Symbol	Value
Nominal voltage	U_N	12V
Terminal resistance	R	1,9 Ω
Output power	$P_{2nom.}$	17W
Efficiency, max.	$\eta_{max.}$	%80
No-load speed	η_0	8100rpm
No-load current	I_0	0,075A
Stall torque	M_H	80mNm
Friction torque	M_R	1mNm
Speed constant	k_n	713rpm/V
Back-EMF constant	k_E	mV/rpm
Torque constant	k_M	13,4mNm/A
Current	k_i	0,075A/mNm

constant		
Slope of n-M curve	$\Delta n / \Delta M$	101rpm/mNm
Rotor inductance	L	65 μ H
Mechanical time constant	τ_m	6ms
Rotor inertia	J	5,7gcm ²

Since a DC motor is an electromechanical system, its electrical and mechanical equations must be obtained when obtaining the transfer function. The electrical equation of the DC motor is given by (3), while its mechanical equation is given by (4).

$$V(s) = LsI(s) + RI(s) + k_s\omega(s) \quad (3)$$

$$Js\omega(s) + B\omega(s) = k_T I(s) \quad (4)$$

If the current is eliminated, (5) is obtained.

$$\frac{\omega(s)}{V(s)} = \frac{k_T}{(Ls+R)(Js+B)+k_Tk_e} \quad (5)$$

If the denominator of (5) is expanded, (6) is obtained.

$$(Ls + R)(Js + B) = LJs^2 + (LB + RJ)s + RB \quad (6)$$

As a result of the above equations, velocity transfer function is obtained as given by (7).

$$\frac{\omega(s)}{V(s)} = \frac{k_T}{LJs^2 + (LB + RJ)s + (RB + k_Tk_e)} \quad (7)$$

The transition from the speed transfer function to the position transfer function is achieved as shown by (8).

$$\frac{\theta(s)}{V(s)} = \frac{1}{s} \cdot \frac{\omega(s)}{V(s)} = \frac{k_T}{s[LJs^2 + (LB + RJ)s + (RB + k_Tk_e)]} \quad (8)$$

By assigning the values in Table (2) to the expression in (8), transfer function of the DC motor is obtained by (9).

$$\frac{\theta(s)}{V(s)} = \frac{0.0134}{[3.705e-11s^3 + 1.089e-6s^2 + 3.595e-4s]} \quad (9)$$

Overall configuration of the numerical applications is given in Fig. 2.

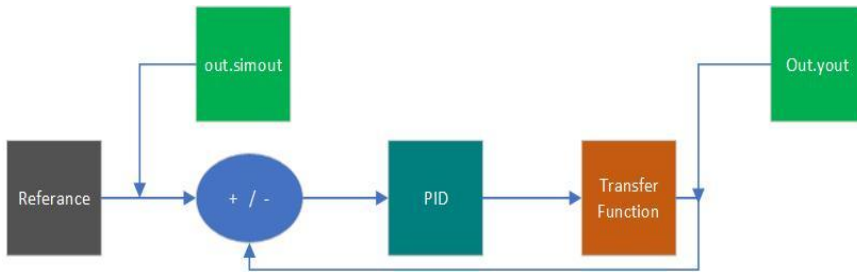


Fig. 2. System block diagram.

4. Metaheuristic Optimization Algorithms

Gradient-based optimization methods utilize the derivative information for parameter optimization. Derivative of the cost function of interest wrt to the associated parameters are used to compute parameter update term, which is based on mathematical foundations regarding extremum points of a function. Thus, when conditions imposed on the update term are met, they guarantee convergence of the parameters to the nearest optimum point. This point highly depends on the initial coordinates and convergence to global optimum cannot be ensured. If the initial parameter vector denotes a point in the parameter space which is closer to any local minimum rather than the global one, cost function is minimized to that local minimum. This is referred to as ‘getting stuck at local minima’, an issue that may make gradient-based optimization problematic sometimes. Although gradient-based algorithms have strong capability of exploitation in parameter space, they cannot provide large exploration for optimum parameters. This is a drawback of these algorithms. However, metaheuristic algorithms perform search in the entire parameter space via multiple populations. Considering the fact that those algorithms possess random assignments at some stages of the optimization process, any individual population can converge to further optimum points which do not necessarily be the closest one. Because of this, these algorithms can converge to a neighborhood of the global optimum within a satisfactory level. This is established via the approach of finding the best individual within an entire solution set (Ting et al., 2014). An outstanding class of metaheuristic algorithms consists of nature-inspired ones.

The ability of living creatures in nature to sustain their own lives and continue their lineage has been a source of inspiration for scientists. Metaheuristic algorithms draw their inspiration from nature. Metaheuristic algorithms can be considered relatively simple in mathematical structure. They can be used for different types of problems. With these advantages, they can be easily applied to different types of problems. Living creatures that live in swarms, such as ants,

particles, bears, bees and wolves can be classified based on swarm-based gravity, electricity, magnetism, physics-based genetic structure and population based mutation evolution (Gandomi et al., 2013; Kareem et al., 2022).

4.1. Particle Swarm Optimization

PSO is a population-based metaheuristic inspired by social behavior in biological swarms and is widely used for controller tuning problems (Clerc and Kennedy, 2002; Kennedy and Eberhart, 1995). In PSO, each particle is characterized by its position x_i^k and velocity v_i^k at iteration k . The speed update equation is as follows.

$$v_i^{k+1} = wv_i^k + c_1r_1(p_{best,i} - x_i^k) + c_2r_2(g_{best} - x_i^k) \quad (10)$$

In this equation, w is the inertia weight controlling momentum, c_1 and c_2 are cognitive and social coefficients, r_1 and r_2 are random variables in $[0,1]$, $p_{best,i}$ is the personal best position and g_{best} is the global best position. The particle position is updated according to the (11).

$$x_i^{k+1} = x_i^k + v_i^{k+1} \quad (11)$$

Linear decreasing inertia weight is achieved with (12).

$$w(k) = w_{max} - \frac{w_{max}-w_{min}}{k_{max}}k \quad (12)$$

Here, w_{max} and w_{min} are maximum and minimum inertia weights and k_{max} is the maximum iteration count. Pseudocode for PSO is given in Algorithm 1 in Appendix A1.

Note that, PSO finds itself in a variety of engineering optimization problems. If we consider particular task of optimal PID tuning, each particle encodes PID gains as $x_i = [K_p, K_i, K_d]_i$ (Åström and Hagglund, 1995). A common objective function is the ITAE (integral time absolute error) index, expressed by (13).

$$J = \int_0^T t|e(t)|dt \quad (13)$$

In this expression, $e(t)$ is the tracking error, T is the evaluation horizon and J is the performance index to be minimized (Seborg et al., 2011). Last, cost function is formulated as reference tracking RMSE (root mean squared error).

4.2. Artificial Bee Colony Optimization

ABC is a swarm intelligence-based optimization algorithm inspired by the foraging behavior of honey bees. Owing to its simplicity and robustness, ABC has been widely used in numerical optimization and control parameter tuning problems (Karaboga, 2005; Karaboga and Basturk, 2007). In ABC algorithm,

each potential solution is modeled as a food source. The population consists of employed bees, onlooker bees, and scout bees, which cooperate to explore and exploit the search space. Each food source is represented by the position vector x_i . A new candidate food source v_{ij} is generated using the following neighbor search expression in (14).

$$v_{ij} = x_{ij} + \phi_{ij}(x_{ij} - x_{kj}) \quad (14)$$

In (14), x_{ij} denotes the j_{th} component of the i_{th} food source, x_{kj} is the j_{th} component of a randomly selected food source k , and ϕ_{ij} is a uniformly distributed random number in the range $[-1, 1]$, controlling the search step size.

The fitness value associated with a food source x_i is commonly defined by (15).

$$fitness(x_i) = \frac{1}{1+f(x_i)} \quad (15)$$

Onlooker bees select food sources based on the probability P_i , which is given by (16).

$$P_i = \frac{fitness(x_i)}{\sum_{n=1}^{SN} fitness(x_n)} \quad (16)$$

Here, SN represents the number of food sources in the colony, and food sources with higher fitness values have a higher chance of being selected. If a food source cannot be improved for a predefined number of cycles, it is abandoned and replaced with a new solution randomly generated according to (17).

$$x_i^{new} = x_{min} + rand(0,1)(x_{max} - x_{min}) \quad (17)$$

In (17), x_{min} and x_{max} define the lower and upper bounds of the search space respectively while $rand(0,1)$ denotes a random real number in $(0, 1)$. Pseudocode for ABC is given in Algorithm 2 in Appendix A2.

ABC optimization algorithm has been employed in a wide range of engineering applications, which includes PID controller tuning as well. In this particular task, similar procedure is followed, which is mentioned in Sect. 4.1 for PSO.

4.3. Musk Ox Optimization

Living organisms found in nature serve as a significant source of inspiration for scientists across various scientific disciplines. With the considerable advancement in mathematical analysis tools, the development of different benchmark problems and efforts to better understand nature, the scientific

community has made significant progress in analyzing living organisms and modeling their fundamental behaviors. MO is among the most recent metaheuristic algorithms which is introduced to the literature in 2025 (Yuan et al, 2025). Its success has been proven by comparing its performance with several algorithms on benchmark functions and real-world applications. The new generation metaheuristic MO model was created by analyzing the social behavior of musk oxen and modeling it mathematically. Musk oxen are creatures that largely succeed in migrating, feeding, and defending themselves in order to survive under harsh living conditions. Signals sent to the herd by sentinel individuals direct the collective movement of the herd in the situations critical for survival.

Migration behavior ensures that the herd moves toward the most suitable conditions. Defensive behavior ensures that the herd can continue to live as a whole. Migration and defense are of great importance for maintaining the best position. Random exploration is carried out due to the herd's need for new food sources. Random exploration increases the diversity of the herd and ensures access to new resources.

One of the most fundamental problems of gradient-based optimization algorithms, the issue of getting stuck in local minima, does not occur in MO thanks to the random foraging behavior of mules. This enables MO to generate effective solutions within complex search spaces. The mathematical structure of the MO algorithm is specified in (18)-(26).

$$\mathbf{X} = \left[\begin{array}{ccc} X_{1,1} & \cdots & X_{1,d} \\ \vdots & \ddots & \vdots \\ X_{n,1} & \cdots & X_{n,d} \end{array} \right]_{n \times d} \quad (18)$$

In (18), n represents the population size and d represents the scale of the problem. X , which denotes an individual in the herd, can be computed as shown in (19).

$$X = MO_{lb} + rand (MO_{ub} - MO_{lb}) \quad (19)$$

Here, $rand$ denotes a random real number in (0, 1) while MO_{lb} denotes the lower bound while MO_{ub} denotes the upper bound. The portion of immature individuals is 20% of the musk ox herd. There are two signals in MO algorithm, warning and safety. When there is no threat to the musk oxen, the herd prefers to shelter. In this mode, the individuals in the herd interact with each other to renew their positions. In the renew mode, the defense margin must be high. The solution for the best position is specified in (20)-(22). Besides, the variables $r_4 - r_6$ in the following equations are random numbers.

$$X^{t+1} = X^t + \Delta x_1 \quad (20)$$

$$\Delta x_1 = rand (X_{best}^t - X^t) + r_4^3 W_i (X^t - X_{best}^t) \quad (21)$$

$$W_i = G_1 \frac{1}{\exp\left(\frac{\beta x t}{T}\right)} \quad (22)$$

In (20)-(22), X_{best}^t is the position of the best ox, W_i is the step effect factor, and G_1 and β are hyperparameters of MO algorithm. The position of musk oxen is updated using (23)-(25).

$$X^{t+1} = X^t + \Delta x_2 \quad (23)$$

$$\Delta x_2 = r_5^2 (X_{second}^t - X_{best}^t) + r_5 M_i (X_{best}^t - X_{worst}^t) \quad (24)$$

$$M_i = G_2 \frac{1}{\exp\left(\frac{\beta x t}{T}\right)} \quad (25)$$

In these equations, X_{second}^t represents the position of the second ox in all iterations, while X_{worst}^t represents the position of the worst ox. G_2 is a hyperparameter with a value of 100. M_i is the step control factor. The position update for female and young musk oxen is provided by (26).

$$X^{t+1} = r_6 (X_{best}^t - X^t) LF \quad (26)$$

In (26), LF is a random walk (Levy flight) and is computed by (27). Here a is a constant with a value of 1.5, $x \sim N(0, \sigma_x^2)$, $y \sim N(0, \sigma_y^2)$ where σ_x^2 and σ_y^2 denote the standard deviations. Reader may refer to (Yuan et al, 2025) for the computation of the standard deviations. Pseudocode for MO is given in Algorithm 3 in Appendix A3.

$$Levy(a) = 0.05 \frac{x}{|y|^{\frac{1}{a}}} \quad (27)$$

In this study, MO optimization algorithm is proposed for tuning the parameters of a PID controller employed in DC motor position control. To do this, similar procedure given for PSO and ABC in the previous subsections is followed.

5. Numerical Applications

In this study, it is observed that, PSO and ABC algorithms determined optimal parameters of PID controller that can drive the DC motor to the reference trajectory with a satisfactory performance. PID tuning strategy based on those algorithms has been growing for a while and today it is a quite mature topic in the associated scientific community. Keeping that in mind, this study proposes

the idea of adopting MO algorithm for tuning PID controller of a DC motor. As will be seen from the results of numerical applications in the following subsections, MO is proven to be a good alternative to the existing state-of-the-art methods in the scheme of PID tuning. In the applications, two different reference signals were considered. The first reference is a step signal with magnitude of 45° while the second one is a combination of sine and square waves. Note that, in the simulations, search space for the PID parameters is initially defined in $[0, 150]$.

5.1. Tracking of 45° Reference

Tracking of the first reference trajectory is given in Fig. 3.

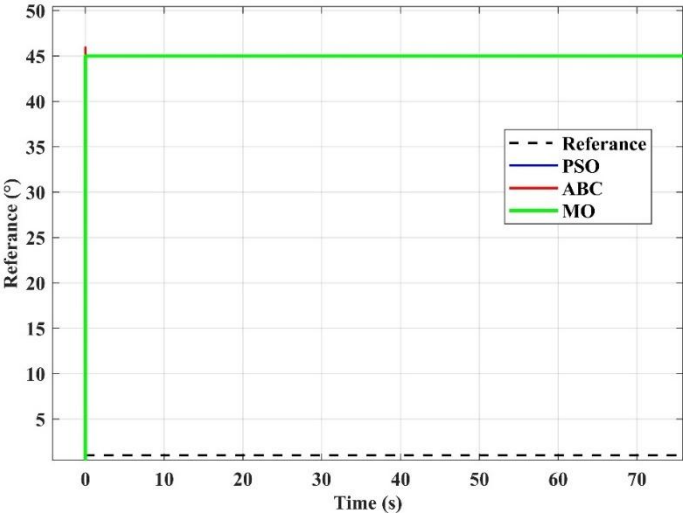


Fig. 3. Implementation of DC motor position control with a reference of 45° .

Fig. 3 shows that all algorithms can perform well in the task of PID parameter tuning, thus the optimally tuned controller provided by them can yield a satisfactory level of trajectory tracking performance. However, there is overshoot in the position of DC motor in the case of PID tuning by PSO and ABC. Initial transient in system output is zoomed and plotted in Fig. 4.

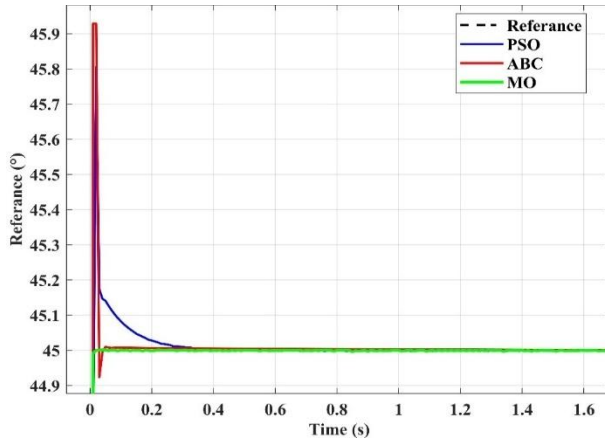


Fig. 4. Initial transient of the DC motor position while tracking 45° reference.

In Fig. 4, it is evident that, MO algorithm-based parameter tuning results in a PID controller which drives the motor position to the desired reference i) faster (settling time is minimum) and ii) no overshoot exists (smoother trajectory tracking). Hence, MO-tuned PID controller gives superior tracking performance compared to the case of PSO and ABC. The resulting K_p , K_i , K_d parameters and tracking RMSE are given in Table 3.

Table 3. K_p , K_i , K_d parameters obtained by each algorithm for 45° reference tracking in addition to tracking RMSE.

	K_p	K_i	K_d	RMSE
PSO	52.73	562.72	Kd=0.01	43.9970
ABC	164.66	187.99	Kd=0.01	43.9975
MO	1150.74	1500.00	Kd=0.19	43.9963

Considering the RMSE values, it can be seen that, MO algorithm gives the best performance followed by PSO and then ABC. Convergence curve shows the rate of minimization of the cost function over iterations. The convergence curves of the optimization algorithms are shown in Fig. 5.

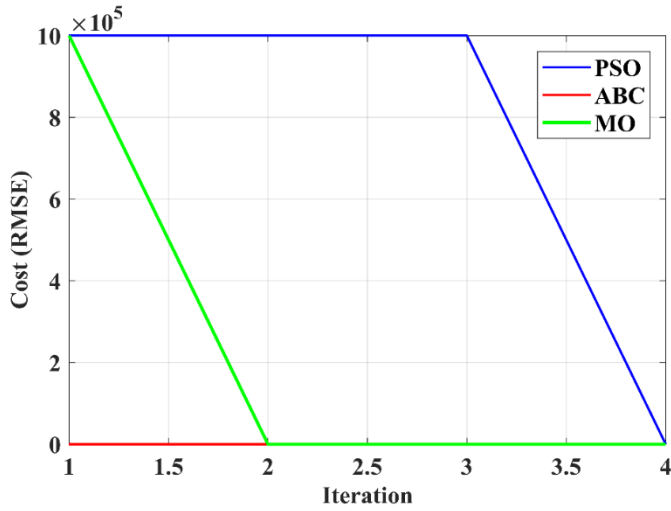


Fig. 5. Convergence curve of optimization algorithms for 45° reference signal.

5.2. Tracking of Complex Reference

As the second numerical application, a reference signal that is combination of sinusoidal and square waves has been considered to demonstrate the PID tuning capability of MO in a relatively complex reference case. The total simulation time is 80 seconds. The first 30 seconds feature a sine wave. Between 30 and 60 seconds, a square wave is generated. The sine and square waves vary within a 60-degree range. The DC motor's position is expected to follow these signals throughout the simulation. According to the results, PSO, ABC and MO yielded acceptable tracking performances, which can be seen in Fig. 6.

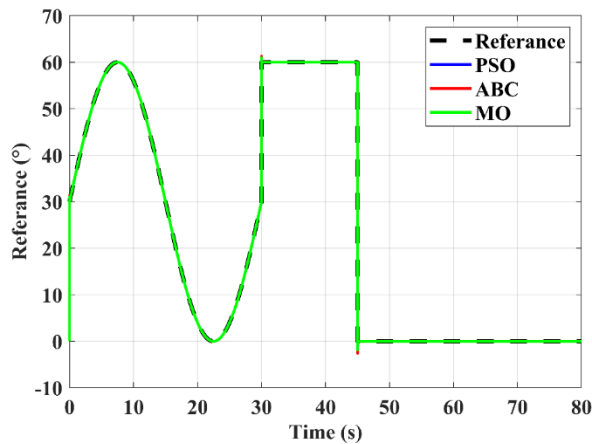


Fig 6. Implementation of DC motor position control in the case of complex reference signal.

In Fig. 6, two overshoots are observed at the point where the square wave reaches 60 degrees first and at the point where the square wave reaches zero. Zoomed plot for the first overshoot is given in Fig. 7 while zoomed plot for the second one is given in Fig. 8.

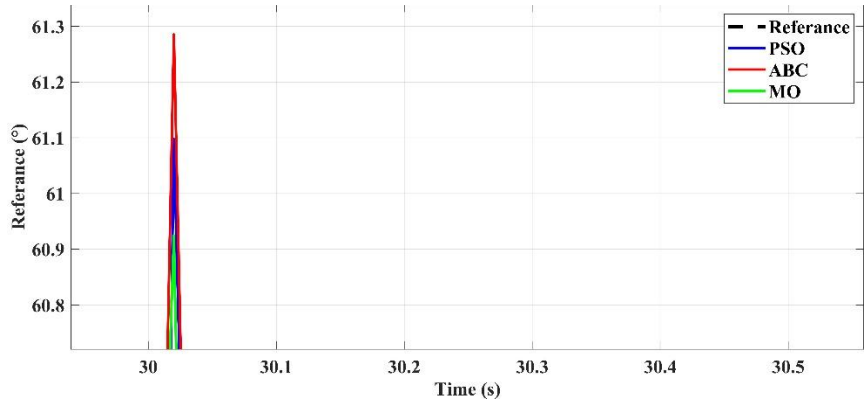


Fig. 7. First overshoot at 60° in the case of complex reference signal.

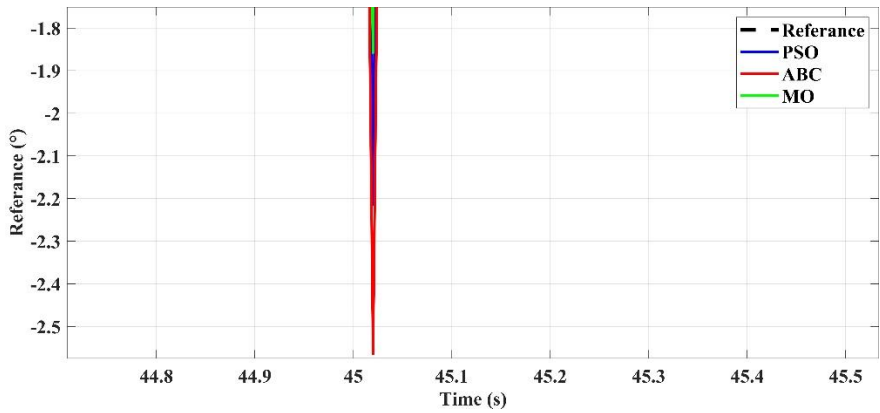


Fig. 8. Second overshoot at 0° in the case of complex reference signal.

It is observed that MO-tuned PID controller gives the least overshoot and thus, the smoothest tracking of the reference. Tracking RMSE and K_p , K_i and K_d parameters obtained by each algorithm are given in Table 4 in the case of complex reference signal. Also, convergence curves of the optimization algorithms in the case of complex reference signal are shown in Fig. 9.

Table 4. K_p , K_i , K_d parameters obtained by each algorithm for tracking of complex reference signal in addition to tracking RMSE.

	K_p	K_i	K_d	RMSE
PSO	23.1038	0.0000	0.0000	0.4755
ABC	22.4472	45.8871	0.0000	0.4758
MO	56.6977	0.0000	1.1033	0.4751

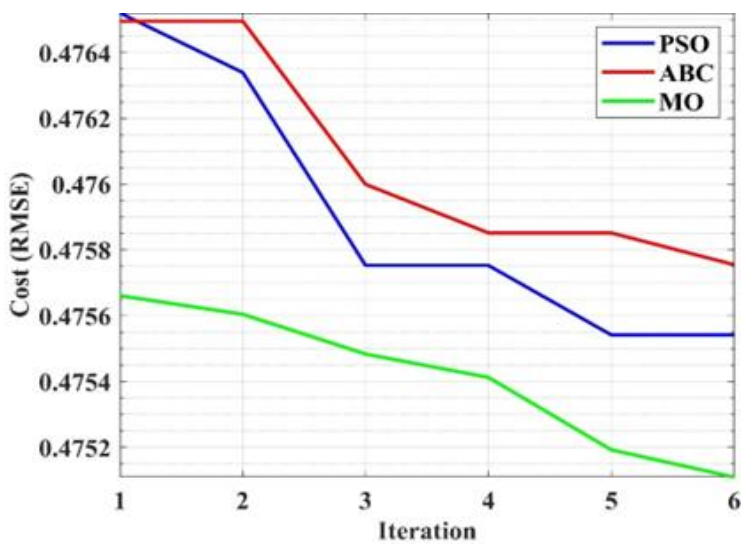


Fig. 9. Convergence curves of algorithms for sine and square wave references.

6. Conclusion

In this study, first the transfer function of a DC motor equipped with an encoder, Faulhaber model 2342I012CR, is obtained. Then, this transfer function is used to simulate the DC motor in numerical application of PID position control where the controller parameters are optimally tuned by PSO, ABC and MO algorithms. The numerical applications were performed using two different types of references. The proposed idea of adopting MO algorithm for PID tuning produced superior DC motor position control performance compared to the well proven PSO and ABC. MO-tuned PID controller resulted in the least overshoot, least settling time and least tracking RMSE. Obtained results demonstrate that, MO algorithm can be used a good alternative in PID tuning task. Given its success in this study, MO algorithm can be used to solve optimization problems in different disciplines. Scientists in different fields can employ MO optimization in different types of engineering problems.

7. References

- Aström, K. J. (1995). PID controllers: theory, design, and tuning. *The international society of measurement and control*.
- Aström, K. J., and Wittenmark, B. (2013). *Computer-controlled systems: theory and design*. Courier Corporation.
- Bequette, B. W. (2003). *Process control: modeling, design, and simulation*. Prentice Hall Professional.
- Brown, W. (2002). Brushless dc motor control made easy. *Microchip Technology Inc*, 1.
- Camacho, E. F., and Bordons, C. (1999). Model predictive control (advanced textbooks in control and signal processing).
- Clerc, M., and Kennedy, J. (2002). The particle swarm-explosion, stability, and convergence in a multidimensional complex space. *IEEE transactions on Evolutionary Computation*, 6(1), 58-73.
- Çakar, Y., Orman, Y., and Kizir, S. (2020). Design Of A Model Reference Adaptive Pid Controller For Dc Motor Position Control: Compared With Pid And Fuzzy Controllers. *Mugla Journal of Science and Technology*, 6, 25-35.
- Doruk, R. Ö., and Zuglem, A. (2018). Geri adımlama tekniği ile bir DC motorun konum ve hız kontrolü. *Süleyman Demirel Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 22, 65-82.
- Fadhil, H. (2025). Metaheuristic Algorithms in Optimization and its Application: A Review. Proceedings of the 3rd International Conference on Engineering and Innovative Technology,
- Franklin, G. F., Powell, J. D., Emami-Naeini, A., and Powell, J. D. (2002). *Feedback control of dynamic systems* (Vol. 4). Prentice hall Upper Saddle River.
- Gandomi, A. H., Yang, X.-S., Talatahari, S., and Alavi, A. H. (2013). Metaheuristic algorithms. *Metaheuristic applications in structures and infrastructures*, 1-24.
- Kaplan, K., Kuncan, M., Polat, H., Tepe, B., and Ertunç, H. M. (2020). PID ve bulanık mantık tabanlı DC motorun gerçek zamanlı konum kontrolü. *Journal of the Institute of Science and Technology*, 10(2), 900-916.
- Karaboga, D. (2005). An idea based on honey bee swarm for numerical optimization.
- Karaboga, D., and Basturk, B. (2007). A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *Journal of global optimization*, 39(3), 459-471.

- Kareem, S. W., Ali, K. W. H., Askar, S., Xoshaba, F. S., and Hawezi, R. (2022). Metaheuristic algorithms in optimization and its application: A review. *JAREE (Journal on Advanced Research in Electrical Engineering)*, 6(1).
- Kennedy, J., and Eberhart, R. (1995). Particle swarm optimization. Proceedings of ICNN'95-international conference on neural networks,
- Lange, K. (2013). *Optimization* (Vol. 95). Springer Science and Business Media.
- Matsui, N., and Shigyo, M. (2002). Brushless DC motor control without position and speed sensors. *IEEE Transactions on Industry Applications*, 28(1), 120-127.
- Müftü, S., and Gökçe, B. (2024). Design and Implementation of an Optimized PID Controller for Two-Limb Robot Arm Control. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 13(1), 192-204.
- Ogata, K. (2010). *Modern control engineering*. Prentice hall.
- Passino, K. M., Yurkovich, S., and Reinfrank, M. (1998). *Fuzzy control* (Vol. 42). Addison-wesley Reading, MA.
- Pedregal, P. (2004). *Introduction to optimization* (Vol. 46). Springer.
- Seborg, D. E., Edgar, T. F., Mellichamp, D. A., and Doyle III, F. J. (2016). *Process dynamics and control*. John Wiley and Sons.
- Şahal, S., and Ortatepe, Z. (2025). Comparative Analysis of PID Tuning Methods for DC Motor Speed Control. 2025 9th International Symposium on Innovative Approaches in Smart Technologies (ISAS),
- Şahin, A. K., Akyazı, Ö., Sahin, E., and Çakır, O. (2021). DC Motorun hız kontrolü için meta-sezgisel algoritma tabanlı PID denetleyici tasarımı. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi*, 10(2), 533-549.
- Ting, T., Yang, X.-S., Cheng, S., and Huang, K. (2014). Hybrid metaheuristic algorithms: past, present, and future. *Recent advances in swarm intelligence and evolutionary computation*, 71-83.
- Yang, X.-S. (2011a). Metaheuristic optimization. *Scholarpedia*, 6(8), 11472.
- Yang, X.-S. (2011b). Metaheuristic optimization: algorithm analysis and open problems. International symposium on experimental algorithms,
- Yaren, T., and Kizir, S. (2022). Model Öngörülü Kontrol ile DA Motor Konum Kontrolü. *Duzce University Journal of Science and Technology*, 10(3), 1151-1164.
- Yuan, Y., Chong, G., Ren, J., Zhao, W., Li, Y., Wang, Z., and Mirjalili, S. (2025). Musk ox optimizer (MO): a novel optimization algorithm and its application. *Cluster Computing*, 28(16), 1041.

APPENDIX

A1. PSO Pseudocode

Pseudocode of PSO algorithm is given as follows.

```
1  // Inputs: swarm size  $N$ , dimension  $D$ ,
2  // max iterations  $k_{max}$ 
2  // Bounds:  $x_{min,j} \leq x_{ij} \leq x_{max,j}$ 
3
4  for ( $i = 1$ ;  $i \leq N$ ;  $i++$ ) {
5      initialize  $x_i^0$  uniformly within bounds;
6      initialize  $v_i^0$  with small random values;
7       $p_{best,i} = x_i^0$ ;
8  }
9   $g_{best} = \text{argmin}_i f(p_{best,i})$ ;
10  $k = 0$ ;
11
12 while ( $k < k_{max}$ ) {
13     for ( $i = 1$ ;  $i \leq N$ ;  $i++$ ) {
14         draw  $r_1, r_2 \in [0, 1]$ ;
15          $v_i^{k+1} = w * v_i^k + c_1 * r_1 * (p_{best,i} - x_i^k) +$ 
            $c_2 * r_2 * (g_{best} - x_i^k)$ ;
16          $x_i^{k+1} = x_i^k + v_i^{k+1}$ ;
17         if ( $f(x_i^{k+1}) < f(p_{best,i})$ )
18              $p_{best,i} = x_i^{k+1}$ ;
19     }
20     for ( $i = 1$ ;  $i \leq N$ ;  $i++$ ) {
21         if ( $f(p_{best,i}) < f(g_{best})$ )
22              $g_{best} = p_{best,i}$ ;
23     }
24      $k++$ ;
25 }
26 return  $g_{best}$ ;
```

Algorithm 1. Pseudocode of PSO.

A2. ABC Pseudocode

Pseudocode of ABC algorithm is given as follows.

```
1  // Inputs: number of food sources  $SN$ , dimension  $D$ ,
2  // max cycles  $MCN$ , limit  $limit$ 
   // Bounds per dimension:  $x_{min,j} \leq x_{ij} \leq x_{max,j}$ 
   (j=1.. $D$ )
3  // Objective: minimize  $f(x)$ 
4
5  for (i = 1; i <=  $SN$ ; i++) {
6      initialize  $x_i \sim U(x_{min}, x_{max})$ ;
7       $trial_i = 0$ ;
8  }
9
10  $best = \operatorname{argmin}_i f(x_i)$ ;
11  $cycle = 1$ ;
12
13 while ( $cycle \leq MCN$ ) {
14     // Employed bee phase
15     for (i = 1; i <=  $SN$ ; i++) {
16         choose  $k$  randomly,  $k \neq i$ ; choose di-
17         mension  $j$  in  $\{1..D\}$ ;
18         draw  $\phi_{ij} \sim U(-1, 1)$ ;
19          $v_{ij} = x_{ij} + \phi_{ij} * (x_{ij} - x_{kj})$ ;
20         construct  $v_i$  by replacing component  $j$ 
21         with  $v_{ij}$ ;
22         // clamp  $v_i$  to bounds
23         for (d = 1; d <=  $D$ ; d++) {
24             if ( $v_{id} < x_{min,d}$ )
25                  $v_{id} = x_{min,d}$ ;
26             else if ( $v_{id} > x_{max,d}$ )
27                  $v_{id} = x_{max,d}$ ;
28             else { /* keep  $v_{id}$  */ }
29         }
30         // Greedy selection
31         if (  $f(v_i) < f(x_i)$  ) {
32              $x_i = v_i$ ;  $trial_i = 0$ ;
33         } else {  $trial_i++$ ; }
34     }
35
36     // Compute probabilities for onlookers
37     for (i = 1; i <=  $SN$ ; i++) {
38          $fit_i = 1 / (1 + f(x_i))$ ;
39     }
```

```

36  sumfit = sum_i fiti;
37  for (i = 1; i <= SN; i++) { Pi = fiti/sumfit; }
38
39      // Onlooker bee phase
40  t = 0;
41  while (t < SN) {
42      select i using roulette wheel based on
         Pi;
43      choose k randomly, k != i; choose di
         mension j; draw  $\phi_{ij}$ ;
44      generate neighbor vi (same as employed
         phase);
45      if ( f(vi) < f(xi) ) { xi=vi; triali=0; }
46      else { triali++; }
47      t++;
48  }
49
50      // Scout bee phase
51  for (i = 1; i <= SN; i++) {
52      if (triali >= limit) {
53  xi = xmin + rand(0,1)*(xmax - xmin);
54  triali = 0;
55      } else { /* keep xi */ }
56  }
57
58      // Update global best
59  candidate = argmin_i f(xi);
60  if ( f(candidate) < f(best) ) { best = candidate; }
61  else { /* keep best */ }
62      cycle++;
63  }
64  return best;

```

Algorithm 2. Pseudocode of ABC.

A3. MO Pseudocode

Pseudocode of MO algorithm is given as follows.

```
1  // Inputs: population size  $N$ , dimension  $D$ , max
    iterations  $k_{max}$ 
2  // Parameters: wintering factor  $\alpha$ , foraging
    factor  $\beta$ 
3  // Bounds:  $x_{min,j} \leq x_{ij} \leq x_{max,j}$ 
4  // Objective: minimize  $f(x)$ 
5
6  for (i = 1; i <=  $N$ ; i++) {
7      initialize  $x_i^0 \sim U(x_{min}, x_{max})$ ;
8  }
9
10  $x_{best} = \text{argmin}_i f(x_i^0)$ ;
11  $k = 0$ ;
12
13 while ( $k < k_{max}$ ) {
14     for (i = 1; i <=  $N$ ; i++) {
15         draw  $r \sim U(0,1)$ ;
16         if (  $r < 0.5$  ) { // wintering (defense)
17              $x_i^{k+1} = x_i^k + \alpha * r * (x_{best}^k - x_i^k)$ ;
18         } else { // foraging (exploration)
19             select random  $x_{rand}^k$  from population;
20              $x_i^{k+1} = x_i^k + \beta * r * (x_{rand}^k - x_i^k)$ ;
21         }
22         // Boundary handling
23         for (j = 1; j <=  $D$ ; j++) {
24             if ( $x_{ij}^{k+1} < x_{min,j}$ ) {  $x_{ij}^{k+1} = x_{min,j}$ ; }
25             else if ( $x_{ij}^{k+1} > x_{max,j}$ ) {  $x_{ij}^{k+1} = x_{max,j}$ ; }
26         }
27     }
28     // Update global best
29      $candidate = \text{argmin}_i f(x_i^{k+1})$ ;
30     if (  $f(candidate) < f(x_{best})$  ) {  $x_{best} = candidate$ ; }
31      $k++$ ;
32 }
33 return  $x_{best}$ ;
```

Algorithm 3. Pseudocode of MO.

CHAPTER 21

AI in Combating Digital Fraud in Banking: Cases, Signals, and Methods

Ali Keleş¹ & Aytürk Keleş²

¹ Assoc. Prof. Dr., Ağrı İbrahim Çeçen University, Faculty of Education, Department of Computer and Education Technologies Education, ORCID: 0000-0001-9755-295X

² Assoc. Prof. Dr., Ağrı İbrahim Çeçen University, Faculty of Education, Department of Computer and Education Technologies Education, ORCID: 0000-0002-0785-9593

1. Introduction

The banking sector has undergone a profound transformation driven by digitalization and the widespread adoption of real-time payment infrastructures (Gomber et al., 2018; BIS, 2020). Mobile banking, internet banking, card-based payment systems, instant transfer mechanisms, and open banking integrations have significantly enhanced customer convenience and financial accessibility (Arner et al., 2017). At the same time, these developments have expanded the attack surface available to fraudsters, exposing financial institutions to increasingly sophisticated and fast-evolving fraud threats (Kahn and Roberds, 2009; Dal Pozzolo et al., 2018). As a result, banks are required to deploy fraud detection mechanisms capable of operating under high transaction volumes, low latency constraints, and strict reliability requirements in order to preserve operational continuity and customer trust (Lucas et al., 2020).

One of the fundamental challenges of fraud detection in banking lies in the extreme class imbalance between legitimate and fraudulent transactions (Dal Pozzolo et al., 2015). Fraud cases typically represent only a tiny fraction of overall transaction volumes, while legitimate user behavior exhibits substantial diversity (Bolton and Hand, 2002). Fraudsters actively exploit this diversity by mimicking normal transaction patterns, thereby bypassing static rule-based controls (Fawcett and Provost, 1997). Moreover, many fraud incidents do not manifest as isolated anomalous transactions but rather unfold as multi-stage processes, including account takeover, credential manipulation, beneficiary creation, high-value transfers, and rapid cash-out activities (Dal Pozzolo et al., 2018; Bahnsen et al., 2015). This process-oriented nature of fraud significantly complicates detection based solely on individual transaction attributes (Whitrow et al., 2009).

Decision errors impose asymmetric costs, which further complicate fraud detection in banking (Bahnsen et al., 2013). False negatives, in which fraudulent activities go undetected, lead to direct financial losses, reputational damage, and regulatory consequences (Kirkos et al., 2007). Conversely, false positives, where legitimate transactions are incorrectly blocked or delayed, negatively affect customer experience, reduce transaction completion rates, and erode trust in digital banking services (Dal Pozzolo et al., 2018). Consequently, modern fraud detection systems must balance sensitivity and precision while also ensuring explainability, operational action ability, and adaptability to emerging fraud patterns (Guidotti et al., 2019).

In this context, machine learning and artificial intelligence (AI)–based approaches have become central to contemporary banking fraud detection systems (Bolton and Hand, 2002; Dal Pozzolo et al., 2015). Supervised learning models are effective in identifying known fraud typologies when sufficient labeled data are available, whereas unsupervised and semi-supervised techniques enable the detection of novel or rare attack patterns through anomaly identification (Chandola et al., 2009). Beyond transaction-level modeling, advanced approaches such as behavioral profiling, sequence-based event modeling, device and session intelligence, and graph-based analysis capture the temporal and relational structure of fraud activities (Akoglu et al., 2015; Van Vlasselaer et al., 2015). As a result, researchers and practitioners increasingly treat fraud detection as a multi-layered risk assessment problem rather than a single-model classification task (Lucas et al., 2020).

In this chapter, fraud attempts that have occurred in the banking sector to date are examined from a holistic perspective, with a focus on the shared behavioral and structural characteristics observed across these incidents, the core challenges encountered in fraud detection processes, and the potential artificial intelligence–based solutions that can be employed to address these challenges. By linking real-world banking fraud cases with methodological insights, this chapter aims to provide a coherent and practically relevant framework to support the design of robust and adaptive fraud detection systems in modern financial institutions.

2. Large-Scale Banking and Financial Fraud Cases

Over the past decade, the banking sector has experienced multiple large-scale fraud and cybercrime incidents that highlight the evolving tactics of adversaries and the limitations of traditional security controls. One of the earliest and most impactful cases was the **Target payment card data breach (2013)**, where attackers compromised point-of-sale (POS) systems to exfiltrate payment card data. This incident revealed the importance of transaction-level monitoring, velocity and geography-based rules, and network segmentation within retail payment environments.

The **GameOver Zeus botnet disruption (2014)** demonstrated how credential-stealing malware combined with botnet infrastructure could enable large-scale banking fraud and mule-based cash-out operations. Detection relied heavily on threat intelligence feeds, malware indicators of compromise (IOCs), abnormal session behavior, and transaction risk scoring, emphasizing the need for strong multi-factor authentication and device-level security controls.

Between 2014 and 2015, the **Carbanak cyber heists** exposed vulnerabilities within internal banking networks. Attackers leveraged spear-phishing and lateral movement to manipulate transaction systems and ATM operations directly. These incidents underscored the necessity of endpoint detection and response (EDR), privileged access management (PAM), and continuous monitoring of internal operator activity.

A landmark example of payment infrastructure abuse was the **Bangladesh Bank SWIFT heist (2016)**, in which malware was used to manipulate SWIFT transfer messages. Anomalous beneficiary creation, out-of-hours activity, and unusual transfer corridors were key signals. This case accelerated the adoption of SWIFT-specific analytics, multi-level approval workflows, and out-of-band verification mechanisms.

Business Email Compromise (BEC) attacks, which have escalated since 2015, exploit email account takeovers and social engineering to induce fraudulent wire transfers. These attacks typically involve urgency cues, beneficiary changes, and broken approval workflows, driving the integration of email security, payment behavior analytics, and cooling-off periods for beneficiary modifications.

Major data breaches such as the **Equifax incident (2017)** enabled downstream identity theft and account-opening fraud by exposing large volumes of personally identifiable information (PII). Detection approaches increasingly combined electronic know-your-customer (eKYC), identity resolution, and graph-based linkage analysis to identify synthetic and compromised identities.

Synthetic identity fraud, particularly prominent after 2015, represents a long-term abuse strategy in which fabricated identities are gradually matured before executing high-impact “bust-out” events. Thin-file characteristics, shared devices, and relationship graphs have become critical signals for detection.

The **Capital One cloud breach (2019)** highlighted the risks associated with cloud misconfigurations and excessive privileges. This case reinforced the role of cloud security posture management, anomaly detection in data access patterns, and least-privilege enforcement.

More recently, **AI-enabled social engineering attacks**, including voice deepfake-based CEO fraud (2019) and deepfake video-conference impersonation (2024), have demonstrated how generative AI can be weaponized to bypass human trust mechanisms. These incidents stress the importance of procedural

controls such as call-back verification, dual authorization, and independent identity confirmation for high-value transactions.

Regulatory and industry reports, such as **ENISA’s SIM-swapping and account takeover analyses (2021)** and investigations into **crypto investment fraud and “pig butchering” scams (2022–2023)**, further illustrate the shift toward multi-channel, cross-platform fraud ecosystems. Consequently, modern fraud detection is increasingly treated as a layered risk management problem integrating transaction analytics, behavioral modeling, graph analysis, and human-in-the-loop controls.

3. Case-Signal-AI Method Mapping Fraud Detection

The case–signal–AI mapping framework structures major banking fraud incidents by explicitly linking fraud case types to their dominant observable signals and the artificial intelligence approaches best suited for their detection. Rather than classifying fraud solely by attack vectors or affected channels, this framework focuses on how fraudulent behavior materializes in data. Accordingly, fraud manifestations are grouped into transaction-level irregularities, behavioral deviations, multi-step temporal processes, and relational network patterns, each requiring distinct analytical treatment.

This perspective reflects the operational reality that many contemporary fraud schemes—such as account takeover, business email compromise, SWIFT-related treasury fraud, and deepfake-assisted social engineering—do not present themselves as isolated anomalous transactions. Instead, they unfold through sequences of actions characterized by changes in device usage, session behavior, beneficiary management, authorization paths, and transaction timing. Detecting such cases therefore necessitates behavioral profiling and sequence-aware models, including time-series analysis, state-transition modeling, and event-sequence learning techniques.

Table 1. Case-Signal-AI Method Mapping for Banking Fraud Detection

Fraud Case Type	Key Observable Signals	Effective AI Approaches
Account Takeover (ATO) SIM Swap	-New device/IP -Session anomalies -Rapid beneficiary & transfer changes	-Behavioral profiling -Sequence models (LSTM/HMM) -Gradient Boosted Trees
Business Email Compromise (BEC)	-New / modified payees -Urgency language -Workflow deviations	-NLP-based email analysis -Payment behavior ML -Rule + ML fusion
SWIFT / Treasury Fraud	-Out-of-hours activity -New corridors -Abnormal amounts	-Sequence-based risk modeling -Supervised ML -Human-in-the-loop AI
Synthetic Identity Fraud	-Thin-file identities -Shared devices/PII -Long warm-up period	-Graph/link analysis -Supervised tabular ML -Behavioral time-series models
Mule Networks & Cash-Out	-Rapid fund dispersion -Shared accounts/devices -Velocity spikes	-Graph ML -Unsupervised anomaly detection -Cluster-based risk scoring
Deepfake Social Engineering	-Abnormal authorization paths -Atypical timing -Identity inconsistencies	-Voice/video deepfake detection -Process anomaly detection -Step-up authentication AI

The mapping further illustrates that fraud types involving coordination across multiple entities, such as mule networks, synthetic identity fraud, and crypto-enabled investment scams, exhibit weak signals at the individual transaction level but strong relational patterns across accounts, devices, and identities. Graph-based learning and link analysis methods are consequently essential for exposing these hidden structures and coordinated activities.

By aligning each fraud case with its characteristic signals and appropriate AI methods, the framework underscores the need for multi-layered detection architectures in modern banking environments. While rule-based and supervised models are effective, they are not sufficient on their own and must be complemented by behavioral, sequential, and relational analyses. Collectively, this case-signal-AI mapping supports a shift away from single-model classification toward integrated risk assessment frameworks that combine multiple AI paradigms with operational controls and human-in-the-loop decision mechanisms to address the scale, complexity, and adaptability of modern financial fraud.

4. Conclusion and Future Work

Fraud detection in the banking sector should be approached as a multi-dimensional and dynamic risk management problem rather than a standalone classification task. Evidence from real-world fraud cases demonstrates that effective detection relies on the integrated analysis of behavioral patterns, temporal attack sequences, and relational structures across accounts, devices, and beneficiaries. Artificial intelligence–based approaches provide the flexibility and scalability required to address this complexity, particularly when deployed within hybrid architectures that combine automated decision-making with human expertise.

Despite these advances, the continuously evolving nature of fraud ecosystems presents ongoing challenges for both research and practice. Future work should prioritize the development of adaptive AI models that are robust to concept drift, capable of learning from limited or delayed labels, and suitable for real-time deployment in high-throughput banking environments. In addition, increasing emphasis should be placed on explainable and transparent AI techniques to support regulatory compliance, operational trust, and human oversight in fraud-related decision processes.

Furthermore, privacy-preserving learning paradigms such as federated learning and secure multi-party computation represent promising research directions for enabling cross-institutional collaboration without compromising sensitive financial data. Finally, fraud detection systems in the future should assess performance not only by predictive accuracy but also by operational effectiveness, interpretability, and long-term sustainability, thereby strengthening resilience and trust in modern digital banking infrastructures.

References

- Akoglu, L., Tong, H., & Koutra, D. (2015). Graph-based anomaly detection and description: A survey. *Data Mining and Knowledge Discovery*, 29(3), 626–688.
- Arner, D. W., Barberis, J., & Buckley, R. P. (2017). FinTech, RegTech, and the reconceptualization of financial regulation. *Northwestern Journal of International Law & Business*, 37(3), 371–413.
- Bahnsen, A. C., Aouada, D., & Ottersten, B. (2013). Example-dependent cost-sensitive decision trees. *Expert Systems with Applications*, 40(18), 7035–7044.
- Bahnsen, A. C., Stojanovic, A., Aouada, D., & Ottersten, B. (2015). Cost-sensitive decision trees for fraud detection. *Expert Systems with Applications*, 42(3), 735–746.
- Bank for International Settlements. (2018). *Reducing the risk of wholesale payments fraud*. Bank for International Settlements.
- Bank for International Settlements. (2020). *Sound practices: Implications of fintech developments for banks and bank supervisors*. Bank for International Settlements.
- Bolton, R. J., & Hand, D. J. (2002). Statistical fraud detection: A review. *Statistical Science*, 17(3), 235–255.
- Chandola, V., Banerjee, A., & Kumar, V. (2009). Anomaly detection: A survey. *ACM Computing Surveys*, 41(3), 1–58.
- Dal Pozzolo, A., Bontempi, G., Snoeck, M., & Brossard, F. (2015). Adaptive machine learning for credit card fraud detection. *IEEE Intelligent Systems*, 30(4), 37–41.
- Dal Pozzolo, A., Boracchi, G., Bontempi, G., & Snoeck, M. (2018). Adversarial drift detection. In *Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN)*.
- ENISA. (2021). *Threat landscape for telecoms: SIM swapping*. European Union Agency for Cybersecurity.
- Europol. (2015). *The Internet Organised Crime Threat Assessment (IOCTA)*.
- Europol. (2023). *Investment fraud and cryptocurrency scams*.
- Fawcett, T., & Provost, F. (1997). Adaptive fraud detection. *Data Mining and Knowledge Discovery*, 1(3), 291–316.

- Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2019). A survey of methods for explaining black box models. *ACM Computing Surveys*, 51(5), 1–42.
- Kahn, C. M., & Roberds, W. (2009). Why pay? An introduction to payments economics. *Journal of Financial Intermediation*, 18(1), 1–23.
- Kirkos, E., Spathis, C., & Manolopoulos, Y. (2007). Data mining techniques for the detection of fraudulent financial statements. *Expert Systems with Applications*, 32(4), 995–1003.
- Krebs, B. (2013). *The Target breach, by the numbers*. Krebs on Security.
- Lucas, Y., Portier, P.-E., Laporte, L., Calabretto, S., He-Guelton, L., Granitzer, M., & Jurgovsky, J. (2020). Adversarial drift detection for fraud detection. In *Proceedings of the 26th ACM SIGKDD Conference*.
- Mandiant. (2015). *Carbanak cybercrime campaign*.
- U.S. Department of Justice. (2014). *GameOver Zeus botnet takedown*.
- Van Vlasselaer, V., Eliassi-Rad, T., Akoglu, L., Snoeck, M., & Baesens, B. (2015). GOTCHA! Network-based fraud detection. *Management Science*, 63(9), 3090–3110.
- Verizon. (2023). *Data Breach Investigations Report*.
- Whitrow, C., Hand, D. J., Juszczak, P., Weston, D., & Adams, N. (2009). Transaction aggregation as a strategy for credit card fraud detection. *Data Mining and Knowledge Discovery*, 18(1), 30–55.

CHAPTER 22

Hydrogen Combustion Technologies: From Engine Applications to Gas Turbines and Safety-Oriented Modeling

Şafak Melih Şenocak¹ & Yasin Varol²

¹ Lecturer, Osmaniye Korkut Ata University, ORCID: 0000-0003-0602-2836

² Prof. Dr., Fırat University, ORCID: 0000-0003-2989-7125

1. Introduction

The increasing global energy demand and the tightening of regulations aimed at reducing greenhouse gas emissions in the context of climate change mitigation have made the investigation of alternative fuels indispensable in the transportation and power generation sectors. In this context, hydrogen has emerged as one of the promising energy carriers of the future due to its carbon free nature, high specific energy content, and broad application potential (Verhelst & Wallner, 2009, White et al., 2006). The fact that hydrogen combustion produces only water vapor provides a significant advantage, particularly in terms of eliminating carbon dioxide and particulate matter emissions at the source (Akal et al., 2020). Figure 1 presents a schematic representation of a representative hydrogen production pathway based on a renewable energy source.

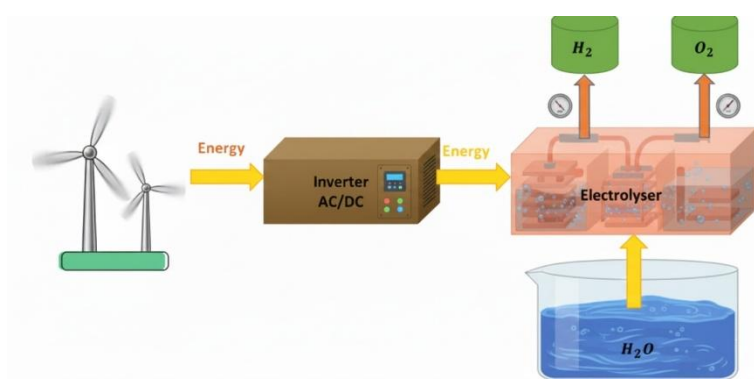


Figure 1. Hydrogen production using wind energy (Created by authors).

Although research on the use of hydrogen in energy systems has largely focused on fuel cells, the direct utilization of hydrogen as a fuel in conventional combustion systems such as internal combustion engines and gas turbines has attracted renewed interest in recent years (White et al., 2006, Stepień, 2021). The main reasons for this renewed interest include the high technological maturity of engine and turbine technologies, the possibility of preserving existing manufacturing and service infrastructures, and the potential to offer more cost effective solutions compared to fuel cells, particularly in heavy duty applications (Verhelst & Wallner, 2009).

However, the combustion characteristics of hydrogen differ significantly from those of conventional hydrocarbon fuels. Its wide flammability limits, extremely low minimum ignition energy, and high laminar flame speed enable hydrogen to offer high efficiency potential in both engines and gas turbines, while simultaneously introducing specific combustion related challenges such as backfire, pre ignition, and thermal NO_x formation (White et al., 2006, Stepień,

2021). Especially under high load operating conditions, combustion control and emission management remain among the primary technical barriers to the widespread adoption of hydrogen based systems (Yip et al., 2019).

The use of hydrogen in internal combustion engines has been investigated through various approaches, including spark ignition, compression ignition, and dual fuel configurations. Fuel delivery strategies such as port fuel injection and direct injection, together with combustion phasing and load control methods, have been extensively examined in the literature as the key parameters governing the performance and emission characteristics of hydrogen fueled engines (Verhelst & Wallner, 2009, Akal et al., 2020). In a similar manner, hydrogen combustion in gas turbines requires dedicated combustion concepts and advanced mixing techniques due to specific challenges such as flame flashback, autoignition, and thermoacoustic instabilities (White et al., 2006).

These technical challenges have made it necessary to investigate hydrogen combustion in detail using both experimental and numerical methods. Chemical kinetic mechanisms, laminar flame speed correlations, and computational fluid dynamics based models play a critical role in understanding hydrogen combustion behavior in engines, gas turbines, and safety related scenarios (Stepien, 2021). In particular, the validation of modeling approaches employed for emission prediction and combustion stability analysis continues to be an important research focus in the literature.

This book chapter aims to address hydrogen combustion technologies from a holistic perspective, with a focus on internal combustion engines and gas turbines. The fundamental combustion properties of hydrogen, its combustion behavior in engine and turbine applications, NO_x formation mechanisms, and numerical modeling approaches are reviewed and discussed comparatively based on the current literature. In this way, the existing body of knowledge regarding the safe, efficient, and low emission use of hydrogen in conventional combustion systems is presented, and future research needs are identified.

2. Fundamental Combustion Characteristics of Hydrogen

The combustion behavior of hydrogen differs markedly from that of conventional hydrocarbon fuels, and these differences directly influence both design and operating strategies in internal combustion engines and gas turbines. The wide flammability limits of hydrogen, its low minimum ignition energy, and its high laminar flame speed provide high combustion efficiency and enable ultra lean operation. At the same time, these properties give rise to specific challenges such as pre ignition, backfire, and elevated NO_x formation associated with high combustion temperatures (Verhelst & Wallner, 2009, White et al., 2006).

Hydrogen air mixtures exhibit flammability over an approximate volumetric concentration range of 4 to 75 percent. This range is exceptionally broad compared with gasoline and diesel fuels and allows hydrogen to sustain stable combustion even under very lean mixture conditions (Akal et al., 2020). The reduction in combustion temperature under lean operation constitutes a major advantage in terms of suppressing thermal NO_x formation. This characteristic enhances the attractiveness of hydrogen based systems, particularly under increasingly stringent emission regulations (Stepien, 2021).

The extremely low minimum ignition energy of hydrogen, on the order of 0.02 mJ, facilitates easy initiation of combustion. However, it also increases the likelihood of undesired pre ignition events caused by hot surfaces or residual gases. This effect is especially pronounced in systems employing port fuel injection, where the presence of a premixed combustible charge in the intake system increases the risk of backfire (White et al., 2006). In addition, the short quenching distance of hydrogen enables flame propagation through narrow geometries, thereby increasing the probability of ignition near valves and confined regions within the combustion chamber (Verhelst & Wallner, 2009).

The laminar flame speed is one of the most distinctive features of hydrogen combustion. Under stoichiometric conditions, the laminar flame speed of hydrogen is approximately an order of magnitude higher than that of gasoline or diesel fuels. This high flame speed shortens the combustion duration and reduces heat losses, contributing to improved thermal efficiency. At the same time, the rapid rate of heat release requires careful control of combustion phasing, particularly under high load operating conditions (White et al., 2006).

The low density and limited volumetric energy content of hydrogen lead to a reduction in volumetric efficiency, especially in premixed systems. This limitation resulted in reduced power density in early hydrogen engine studies and subsequently motivated the development of strategies such as direct injection to mitigate this drawback (Verhelst & Wallner, 2009, Yip et al., 2019). In gas turbine applications, these properties make the control of fuel air mixing ratios and mixture homogeneity critical from a design perspective.

These fundamental combustion characteristics demonstrate that hydrogen has substantial potential as a fuel for both internal combustion engines and gas turbines. However, the safe, stable, and low emission utilization of this potential depends on the systematic management of the associated combustion phenomena. In the following sections, the combustion behavior of hydrogen in engine and turbine applications is examined in detail within the context of these fundamental characteristics. Table 1 summarizes the key properties of hydrogen in comparison with gasoline and diesel fuels.

Table 1. Hydrogen and Conventional Fuels: Fundamental Combustion Properties (White et al. 2006; Verhelst & Wallner 2009; Akal et al. 2020).

Property	Hydrogen (H ₂)	Gasoline	Diesel
Flammability limits (vol. %)	4-75	1-7.6	0.6-5.5
Minimum ignition energy (mJ)	0.02	0.24	0.25
Laminar flame speed (m/s, $\phi \approx 1$)	2.0-3.0	0.3-0.4	0.3
Auto-ignition temperature (°C)	585	280	210
Quenching distance (mm)	0.6	2.0	2.0
Carbon-based emissions	None	CO ₂ , CO, HC	CO ₂ , CO, soot

3. Hydrogen Combustion in Internal Combustion Engines

The use of hydrogen in internal combustion engines has been extensively investigated in the literature within the framework of different combustion concepts and fuel supply strategies. While the high flame speed and wide flammability limits of hydrogen provide significant advantages in terms of engine efficiency, its low minimum ignition energy and high combustion temperature require careful control with respect to combustion stability and NO_x formation (White et al., 2006, Stepien, 2021).

In internal combustion engine applications, hydrogen has primarily been considered in spark ignition, compression ignition, and dual fuel configurations. Early studies demonstrated that premixed hydrogen engines employing port fuel injection can achieve high efficiency under lean mixture conditions. However, in such configurations, backfire and pre ignition phenomena have been identified as major factors limiting engine operation, particularly under high load conditions (White et al., 2006).

To overcome these limitations, direct injection strategies have been developed in which hydrogen is introduced into the cylinder at a late stage of the compression stroke. This approach prevents the formation of a combustible mixture in the intake system and significantly reduces volumetric efficiency losses. In direct injection hydrogen engines, more precise control of combustion phasing has enabled increased power density while suppressing the tendency for pre ignition (Verhelst & Wallner, 2009, Yip et al., 2019). Consequently, direct injection has emerged as the dominant solution in contemporary hydrogen engine research.

The direct use of hydrogen in compression ignition engines is limited due to its high autoignition temperature. For this reason, hydrogen is often considered in combination with a pilot fuel in dual-fuel configurations. Reactivity controlled compression ignition (RCCI) is an advanced combustion concept that uses a low-reactivity fuel premixed through the intake and a high-reactivity fuel directly injected near top dead center to control combustion phasing. RCCI studies commonly employ hydrocarbon fuels such as gasoline and diesel rather than hydrogen. These studies still provide a useful reference point because they illustrate how reactivity stratification can reduce soot, CO, and HC emissions while potentially increasing NO_x due to higher peak temperatures (Firat et al., 2022).

Recent work has also examined oxidant-assisted RCCI, such as ozone seeding, to influence low-temperature chemistry and combustion initiation. Experiments with diesel and iso-octane in RCCI have shown that ozone can reduce HC and smoke opacity and may reduce NO_x under specific operating conditions (Firat et al., 2023). More broadly, additive-assisted and multi-fuel approaches in compression ignition engines, including studies with oxygenated additives, demonstrate that modifying fuel chemistry can shift combustion and emission behavior. This is relevant when assessing future pathways for integrating hydrogen as an additional energy carrier in multi-fuel systems (Öztürk et al., 2024).

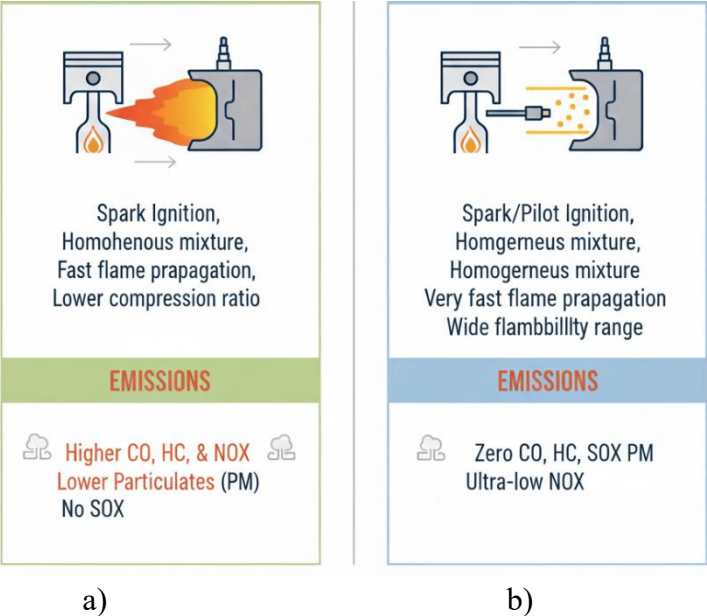


Figure 2. Emission comparison between a) gasoline and b) hydrogen (Created by authors).

In RCCI engines, combustion control is achieved by introducing the low reactivity fuel through the intake system while injecting the high reactivity fuel directly into the cylinder. Experimental RCCI studies commonly employ a low reactivity fuel such as gasoline together with a high reactivity fuel such as diesel, and report that RCCI operation can reduce pollutant emissions including NO_x and soot compared to conventional diesel operation, while also improving thermal efficiency (Firat et al., 2022).

One of the recent approaches proposed to improve combustion control and emission performance in RCCI engines is ozone assisted operation. Ozone addition can accelerate low temperature chemistry, shorten ignition delay, and promote a more controllable combustion onset. In an experimental RCCI study using diesel as the high reactivity fuel and iso octane as the low reactivity fuel, ozone seeding in the intake air was reported to reduce HC and smoke opacity, and to reduce NO_x under specific operating conditions (Firat et al., 2023). These findings indicate that ozone assisted strategies can be a practical option for addressing combustion control challenges and emission deterioration in RCCI operation (Firat et al., 2023).

Similarly, the effects of alternative fuel additives on combustion behavior have also been investigated in the literature. Oxygenated fuel additives such as 1,2 dimethoxy ethane have been shown to influence combustion rates and emission formation in diesel engines. These studies indicate that evaluating hydrogen not only as a single fuel but also within multi fuel or additive assisted systems can provide new opportunities for improved combustion control (Öztürk et al., 2024).

Overall, the use of hydrogen in internal combustion engines is strongly dependent on the fuel delivery method, combustion concept, and auxiliary control strategies. While direct injection spark ignition engines offer high power density, RCCI and dual fuel systems provide advantages in terms of emission control. Nevertheless, the trade off between NO_x formation and combustion stability remains a key research challenge in the widespread implementation of hydrogen fueled engines.

Table 2. Combustion Strategies in Internal Combustion Engines (White et al. 2006; Verhelst & Wallner 2009)

Engine concept	Hydrogen role	Main benefits	Main limitations
SI - Port fuel injection	Primary fuel (premixed)	Lean combustion, low CO/HC	Backfire, early ignition, low power density
SI - Direct injection	Primary fuel (late injection)	High power density, stable combustion	Injector complexity, NO _x at high load
CI - Dual fuel	Secondary fuel	Soot and CO ₂ reduction	NO _x increase, limited H ₂ ratio

4. Hydrogen Combustion in Gas Turbines

Gas turbines present different physical and chemical challenges compared to internal combustion engines when hydrogen is considered for high power density systems requiring continuous combustion. While combustion in internal combustion engines occurs in a time dependent manner within a closed volume, combustion in gas turbines takes place under steady flow conditions in an open system. This fundamental difference causes the properties of hydrogen, such as its high flame speed and low ignition energy, to have more critical consequences in gas turbine combustors (Lieuwen & Yang, 2005).

The use of hydrogen in gas turbines has been intensively investigated in recent years, particularly in the context of low carbon energy production and fuel flexible concepts. The conversion of natural gas turbines to operate with hydrogen offers a significant advantage in terms of preserving existing energy infrastructure. However, the combustion characteristics of hydrogen generally prevent the direct use of conventional natural gas combustor designs without substantial modifications (Chiesa et al., 2005; Lefebvre & Ballal, 2010). A simplified illustration of hydrogen combustion in a gas turbine is shown in Figure 3.

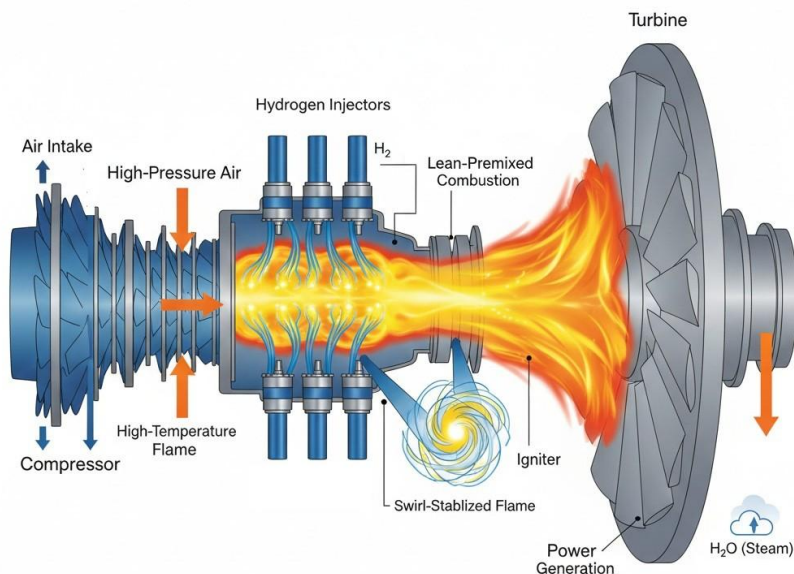


Figure 3. Schematic representation of hydrogen combustion in a gas turbine (Created by authors).

The most fundamental challenge associated with hydrogen combustion in gas turbines is the tendency for flashback. Due to the high laminar flame speed of hydrogen, the flame may propagate upstream against the flow direction in regions where the local flow velocity decreases. In lean premixed combustors, flame propagation toward the injector or mixing passages can result in severe structural damage and significant safety risks (Lieuwen & Yang, 2005). Consequently, understanding and preventing flashback mechanisms constitutes a central aspect of hydrogen gas turbine design.

Autoignition represents another critical issue. In gas turbines, hydrogen exposed to compressed and high temperature air may undergo chemical ignition before a stable flame is established, particularly in systems characterized by long mixing times. This behavior necessitates careful selection of combustor geometry and fuel injection locations (Lefebvre & Ballal, 2010). In addition, the low minimum ignition energy of hydrogen makes autoignition more likely in the presence of hot surfaces and localized temperature increases induced by turbulence.

To address these challenges, several advanced combustion concepts have been developed in the literature. Micro mixing based combustors promote rapid mixing of hydrogen and air over very short distances using small scale jets, thereby preventing the formation of extended premixed regions. This approach

eliminates long mixing channels susceptible to upstream flame propagation and significantly reduces flashback risk (Cecere et al., 2023, Lieuwen & Yang, 2005).

In addition, dilution based combustion concepts are widely considered for hydrogen fueled gas turbines. By introducing steam, nitrogen, or recirculated combustion products into the combustor, the flame temperature is reduced, which suppresses both flashback tendencies and thermal NO_x formation. However, such approaches require careful optimization due to their potential impact on turbine efficiency and overall system complexity (Lefebvre & Ballal, 2010).

Another important aspect of hydrogen combustion in gas turbines is thermoacoustic instability. The fast chemical kinetics of hydrogen can lead to coupling between heat release oscillations and acoustic pressure waves, resulting in high amplitude pressure fluctuations within the combustor. These oscillations increase the risk of mechanical damage and limit stable operating ranges. The literature emphasizes that stable operation of hydrogen gas turbines requires an integrated consideration of combustor geometry, fuel injection configuration, and flow field characteristics (Lieuwen & Yang, 2005).

NO_x emissions constitute a key parameter in assessing the environmental performance of hydrogen gas turbines. The high adiabatic flame temperature of hydrogen makes thermal NO_x formation unavoidable under certain operating conditions. For this reason, hydrogen gas turbines are almost exclusively operated in lean premixed or diluted combustion regimes. Under appropriately designed conditions, NO_x emissions from hydrogen fueled gas turbines can be reduced to levels comparable to those of natural gas turbines, as reported in the literature (Correa, 1993, Lefebvre & Ballal, 2010).

In conclusion, the use of hydrogen in gas turbines offers substantial technical potential, but it requires a comprehensive treatment of specific challenges including combustion stability, flashback, autoignition, and thermoacoustic instabilities. Compared to engine applications, hydrogen gas turbines rely more heavily on advanced combustion concepts, numerical modeling, and safety oriented design strategies to ensure reliable and low emission operation.

Table 3. Hydrogen Combustion Challenges and Mitigation Strategies in Gas Turbines

Issue	Physical origin	Impact on operation	Mitigation strategies	Key references
Flashback	High laminar flame speed, low flow velocity regions	Injector and combustor damage	Micro-mixing, high-velocity jets, geometric isolation	Correa (1993); Cecere et al. (2023)
Autoignition	High mixture temperature, long residence time	Uncontrolled ignition, instability	Reduced mixing time, staged injection	Lefebvre & Ballal (2010)
Thermo-acoustic instability	Fast heat release coupling with pressure waves	Structural damage, noise	Combustor geometry optimization, damping	Lieuwen & Yang (2005)
Thermal NO _x	High flame temperature	Environmental impact	Lean premixed combustion, dilution	Correa (1993); Lefebvre & Ballal (2010)

5. NO_x Formation and Mitigation in Hydrogen Combustion

The use of hydrogen in internal combustion engines and gas turbines offers a significant advantage in terms of reducing carbon based emissions. However, the formation of nitrogen oxides remains the primary limiting factor governing the environmental performance of these systems. Since hydrogen does not contain carbon, carbon dioxide formation is eliminated at the source. Nevertheless, under high temperature conditions, the oxidation of atmospheric nitrogen leads to the formation of NO and NO₂ (Wróbel et al., 2022, Teoh et al., 2023). For this reason, the clean objective of hydrogen combustion technologies is, in practice, largely defined by effective NO_x management (Stepien, 2021).

In hydrogen combustion, NO_x formation is predominantly associated with thermal mechanisms because the fuel itself does not contain nitrogen. The formation rate is mainly governed by the peak combustion temperature and the residence time at elevated temperatures (Chiesa, Lozza, & Mazzocchi, 2005, Cecere et al., 2023). The chemical framework of NO_x formation is discussed in the literature through several pathways, with the Zeldovich mechanism

representing the dominant thermal route. In pure hydrogen combustion, the prompt NO pathway is of secondary importance compared to hydrocarbon fuels. However, in blended fuel systems such as methane hydrogen mixtures, alternative kinetic pathways may regain relevance (Cecere et al., 2023). Accordingly, the fundamental principle of NO_x control in both engine and turbine applications can be summarized as reducing flame temperature and, where necessary, limiting the residence time at high temperatures (Chiesa et al., 2005, Cecere et al., 2023).

5.1. NO_x control in internal combustion engines

The primary approach for reducing NO_x emissions in hydrogen fueled internal combustion engines is the reduction of peak combustion temperature through lean operation. The wide flammability limits of hydrogen allow stable combustion to be maintained even at very low equivalence ratios. Under suitable operating conditions, this feature contributes to the suppression of thermal NO_x formation (Stepien, 2021, Teoh et al., 2023). Nevertheless, lean operation alone may not be sufficient under all conditions. Power demand, combustion stability, and efficiency requirements often necessitate the combined application of multiple NO_x control strategies (Wróbel et al., 2022).

Exhaust gas recirculation is frequently employed to reduce NO_x emissions in hydrogen engines. By introducing inert components into the mixture and increasing the overall heat capacity, EGR lowers flame temperature and limits thermal NO_x formation (Teoh et al., 2023, Stepien, 2021). However, depending on the operating point, EGR may reduce combustion speed and lead to trade offs such as reduced stability or performance losses (Wróbel et al., 2022). Similarly, water injection or intake air humidification can effectively reduce NO_x by lowering combustion temperature. These approaches, however, require careful assessment due to increased system complexity and operating costs (Teoh et al., 2023).

Fuel delivery and combustion phasing strategies also play a critical role in NO_x control. In particular, direct injection hydrogen systems enable more precise control of mixture formation and combustion timing, thereby providing additional flexibility in balancing emissions and performance (Wróbel et al., 2022). Consequently, NO_x control in hydrogen engines is typically addressed through combined strategies involving lean operation, optimized injection and phasing, and, where required, EGR or water injection (Stepien, 2021, Teoh et al., 2023).

5.2. NO_x control in gas turbines

In gas turbines, NO_x control in hydrogen combustion generally follows two main approaches. The first involves diffusion flames diluted with water, steam, or inert gases. The second relies on lean premixed combustion (Chiesa et al., 2005, Cecere et al., 2023). Dilution based approaches can effectively reduce NO_x by lowering flame temperature, but they may impose significant penalties in terms of efficiency and system complexity. Lean premixed combustors aim to control flame temperature and achieve low NO_x emissions. However, due to the high reactivity of hydrogen, these systems are susceptible to operational and safety risks such as flashback and autoignition (Cecere et al., 2023).

These risks directly influence the applicability of NO_x control strategies. As the hydrogen fraction increases, flame speed rises, mixing times decrease, and the geometric design window becomes narrower. As a result, gas turbine combustor design requires a more delicate balance between mixture quality and combustion stability (Cecere et al., 2023). To improve this balance, modern studies have explored micro mixing based concepts, which aim to reduce NO_x formation by minimizing residence time and distributing combustion across multiple small scale flame structures (Giacomazzi et al., 2023).

5.3. Integrated perspective and practical implications

When internal combustion engines and gas turbines are considered together, it becomes evident that NO_x control in hydrogen combustion cannot be achieved through a single universal solution. Instead, the selection of appropriate mitigation strategies depends on the application type, load range, and targeted performance criteria. Effective NO_x reduction requires the integrated design of primary measures within the combustion process and secondary measures applied downstream where necessary (Stepien, 2021, Cecere et al., 2023). Consequently, NO_x control remains a central research area in the development and deployment of hydrogen combustion technologies, and the advancement of novel combustion concepts and design strategies continues to be of critical importance (Giacomazzi et al., 2023). Table 4 summarizes the dominant NO_x formation mechanisms and mitigation strategies reported in the literature for hydrogen fueled internal combustion engines and gas turbines, emphasizing the trade offs between effectiveness, system complexity, and operational constraints.

Table 4. Overview of NO_x formation and mitigation strategies in hydrogen combustion systems

Application	Combustion regime	Primary NO _x formation driver	Mitigation strategy	Key limitations
SI hydrogen ICE (Verhelst & Wallner (2009); Stepien (2021))	Lean premixed	High adiabatic flame temperature	Lean operation (high λ)	Power density reduction, stability limits
SI hydrogen ICE (Teoh et al. (2023); Wróbel et al. (2022))	Stoichiometric / near-stoichiometric	Thermal NO _x	EGR	Volumetric efficiency loss, control sensitivity
SI hydrogen ICE (Teoh et al. (2023))	Stoichiometric	Thermal NO _x	Water injection	System complexity, water management
CI / hydrogen-assisted engines (Teoh et al. (2023); Wróbel et al. (2022))	Dual-fuel, premixed H ₂ + pilot	Local high-temperature zones	EGR + injection phasing	Limited hydrogen substitution ratio
Gas turbine (GT) (Chiesa et al. (2005); Cecere et al. (2023))	Lean premixed (DLE/DLN)	High flame temperature	Lean premixed combustion	Flashback, autoignition, instability
Gas turbine (GT) (Chiesa et al. (2005); Lefebvre & Ballal (2010))	Diffusion flame	Thermal NO _x	Steam / water dilution	Efficiency penalty, system complexity
Gas turbine (GT) (Giacomazzi et al. (2023); Cecere et al. (2023))	Micro-mixing / LDI	Thermal NO _x	Reduced residence time, rapid mixing	Manufacturing and cooling challenges
Generic H ₂ systems (Kozhukhova et al. (2021))	Catalytic combustion	Low-temperature oxidation	Catalytic hydrogen combustion	Catalyst durability, limited load range

6. Modeling and Simulation of Hydrogen Combustion

The safe, stable, and low emission implementation of hydrogen combustion in internal combustion engines and gas turbines depends not only on experimental investigations but also on the maturity of modeling and simulation tools. This is mainly due to the fact that hydrogen properties such as high flame speed, low ignition energy, and wide flammability range exert strong and often coupled influences on mixture formation, flame propagation, NO_x formation, and combustion instability mechanisms (Ströhle & Myhrvold, 2007; Knop et al., 2008). For this reason, hydrogen combustion has been investigated in the

literature using three main approaches: detailed chemical kinetic mechanisms, CFD based engine and gas turbine simulations incorporating combustion models, and CFD frameworks that address the coupled leakage, dispersion, and combustion chain in safety related scenarios (Venetsanos et al., 2008).

6.1. Chemical kinetic mechanisms and validation requirements

Under high pressure combustion conditions such as those encountered in gas turbines, detailed H_2 - O_2 kinetic mechanisms become critical for the accurate prediction of ignition delay, flame speed, and mixture reactivity. Ströhle and Myhrvold (2007) emphasized that hydrogen chemical kinetics gain particular importance within pressure ranges representative of gas turbine operation, typically between 15 and 35 atm. In their study, selected kinetic mechanisms were evaluated through comparisons with shock tube ignition delay measurements and laminar flame speed data. The results indicated that the Li and Ó Conaire mechanisms provided better agreement with experimental ignition delay data at elevated pressures up to 33 bar, while also demonstrating superior performance in predicting laminar flame speeds under high pressure conditions (Ströhle & Myhrvold, 2007). These findings clearly show that mechanism selection in gas turbine applications is not a purely formal step, since the chosen kinetic sub mechanism can directly influence predictions related to autoignition and flashback behavior.

In contrast, in three dimensional CFD simulations of internal combustion engines, the direct solution of detailed chemical kinetics is often computationally prohibitive. As a result, model based approaches are commonly adopted for combustion and emission prediction. In this context, fundamental kinetic validation studies such as ignition delay and laminar flame speed measurements generally serve as reliable inputs for CFD combustion models, ensuring that simplified combustion representations remain physically consistent with hydrogen specific chemical behavior.

6.2. Hydrogen in engine CFD: adaptation and validation of combustion models

In multidimensional simulations of hydrogen engines, a common approach is to employ flame surface or flamelet based models, as well as characteristic combustion models, to represent turbulent premixed combustion. However, due to the specific properties of hydrogen, the input parameters and certain model terms require dedicated adaptation to ensure physically consistent predictions.

In this context, Knop et al. (2008) adapted the Extended Coherent Flame Model within a three dimensional CFD framework to simulate hydrogen combustion. The study explicitly states that a new laminar flame speed correlation and a corresponding laminar flame thickness formulation were

introduced for hydrogen (Knop et al., 2008). In addition, the authors emphasized that the extended Zeldovich approach used for NO_x prediction cannot be directly applied under hydrogen conditions when using parameter sets developed for hydrocarbon fuels. For this reason, parameter adjustments were implemented to adapt the model to hydrogen combustion (Knop et al., 2008). Importantly, the model was not presented only as a theoretical formulation. Validation was performed by considering two limiting mixture formation cases. The first case corresponded to a cryogenic port fuel injection configuration representing approximately homogeneous mixture conditions. The second case involved late direct injection, resulting in a highly stratified mixture. In both cases, combustion behavior and NO_x prediction were evaluated (Knop et al., 2008). This approach demonstrates that a single calibration cannot be expected to accurately represent all operating conditions in hydrogen engine CFD, and that models must be tested under both homogeneous and stratified mixture regimes.

Rakopoulos et al. (2010) evaluated a CFD based combustion model for hydrogen spark ignition engine simulations and provided a detailed description of the model components. Reaction rates were calculated using a characteristic conversion time scale approach. Flame development was tracked through laminar and turbulent burning velocity correlations, while the turbulent flame speed was represented using the Zimont Lipatnikov formulation. NO formation was calculated using the Zeldovich mechanism (Rakopoulos et al., 2010). The study also discussed the wide range of available laminar flame speed correlations for hydrogen and pointed out that certain correlations should not be preferred due to their limited validity with respect to pressure and temperature ranges (Rakopoulos et al., 2010). Such analyses clearly indicate that hydrogen engine CFD requires the combined consideration of combustion model selection, correlation choice, and explicit acknowledgment of model limitations.

When these two studies are considered together, a consistent conclusion emerges from the literature. In hydrogen engine simulations, predictive accuracy depends less on the general structure of the combustion model and more on the correct definition and validation of hydrogen specific inputs, including flame speed, flame thickness, and NO_x formation related parameters, within the targeted operating regimes (Knop et al., 2008, Rakopoulos et al., 2010).

6.3. Numerical verification discipline

One of the critical factors that determines the reliability of numerical modeling of hydrogen combustion is numerical error control, including mesh sensitivity and time step sensitivity, together with the influence of geometric dimensionality.

Pashchenko (2017) compared two dimensional, two dimensional axisymmetric, and three dimensional simulations for premixed hydrogen air

combustion in a micro cylindrical combustor. The study reported that a two dimensional approach can be used for preliminary evaluation, but that results may deviate substantially from the three dimensional behavior. For example, the authors reported that differences of more than 25 percent can occur in outlet temperature when comparing three dimensional and two dimensional simulations. In the flame formation region, temperatures were reported to differ by several times (Pashchenko, 2017). Although these findings cannot be transferred directly to engine or gas turbine geometries, they provide methodological evidence that diffusion, mixing, and turbulence interaction effects in hydrogen combustion have an inherently three dimensional character and that critical physics may be missed in two dimensional reductions.

Similarly, Halouane and Dehbi (2017) simulated hydrogen air deflagration experiments using CFD and compared two model families, namely the Eddy Dissipation Model and the Turbulent Flame Closure approach. The study indicated that both models can capture the overall combustion evolution reasonably well. However, for fast dynamics, particularly pressure rise gradients, the Turbulent Flame Closure model was reported to perform better. At the same time, the Turbulent Flame Closure approach requires smaller time steps and finer meshes, which increases computational cost (Halouane & Dehbi, 2017). The authors also stated explicitly that mesh and time step sensitivity studies were conducted as part of CFD best practice procedures (Halouane & Dehbi, 2017). These findings demonstrate that in hydrogen combustion simulations, numerical configuration and verification discipline are as critical as model selection.

6.4. Safety oriented CFD and scenario based approaches

With the expanding implementation of hydrogen technologies, safety considerations have created a distinct simulation domain focusing on the CFD analysis of high pressure hydrogen releases, dispersion processes, and delayed ignition and combustion in vehicle, tunnel, and urban scenarios.

Venetsanos et al. (2008) investigated the potential consequences of releases from compressed hydrogen systems in urban and tunnel environments using CFD. The study considered different storage pressures of 20, 35, and 70 MPa and examined different pressure relief device opening scenarios. The methodology was explicitly structured into five stages, including scenario identification, source term calculations, dispersion calculations, combustion calculations, and analysis of results (Venetsanos et al., 2008). The findings emphasized that, in semi confined environments such as tunnels, a release of the same magnitude can generate a flammable cloud that persists for longer durations compared with open urban environments. The study also highlighted that the ignition location can influence the combustion regime and associated consequences (Venetsanos et al., 2008). Although such studies do not directly represent engine or turbine

combustor models, they provide a scientific framework defining the role of simulation in engineering design decisions related to pressure relief device strategies, ventilation, and risk mitigation.

6.5. Synthesis and good practice principles

The technical outcomes emerging from this body of literature can be summarized through three main principles.

First, objective model matching is essential. For kinetically sensitive problems such as autoignition and flashback, mechanism selection is critical, and mechanisms validated against ignition delay and laminar flame speed data should be preferred (Ströhle & Myhrvold, 2007).

Second, hydrogen specific adaptation is required. Combustion and NO_x models developed for hydrocarbon fuels should not be directly applied to hydrogen without modification. Laminar flame speed, flame thickness, and Zeldovich related parameters must be treated under hydrogen relevant conditions (Knop et al., 2008, Rakopoulos et al., 2010).

Third, numerical verification is mandatory. If the effects of two dimensional versus three dimensional modeling, mesh resolution, time step sensitivity, and model choice are not evaluated together, simulation results can be misleading (Pashchenko, 2017, Halouane & Dehbi, 2017).

Within this framework, the modeling and simulation literature on hydrogen combustion is not limited to describing combustion phenomena. It represents a set of methodological tools that directly guide engineering design decisions, including injection strategy, combustor geometry, NO_x mitigation packages, and safety scenario assessment.

7. Future Outlook and Research Needs

The literature on the application of hydrogen combustion in internal combustion engines and gas turbines has reached a considerable level of maturity in recent years. Nevertheless, it also indicates that several fundamental research challenges remain to be addressed before these systems can be deployed in a widespread and safe manner. Existing studies demonstrate that the combustion characteristics of hydrogen offer strong potential in terms of high efficiency and low carbon emissions. However, they also show that this potential can only be realized in practical applications through appropriate combustion control, effective emission management, and safety oriented design approaches (Stepien, 2021, Wróbel et al., 2022).

From the perspective of internal combustion engines, one of the key future research priorities is the achievement of stable operation with low NO_x emissions

under high load conditions. Although approaches such as lean operation, direct injection, and exhaust gas recirculation have enabled significant progress, the trade off between combustion stability and emission control remains a critical issue, particularly in applications requiring high power density (Verhelst & Wallner, 2009, Teoh et al., 2023). In this context, continued efforts to jointly optimize injection strategies, ignition systems, and combustion phasing remain essential.

On the gas turbine side, the future of hydrogen utilization is closely linked to combustor design and mixture control. Lean premixed and micro mixing based combustor concepts offer promising pathways for reducing NO_x emissions. However, challenges such as flashback, autoignition, and thermoacoustic instabilities significantly constrain the available design window for these systems (Chiesa et al., 2005, Cecere et al., 2023). Consequently, future research is expected to focus on improving the balance between combustion stability and emission performance in gas turbines operating with high hydrogen content fuels.

In the field of modeling and simulation, the literature clearly establishes that reliable representation of hydrogen combustion requires the combined use of validated chemical kinetic mechanisms and combustion models specifically adapted for hydrogen (Ströhle & Myhrvold, 2007, Knop et al., 2008). In addition, mesh and time step sensitivity, two dimensional versus three dimensional geometric effects, and the need for experimental validation are highlighted as critical methodological aspects in numerical studies of hydrogen combustion (Rakopoulos et al., 2010, Halouane & Dehbi, 2017).

Finally, as hydrogen technologies continue to expand, the importance of safety oriented modeling and scenario analysis is increasing. CFD based studies that address high pressure hydrogen storage, leakage, dispersion, and delayed ignition in an integrated manner are increasingly regarded as complementary tools in the design of hydrogen based systems (Venetsanos et al., 2008).

Overall, the literature indicates that hydrogen combustion is technically feasible in both internal combustion engines and gas turbines. At the same time, it emphasizes that achieving performance, emission, and safety objectives simultaneously requires interdisciplinary and multiscale research efforts. In this regard, the future development of hydrogen combustion technologies depends on the coordinated advancement of experimentally validated modeling approaches together with advanced combustion concepts and emission control strategies.

8. Conclusions

In this book chapter, hydrogen combustion technologies have been examined from a holistic perspective, with a focus on internal combustion engines and gas

turbines. Based on a synthesis and comparative evaluation of the existing literature, the following key conclusions can be drawn.

1. Hydrogen is a fuel with the potential to eliminate CO₂, CO, and particulate matter emissions at the source in internal combustion engines and gas turbines due to its carbon free nature. However, NO_x formation associated with high combustion temperatures remains the primary factor limiting the environmental performance of hydrogen based systems.
2. The wide flammability limits and high flame speed of hydrogen provide important advantages for lean combustion and high efficiency operation. At the same time, these properties necessitate careful management of specific challenges such as pre ignition, flashback, and combustion stability.
3. In internal combustion engines, the use of hydrogen is strongly dependent on the fuel delivery strategy. Direct injection systems offer a wider operating window compared to premixed configurations by reducing volumetric efficiency losses and limiting flashback risk.
4. Hydrogen combustion in gas turbines requires design approaches that differ substantially from conventional gas combustors, particularly due to challenges related to flashback, autoignition, and thermoacoustic instabilities. Lean premixed, diluted, and micro mixing based combustion concepts represent the most prominent solutions in this context.
5. NO_x control in both engine and gas turbine applications is primarily based on reducing combustion temperature. Lean operation, exhaust gas recirculation, water or steam dilution, and appropriate injection and phasing strategies are identified in the literature as the most widely applied in cylinder control methods.
6. Modeling of hydrogen combustion requires the combined use of detailed chemical kinetic mechanisms and combustion and NO_x models adapted specifically for hydrogen. Accurate representation of parameters such as laminar flame speed, flame thickness, and ignition delay is critical for the reliability of numerical predictions.
7. CFD based studies have become indispensable tools for understanding hydrogen combustion behavior in engines, gas turbines, and safety related scenarios. However, the literature clearly shows that results obtained without proper consideration of mesh

resolution, time step sensitivity, and two dimensional versus three dimensional modeling effects may have limited validity.

8. With the increasing deployment of hydrogen technologies, safety oriented modeling studies addressing leakage, dispersion, and delayed ignition scenarios have become an integral part of engineering design practice.
9. In general, the literature demonstrates that hydrogen combustion is technically feasible in both internal combustion engines and gas turbines. At the same time, it highlights the need for interdisciplinary and multiscale research to achieve performance, emission, and safety targets simultaneously.
10. In this context, the future of hydrogen combustion technologies depends on the coordinated advancement of experimentally validated modeling approaches together with advanced combustion concepts and effective emission control strategies

References

- Akal, D., Oztuna, S., & Büyükkakın, M. K. (2020). A review of hydrogen usage in internal combustion engines (gasoline-LPG-diesel) from combustion performance aspect. *International Journal of Hydrogen Energy*, 45(5), 35257-35268. <https://doi.org/10.1016/j.ijhydene.2020.02.001>
- Cecere, D., Giacomazzi, E., Di Nardo, A., & Calchetti, G. (2023). Gas turbine combustion technologies for hydrogen blends. *Energies*, 16(19), 6829. <https://doi.org/10.3390/en16196829>
- Chiesa, P., Lozza, G., & Mazzocchi, L. (2005). Using hydrogen as gas turbine fuel. *Journal of Engineering for Gas Turbines and Power*, 127(1), 73-80. <https://doi.org/10.1115/1.1787513>
- Correa, S. M. (1993). A review of NO_x formation under gas-turbine combustion conditions. *Combustion Science and Technology*, 87(1-6), 329-362. <https://doi.org/10.1080/00102209208947221>
- Firat, M., Altun, Ş., Okçu, M., Varol, Y., & Şenocak, E. (2022). Experimental study on reduction of pollutant emissions in a reactivity-controlled compression ignition (RCCI) engine. *International Journal of Automotive Engineering and Technologies*, 11(2), 63-72. <https://doi.org/10.18245/ijaet.1078400>
- Firat, M., Varol, Y., Öztürk, E., & Şenocak, E. (2023). Ozone-assisted combustion and emission control in reactivity controlled compression ignition engine. *Energy*, 284, 129247. <https://doi.org/10.1016/j.energy.2023.129247>
- Giacomazzi, E., Troiani, G., Di Nardo, A., Calchetti, G., Cecere, D., Messina, M., & Carpenella, M. (2023). Hydrogen combustion: Features and barriers to its exploitation in the energy transition. *Energies*, 16(20), 7174. <https://doi.org/10.3390/en16207174>
- Halouane, K., & Dehbi, A. (2017). CFD simulations of premixed hydrogen combustion in a constant volume explosion: Influence of turbulent combustion modelling and numerical differencing schemes. *International Journal of Hydrogen Energy*, 42(34), 21990-22004. <https://doi.org/10.1016/j.ijhydene.2017.07.075>
- Knop, V., Benkenida, A., Jay, S., & Colin, O. (2008). Modelling of combustion and nitrogen oxide formation in hydrogen-fuelled internal combustion engines within a 3D CFD code. *International Journal of Hydrogen Energy*, 33(19), 5083-5097. <https://doi.org/10.1016/j.ijhydene.2008.06.027>
- Kozhukhova, A. E., du Preez, S. P., & Bessarabov, D. G. (2021). Catalytic hydrogen combustion for domestic and safety applications: A critical review of catalyst materials and technologies. *Energies*, 14, 4897. <https://doi.org/10.3390/en14164897>

- Lefebvre, A. H., & Ballal, D. R. (2010). *Gas turbine combustion: Alternative fuels and emissions* (3rd ed.). Boca Raton, FL: CRC Press.
- Lieuwen, T., & Yang, V. (Eds.). (2005). *Combustion instabilities in gas turbine engines: Operational experience, fundamental mechanisms, and modeling*. Reston, VA: AIAA.
- Öztürk, G., Şenocak, Ş. M., Şenocak, N., & Fırat, M. (2024). *An experimental investigation of 1,2-dimethoxy ethane as a fuel additive in biodiesel-fueled diesel engine*. *Journal of the Energy Institute*, 117, 101824. <https://doi.org/10.1016/j.joei.2024.101824>
- Pashchenko, D. (2017). Comparative analysis of hydrogen/air combustion CFD-modeling in micro-cylindrical combustor. *International Journal of Hydrogen Energy*, 42(47), 29545-29556. <https://doi.org/10.1016/j.ijhydene.2017.10.070>
- Rakopoulos, C. D., Kosmadakis, G. M., & Pariotis, E. G. (2010). Evaluation of a combustion model for the simulation of hydrogen spark-ignition engines using a CFD code. *International Journal of Hydrogen Energy*, 35(22), 12545-12560. <https://doi.org/10.1016/j.ijhydene.2010.09.002>
- Stepien, Z. (2021). A comprehensive overview of hydrogen-fueled internal combustion engines: Achievements and future challenges. *Energies*, 14, 6504. <https://doi.org/10.3390/en14206504>
- Ströhle, J., & Myhrvold, T. (2007). Detailed reaction mechanisms for hydrogen combustion under gas turbine conditions. *International Journal of Hydrogen Energy*, 32(2), 125-135. <https://doi.org/10.1016/j.ijhydene.2006.04.005>
- Teoh, Y. H., Zainal, Z. A., Alabdulkarem, A., Yu, Y., Norrizal Mustaffa, M. A., Awad, O. I., & Mohamed, M. N. M. (2023). Advanced hydrogen combustion strategies and emissions control: A review of internal combustion engines and gas turbines. *Fuel*, 333, 126525. <https://doi.org/10.1016/j.fuel.2022.126525>
- Venetsanos, A. G., Baraldi, D., Adams, P., Heggem, P. S., & Wilkening, H. (2008). CFD modelling of hydrogen release, dispersion and combustion for automotive scenarios. *Journal of Loss Prevention in the Process Industries*, 21(2), 162-184. <https://doi.org/10.1016/j.jlp.2007.06.016>
- Verhelst, S., & Wallner, T. (2009). Hydrogen-fueled internal combustion engines. *Progress in Energy and Combustion Science*, 35(6), 490-527. <https://doi.org/10.1016/j.pecs.2009.08.001>
- White, C. M., Steeper, R. R., & Lutz, A. E. (2006). The hydrogen-fueled internal combustion engine: A technical review. *International Journal of Hydrogen Energy*, 31(10), 1292-1305. <https://doi.org/10.1016/j.ijhydene.2005.12.001>

- Wróbel, K., Wróbel, J., Tokarz, W., Lach, J., Podsadni, K., & Czerwiński, A. (2022). Hydrogen internal combustion engine vehicles: A review. *Energies*, 15, 8937. <https://doi.org/10.3390/en15238937>
- Yip, H. L., Srna, A., Yuen, A. C. Y., Kook, S., Taylor, R. A., Yeoh, G. H., Medwell, P. R., & Chan, Q. N. (2019). A review of hydrogen direct injection for internal combustion engines: Towards carbon-free combustion. *Applied Sciences*, 9(22), 4842. <https://doi.org/10.3390/app9224842>

CHAPTER 23

Photovoltaic Solar Panel Technologies: A Comparison of Efficiency, Cost, and Lifetime Based on Material Types

Şafak Melih Şenocak¹ & Yasin Varol²

¹ Lecturer, Osmaniye Korkut Ata University, ORCID: 0000-0003-0602-2836

² Prof. Dr., Fırat University, ORCID: 0000-0003-2989-7125

1. Introduction

With the growth of the global population and the increasing demand for energy, the use of fossil based resources and the associated environmental pollution have reached critical levels. Although various efforts have been undertaken to mitigate the negative impacts of fossil fuel based energy systems, the demand for renewable energy sources continues to rise steadily (Firat et al., 2023; Öztürk et al., 2024). Solar energy is the most abundant, secure, and environmentally clean energy source for sustainable economic growth (Gürgeç et al., 2025a). It can be stored in the form of thermal energy or converted into different forms of energy for various applications (Şenocak et al., 2025; Gürgeç et al., 2025b). Photovoltaic (PV) technology represents one of the most efficient methods for utilizing solar energy by directly converting solar radiation into electricity. In recent years, substantial progress has been achieved in the PV sector, and a rapid increase in installed capacity has been observed since the 2010s (Sampaio & González, 2017). By the end of 2020, the global installed PV capacity had reached several hundred gigawatts and is projected to increase to the terawatt scale by 2050 (Chowdhury et al., 2020). The PV market is currently dominated by crystalline silicon (c-Si) based technologies. However, alternative PV cells based on different material systems have also been developed, attracting significant attention due to their potential for cost reduction and efficiency enhancement (Sampaio & González, 2017). In this context, performance comparisons of widely used PV panel materials such as monocrystalline and polycrystalline silicon, cadmium telluride (CdTe), copper indium gallium selenide (CIGS), and the recently emerging perovskite solar cells are of considerable importance. In this chapter, these technologies are examined in terms of efficiency, cost, lifetime, and environmental impact, while methods aimed at improving the efficiency of PV systems are also discussed. By presenting the current state of the art, the necessity of such a comprehensive comparison is clearly articulated. The findings aim to provide readers at the research level with an in-depth perspective on the current status and future outlook of PV technologies.

2. Classification of PV Panel Technologies

Based on material composition, photovoltaic cell technologies can generally be classified into three generations or categories (Kant & Singh, 2022). Figure 1 schematically illustrates solar panel types according to their material structures.

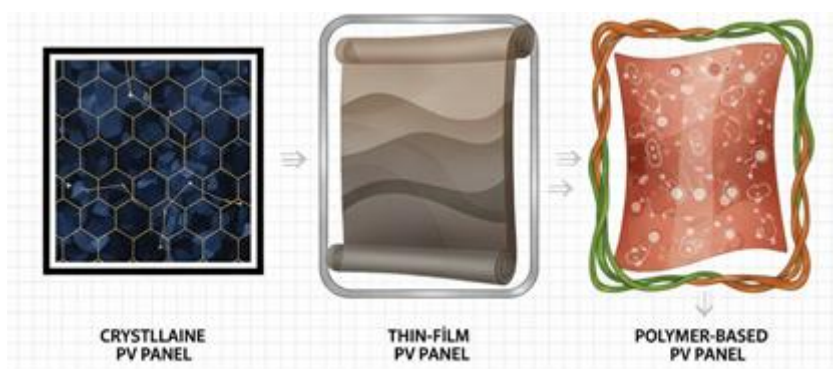


Figure 1. Solar panel types based on their structural characteristics (Created by authors).

First generation photovoltaic technologies, commonly referred to as crystalline based PV systems, represent the most widely commercialized products today and are primarily based on monocrystalline and polycrystalline silicon wafers. Crystalline silicon cells dominate more than 90 percent of the global market due to their high efficiencies, typically exceeding 20 percent, and their proven long term stability (Al-Ezzi & Ansari, 2022). Monocrystalline silicon cells exhibit higher conversion efficiencies as a result of their uniform and well ordered crystal structure, whereas polycrystalline silicon cells show slightly lower efficiencies due to grain boundaries and structural imperfections, although their production is more cost effective (Al-Ezzi & Ansari, 2022).

Second generation photovoltaic technologies, known as thin film PV systems, aim to reduce material usage and energy consumption by employing very thin semiconductor layers. Amorphous silicon, cadmium telluride, and copper indium gallium selenide belong to this category. Thin film technologies typically offer lower cell efficiencies, in the range of approximately 10 to 15 percent, but provide advantages such as reduced manufacturing costs and flexibility in applications (Kant & Singh, 2022). For instance, CdTe and CIGS cells absorb incident light using absorber layers with thicknesses of only 1 to 2 micrometers, significantly reducing material consumption and potentially lowering production costs. In laboratory conditions, CdTe and CIGS cells have achieved record efficiencies of approximately 21 percent and 23 percent, respectively, approaching the performance of crystalline silicon technologies (Fraunhofer ISE, 2025). However, sustainability challenges remain due to the toxic nature of cadmium and the scarcity of tellurium in CdTe cells, as well as the limited availability of critical metals such as indium and gallium in CIGS cells (Al-Ezzi & Ansari, 2022). Another thin film material is amorphous silicon, which is inexpensive to manufacture and suitable for flexible applications but exhibits low standalone cell

efficiencies of around 5 to 8 percent and suffers from light induced degradation known as the Staebler-Wronski effect.

Third generation photovoltaic technologies encompass emerging approaches based on innovative materials such as organic polymer cells, dye sensitized solar cells, quantum dot cells, and particularly perovskite solar cells (Nayak et al., 2019). Some of these technologies are also described as hybrid systems. For example, perovskite materials possess an organic–inorganic hybrid crystal structure. The primary objective of third generation cells is either to surpass the Shockley-Queisser efficiency limit through new physical mechanisms or to achieve significantly lower production costs (Kant & Singh, 2022). Although these technologies are still largely in the research phase, they have demonstrated rapid efficiency improvements at the laboratory scale. For instance, the single junction efficiency of lead halide perovskite solar cells increased from approximately 3 percent in 2009 to more than 25 percent by 2022 (Sharma et al., 2022). Nevertheless, long term stability and large scale manufacturability remain unresolved challenges.

Finally, multi junction or tandem solar cells constitute a category designed to utilize the solar spectrum more efficiently by stacking materials with different band gaps. In particular, triple and quadruple junction gallium arsenide based cells have achieved record efficiencies exceeding 45 percent under concentrated illumination (Fraunhofer ISE, 2025). However, due to their extremely high production costs, these III–V semiconductor cells are mainly used in niche applications such as space technologies (Crowell, 2021). In recent years, perovskite–silicon tandem cells have emerged as a promising and potentially cost effective pathway to bring multi junction concepts into commercial terrestrial applications. This hybrid approach is widely regarded as a precursor to fourth generation photovoltaic technologies (Kant & Singh, 2022).

2.1. Efficiency Comparison

The efficiency with which photovoltaic cells and modules convert solar radiation into electrical energy varies significantly depending on the material system employed. Average efficiencies of commercial modules and laboratory record values are summarized in Table 1. In general, monocrystalline silicon panels currently achieve the highest commercial module efficiencies. With the adoption of next generation n type cell architectures, monocrystalline modules with efficiencies approaching 24 percent have been reported (Svarc, 2025). Polycrystalline silicon panels, by contrast, have reached maximum efficiencies of approximately 18 percent and are considered technologically mature, leading to their gradual replacement by monocrystalline products in most markets (Svarc, 2025).

Among thin film technologies, CdTe and CIGS modules have shown substantial improvements over time. While their efficiencies initially remained within the 10 to 15 percent range, recent advances have increased commercial CdTe module efficiencies to nearly 19 percent (Fraunhofer ISE, 2025), with CIGS modules achieving similar values in the range of 15 to 17 percent. At the laboratory scale, both thin film technologies have exceeded 20 percent efficiency, demonstrating performance levels comparable to crystalline silicon systems (Nayak et al., 2019).

Perovskite solar cells, although still at the research stage, have rapidly emerged as leading contenders in the efficiency race. Single junction perovskite cell efficiencies reached approximately 25 percent during the 2020s, and when integrated into silicon based tandem architectures, efficiencies of around 33 to 35 percent were reported by 2023, surpassing the Shockley–Queisser limit for single junction devices. However, these record efficiencies have thus far been achieved only in small area laboratory cells and have not yet been fully translated to stable module scale devices. The efficiency values presented in Table 1 are derived from up to date literature published after 2015.

Table 1. Efficiency comparison of different PV panel materials (Fraunhofer ISE, 2025; Svarc, 2025; Nayak et al., 2019)

PV Technology	Efficiency (approx., %)
Monocrystalline Si	20-24 (commercial module)
Polycrystalline Si	15-18 (commercial module)
CdTe (thin film)	16-19 (commercial module)
CIGS (thin film)	15-17 (commercial module)
Perovskite (laboratory cell)	20-25 (single junction, laboratory scale)

According to the table above, monocrystalline silicon cells provide the highest efficiency among the evaluated technologies. During the 2020s, the efficiency of monocrystalline silicon modules exceeded the 24 percent range as a result of advanced cell architectures such as PERC, TOPCon, and heterojunction (HJT) designs (Svarc, 2025). Polycrystalline silicon modules, on the other hand, remain slightly behind due to material purity limitations and crystal defects, with efficiencies typically in the range of 15 to 18 percent. CdTe and CIGS thin film technologies initially lagged behind crystalline silicon but have gradually improved through continuous technological development, reaching efficiency levels of approximately 16 to 19 percent. In particular, CdTe modules manufactured by First Solar have approached efficiencies close to 19 percent, making this technology economically competitive as well (Fraunhofer ISE, 2025). Commercial CIGS modules have similarly reported efficiencies of around 17 percent.

The main advantages of thin film technologies lie in their lower production costs and flexibility. However, compared to crystalline silicon, they are not yet considered equally mature technologies, and their full efficiency improvement potential has not been completely realized (Nayak et al., 2019). Regarding perovskite solar cells, Table 1 reports laboratory scale single junction efficiencies in the range of 20 to 25 percent. Achieving such high efficiency levels within only a few years is particularly noteworthy (Sharma et al., 2022). Indeed, perovskite solar cells improved from approximately 3 percent efficiency in 2009 to around 25 percent by 2019 (Sharma et al., 2022). Moreover, when combined with silicon in tandem configurations, efficiencies exceeding 30 percent have been demonstrated (Lira-Cantu, 2025). Nevertheless, the primary limitation of perovskite technology remains long term stability, which is discussed in detail in the lifetime section. Overall, although monocrystalline silicon remains the leading technology in terms of efficiency, the performance gap with alternative technologies has narrowed. In particular, tandem approaches and novel materials suggest that commercial efficiencies above 30 percent may become feasible in the future (Lira-Cantu, 2025).

2.2. Cost Comparison

In the economic evaluation of photovoltaic technologies, the cost per watt is commonly used as a key metric. The manufacturing costs and market prices of cells and modules vary depending on the material system and production technology. Table 2 presents approximate module costs for major PV panel types based on data reported for the year 2020 (Crowell, 2021). The values shown correspond to the minimum sustainable price levels required for economically viable industrial production. According to a comprehensive analysis conducted by the National Renewable Energy Laboratory, crystalline silicon technologies exhibit the lowest cost per watt, followed by CdTe modules (Crowell, 2021). By contrast, CIGS panels continue to have significantly higher production costs. Although perovskite cells have not yet reached full commercial maturity and are currently produced only at the pilot scale, they are expected to be manufactured at very low cost in the future (Crowell, 2021). These cost estimates are summarized in Table 2.

Table 2. Approximate module costs of PV panel technologies (Crowell, 2021)

PV Technology	Module Cost (USD/W)
Monocrystalline Si	0.26 (average)
Polycrystalline Si	0.24 (average)
CdTe thin film	0.28
CIGS thin film	0.48
Perovskite single junction	0.38 (pilot production)

According to Table 2, the manufacturing cost of crystalline silicon modules is approximately 0.25 to 0.27 USD per watt (Crowell, 2021). The cost difference between monocrystalline and polycrystalline modules is minimal, and the higher efficiency of monocrystalline silicon allows it to remain competitive on a cost per watt basis. CdTe thin film modules have reached a sustainable production cost of around 0.28 USD per watt. Manufacturers such as First Solar have successfully reduced CdTe module prices to levels comparable with silicon modules through economies of scale. In contrast, the more complex and costly production processes associated with CIGS technology result in higher module costs of approximately 0.48 USD per watt, which limits its market share (Crowell, 2021).

For perovskite technology, which has not yet achieved full commercial maturity, cost modeling suggests that single junction perovskite modules could be produced at approximately 0.38 USD per watt under small scale pilot production conditions (Crowell, 2021). Researchers further anticipate that if efficiency and stability challenges in perovskite cells are resolved, long term production costs could potentially fall below 0.20 USD per watt (Crowell, 2021). In addition, early stage manufacturing estimates indicate that perovskite on silicon tandem modules could reach cost levels of approximately 0.31 USD per watt. Overall, PV module costs have declined dramatically over the past decade. While crystalline silicon modules were priced at around 1.5 USD per watt in 2010, prices had fallen below 0.3 USD per watt by the early 2020s. This rapid cost reduction has enabled solar electricity to become competitive with fossil fuel based power generation in many regions. Thin film technologies have followed a similar downward cost trend, although crystalline silicon continues to maintain cost leadership due to its strong scale advantages. Emerging technologies, while not yet commercially deployed, may enable further cost reductions in the future through material savings and simplified manufacturing processes.

2.3. Lifetime and Stability Comparison

The economic lifetime and performance stability of solar panels are critical factors for the long term viability of photovoltaic technologies. For current commercial PV modules, a typical service lifetime of 25 to 30 years is widely accepted (Chowdhury et al., 2020). Over this period, module power output gradually decreases due to degradation, with most manufacturers guaranteeing that 80 to 85 percent of the initial rated power will be retained after 25 years of operation. Expected lifetimes of different panel types are summarized in Table 3.

Crystalline silicon modules are well known for their long term reliability. Field data indicate that high quality c-Si panels can operate for 25 to 30 years with average annual power losses of less than 0.5 percent. Thin film technologies are also generally designed for a 25 year lifetime, although slightly higher degradation rates may be observed during the initial years of operation. For example, CdTe modules may experience minor early stage effects related to cadmium telluride corrosion, while CIGS modules can suffer from absorber layer

degradation under moisture ingress. These issues have largely been mitigated through improved encapsulation techniques.

Perovskite solar cells currently represent the weakest link in terms of operational lifetime. According to existing literature, the most stable perovskite cells reported to date achieve continuous operation lifetimes on the order of approximately one year (Sharma et al., 2022). In contrast, crystalline silicon panels routinely operate reliably for 25 years or longer. Sharma et al. (2022) reported that while the maximum demonstrated lifetime of perovskite solar cells is approximately one year, silicon based cells can achieve lifetimes of up to 25 years. This substantial disparity remains the primary barrier to the commercialization of perovskite technologies. Nevertheless, researchers anticipate that perovskite modules may approach 25 year lifetime targets within the next few years as standard qualification tests are successfully met (Lira-Cantu, 2025).

Table 3. Comparison of average lifetime and stability of different PV technologies (Nayak et al., 2019)

PV Technology	Estimated Lifetime (years)	Notes on degradation and stability
Monocrystalline Si	25-30	Annual power loss of approximately 0.3 to 0.5 percent. More than 80 percent of initial power retained after 25 years. High stability
Polycrystalline Si	Approximately 25	Similar to monocrystalline silicon. Slightly higher first year loss may occur due to impurities
CdTe thin film	Approximately 25	Initial performance increase of about 1 to 2 percent during the first years due to light induced effects. Followed by around 0.5 percent annual loss. A 25 year lifetime is achievable with proper encapsulation
CIGS thin film	Approximately 25	Additional loss of about 1 to 2 percent in the first year followed by around 0.5 percent annual degradation. Sensitive to moisture and oxygen. Encapsulation is critical
Perovskite	1-3 (laboratory prototypes)	Highly sensitive to UV radiation, moisture, and heat. Unencapsulated cells degrade within hours or days. Target lifetime exceeds 20 years but has not yet been achieved

In light of Table 3, crystalline silicon modules, both monocrystalline and polycrystalline, clearly demonstrate superior performance in terms of lifetime and durability. Most manufacturers provide performance warranties of 25 to 30 years for silicon based modules. Field data further indicate the existence of silicon panels that have been operating for more than 30 years while still retaining over 70 percent of their initial power output. Thin film technologies are also designed

with similar lifetime targets. For example, CdTe module manufacturers typically offer 25 to 30 year performance guarantees. During the early years of operation, CdTe cells may exhibit a slight performance improvement due to changes in the absorber layer, such as dopant redistribution, after which they display stable performance with degradation rates comparable to crystalline silicon.

For CIGS modules, a 25 year operational lifetime is achievable when appropriate encapsulation and barrier layers are employed. However, field experience has shown that occasional failures can occur due to moisture ingress and temperature related effects. In some thin film systems, power losses of up to approximately 20 percent after around 20 years of operation have been reported. These issues are being addressed through continuous improvements in encapsulation technologies (Nayak et al., 2019).

In contrast, stability remains a major challenge for next generation perovskite and organic solar cells. Perovskite cells are highly sensitive to environmental factors such as moisture, oxygen, ultraviolet radiation, and elevated temperatures. Without protective encapsulation, rapid efficiency loss can occur within hours or days (Dada & Popoola, 2023). Recent studies have demonstrated that advanced encapsulation materials, interface engineering, and device architecture modifications can enable laboratory scale stability exceeding 1000 hours, corresponding to approximately one and a half months of continuous operation. Moreover, some companies have announced that perovskite silicon tandem modules have successfully passed 25 year lifetime qualification tests (Lira-Cantu, 2025). Nevertheless, these claims require independent long term validation under real field conditions. Overall, based on proven field reliability, crystalline silicon technology remains the most dependable option at present. Thin film panels can offer comparable durability when appropriate design and protection measures are implemented. For emerging technologies such as perovskite photovoltaics, substantial progress in long term stability is still required in the coming years.

3. Efficiency Enhancement Methods

One of the main factors limiting PV panel efficiency is high operating temperature. Solar panels convert only a portion of incident radiation into electricity, while the remainder is converted into heat. As the cell temperature rises above the 25 °C reference condition, cell efficiency decreases by approximately 0.4 to 0.5 percent per degree Celsius (Kozak-Jagiela et al., 2023). Therefore, cooling techniques play an important role in improving PV system performance. Passive cooling approaches include adding heat dissipating fins to the rear side of the module, enhancing airflow through natural convection, or applying radiative cooling coatings on the module surface. For example, aluminum heat spreaders attached to the back of a panel can reduce cell temperature by several degrees Celsius without additional energy consumption

and can increase efficiency by approximately 2 to 5 percent. Active cooling methods, in contrast, rely on external energy input to remove heat, such as fans or circulating fluids. The most common active approach is water based cooling, either by circulating water behind the panel or by spraying water onto the module surface. Due to the high heat transfer capability of water, this method can reduce panel temperature substantially. Review studies report that active water cooling can lower panel temperature by around 20 percent and increase output power by 8 to 12 percent (Kozak-Jagięła et al., 2023). Air based cooling is generally less effective. Fan assisted systems can reduce temperature, but their contribution to efficiency is typically limited compared with water cooling (Kozak-Jagięła et al., 2023). In recent years, phase change materials have also gained attention in PV cooling applications (Kozak-Jagięła et al., 2023). When integrated at the rear of the panel, phase change materials absorb and store heat during sunny periods by melting, which helps prevent excessive temperature rise. As they solidify during the night, the stored heat is released to the surroundings. This cycle provides passive thermal buffering and can maintain panel temperature within a narrower range. Overall, cooling techniques represent important auxiliary solutions for improving PV efficiency, particularly in hot climates. Passive approaches are simple and low maintenance, whereas active approaches can be more effective but involve higher costs, and practical selection is usually based on economic optimization.

Another approach to improving PV performance involves surface coating technologies applied to the panel. Anti reflective and anti soiling coatings on the front glass aim to increase light transmission to the cells and to maintain surface cleanliness. A standard glass surface reflects approximately 4 percent of incident sunlight (Elsafi et al., 2025). In desert and semi arid climates, dust accumulation on panels can lead to more than 10 percent energy loss within a few months if cleaning is not performed (Elsafi et al., 2025). Hydrophobic coatings and photocatalytic self cleaning coatings reduce dust adhesion and facilitate removal by rainwater. For instance, nano structured anti soiling coatings have been reported to reduce dust accumulation by 30 to 40 percent compared with uncoated surfaces. Combining anti reflective and anti soiling functionality is an important target, and the literature reports that dual function coatings can maximize PV output while minimizing cleaning requirements (Elsafi et al., 2025). Long term durability of coatings is also crucial because UV exposure, sand abrasion, and thermal cycling can degrade coating performance in outdoor conditions. For this reason, next generation coatings are developed to withstand harsh environments and are evaluated using standardized qualification procedures such as IEC 61215 (Elsafi et al., 2025).

In addition to these methods, solar tracking systems are widely used to enhance energy yield in PV installations. Conventional fixed tilt panels cannot maintain normal incidence to sunlight throughout the day and across seasons due to the apparent motion of the sun. Tracking systems optimize the panel orientation toward the sun to maximize irradiance capture. Single axis trackers typically rotate panels along an east to west trajectory to follow daily solar movement. Dual axis trackers can adjust both azimuth and elevation angles, enabling panels to face the sun more directly throughout the year. These gains make tracking systems particularly attractive for large solar farms in regions where land costs are relatively low. Today, a considerable portion of utility scale PV plants in the United States and Europe employ single axis tracking. The main disadvantages of tracking systems are higher investment and maintenance costs, along with the risk of mechanical failures associated with moving components. However, reliability has improved with technological progress, and over the long term the additional energy yield can offset these costs. Overall, supporting engineering measures such as cooling strategies, advanced coatings, and solar tracking can significantly enhance the performance of existing PV technologies. When combined appropriately, these methods can help maintain module performance closer to laboratory values under real operating conditions and increase the total energy yield of photovoltaic systems.

4. Recycling and Environmental Impacts

Solar energy systems are widely regarded as clean technologies because they do not produce carbon emissions during operation. The life cycle CO₂ emissions of silicon based PV panels per unit of electricity generated are relatively low, typically below 50 g CO₂ per kWh, which is far lower than the 400 to 1000 g CO₂ per kWh associated with coal or natural gas power plants (Chowdhury et al., 2020). Therefore, PV is clearly a climate friendly technology. Nevertheless, end of life waste management and the toxic characteristics of certain materials represent environmental issues that require careful attention. As the PV market continues to expand, the volume of decommissioned panels will also increase. Assuming an average lifetime of 25 years, global PV waste is estimated to reach 1.7 to 8 million tonnes by 2030 and to rise to 60 to 78 million tonnes by 2050 (Chowdhury et al., 2020). This scale makes the development of effective recycling infrastructure essential for the PV sector. Figure 2 schematically illustrates PV panel recycling.

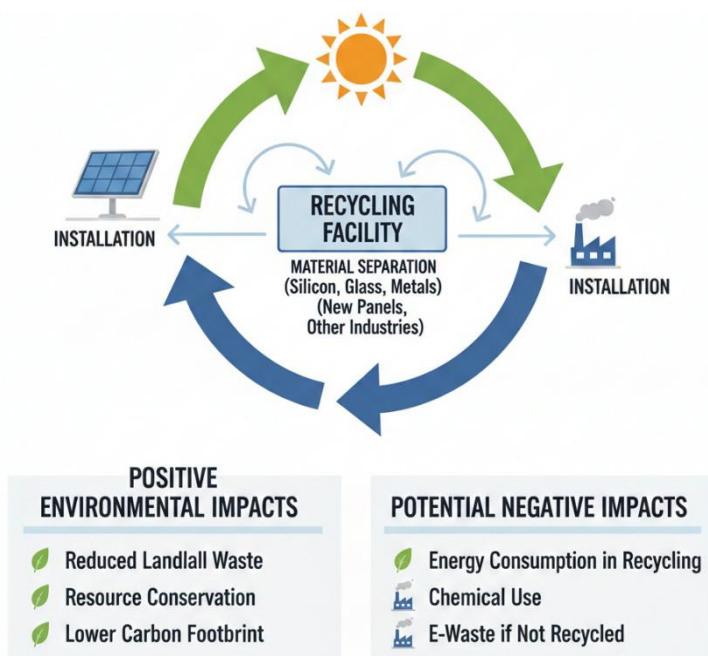


Figure 2. Recycling of PV panels (Created by authors).

The European Union has included photovoltaic panels within the scope of the Waste Electrical and Electronic Equipment Directive. Under this framework, manufacturers placing PV modules on the market are responsible for financing the collection and recycling of panels at the end of their service life (Chowdhury et al., 2020). Accordingly, specific collection and recovery targets have been defined for PV modules. Crystalline silicon panels consist largely of recyclable materials, with approximately 70 to 80 percent glass and about 10 percent aluminum framing by weight. Components such as glass, frames, and junction boxes can be reprocessed using existing recycling facilities. The recovery of silicon cells is also technically feasible, although the process is more complex. Through thermal and chemical treatments, cell layers can be separated and purified silicon can be obtained. Although large scale silicon recovery is not yet widely implemented, ongoing research continues to improve these processes. The reuse of recycled silicon in new cell production would further enhance the contribution of PV technologies to a circular economy.

Thin film panels, including CdTe and CIGS technologies, require specialized recycling processes due to the presence of rare and potentially toxic elements in their composition. In CdTe panels, the recovery of cadmium and tellurium is particularly critical. Cadmium is a toxic heavy metal if released uncontrollably into the environment. However, in CdTe panels it is chemically bound within the

CdTe compound and encapsulated in glass, which significantly reduces the risk of leakage as long as the panel remains intact. Nevertheless, end of life CdTe panels should never be disposed of in landfills and must instead be treated using appropriate chemical recovery processes to reclaim cadmium and tellurium. First Solar, the earliest and largest CdTe module manufacturer, reports that its closed loop recycling program recovers more than 90 percent of the cadmium and tellurium from its panels. This approach enables the use of recycled tellurium in new module production and reduces dependence on primary mineral resources. Similarly, for CIGS panels, the recovery of copper, gallium, and especially the critical element indium is of high importance. Recycling of thin film modules typically involves mechanical crushing of panels, acid leaching to dissolve metallic components, and subsequent separation of target metals such as cadmium, tellurium, indium, and gallium through precipitation or electrochemical methods. While these processes are chemically advanced, further improvements are still required to enhance their economic viability (Chowdhury et al., 2020).

Perovskite and organic solar cells introduce additional challenges in terms of recycling and waste management. Most current perovskite cells contain lead in their structure, for example in compounds such as MAPbI₃. Because lead is a toxic heavy metal, uncontrolled environmental release from perovskite panels is undesirable. In addition, organic components within perovskite cells may degrade over time and potentially release harmful substances (Dada & Popoola, 2023). For these reasons, the development of effective recycling and waste management strategies is considered essential before large scale commercialization of perovskite PV technologies. Research efforts are also directed toward lead free perovskite materials, such as tin based perovskites. However, lead based perovskites currently exhibit superior efficiency and stability. In the future, when perovskite panels reach the end of their operational life, recovery of glass, plastics, metal electrodes, and elements such as lead and tin will be required. Proposed approaches include electrochemical recycling and solvent based dissolution techniques.

From an environmental perspective, the manufacturing processes of PV systems are also subject to evaluation. Although silicon cell production is energy intensive, PV systems typically generate many times the energy required for their manufacture over their operational lifetime. For example, in Mediterranean regions, the energy payback time of a crystalline silicon PV system is approximately one year, meaning that the system produces the energy consumed during its manufacture within its first year of operation (Fraunhofer ISE, 2025). Over a 25 year lifetime, such systems can deliver roughly twenty five times more energy than that invested in their production. Similar trends are observed for thin

film technologies. Consequently, compared with fossil fuel based energy systems, PV technologies exhibit highly favorable profiles in terms of both greenhouse gas emissions and net energy yield. Nevertheless, the extraction of raw materials used in the PV industry, such as quartz for silicon, cadmium by products for CdTe, and indium for CIGS, along with the use of chemicals such as acids and solvents during panel manufacturing, necessitates careful environmental management. In response, practices such as environmentally friendly process development, the use of recycled materials, and treatment of production waste have gained increasing importance. Overall, establishing effective recycling infrastructure, managing toxic components, and ensuring sustainable material supply are considered critical for improving the environmental profile of PV technologies. Policy instruments such as the European Union WEEE Directive play an important guiding role in this regard (Chowdhury et al., 2020).

5. Future Trends and Next Generation Materials

Rapid research and development activities in the field of solar energy have focused on pushing the limits of existing PV technologies and bringing next generation materials closer to practical application. Several key trends are expected to shape the future of photovoltaics in the coming years.

Tandem and Multijunction Solar Cells: The theoretical efficiency limit of single junction solar cells is approximately 33 percent, as defined by the Shockley-Queisser limit (Tanko et al., 2025). To exceed this limit, the tandem cell concept, which combines materials with different band gaps in series, has attracted considerable interest. In particular, perovskite-silicon tandem cells are regarded as the most promising multijunction technology for near term commercial deployment. In 2022, Oxford PV reported the development of a perovskite-silicon tandem cell with an efficiency of 29 percent and established a dedicated production line (Lira-Cantu, 2025). By 2023, laboratory scale single devices had achieved efficiencies as high as 34.7 percent, thereby surpassing the Shockley-Queisser limit (Lira-Cantu, 2025). In tandem architectures, the perovskite layer acts as a top cell that absorbs high energy photons, while the underlying silicon cell converts lower energy photons, allowing more efficient utilization of the solar spectrum and higher overall power conversion. It is anticipated that tandem cells may partially replace single junction silicon panels in the near future. Accordingly, silicon cell manufacturers have intensified research efforts to adapt existing production lines for tandem compatibility. Beyond perovskite based systems, other combinations such as silicon-CIGS tandems and fully thin film tandem architectures are also under investigation. Although multijunction III-V cells have not achieved widespread terrestrial deployment outside space applications, advances in this field continue to inform

and inspire progress in other PV technologies (Nayak et al., 2019). For example, advanced antireflection coatings and interlayer designs developed for GaInP-GaAs-Ge triple junction cells are being adapted for perovskite tandem systems. In the long term, the commercialization of two, three, or even four junction tandem panels with efficiencies exceeding 40 percent remains an important objective.

New Thin Film and Compound Materials: In addition to conventional CdTe and CIGS technologies, significant research efforts are directed toward thin film PV materials composed of more abundant and less toxic elements. Chalcogenide compounds such as copper zinc tin sulfide have been considered sustainable alternatives due to the abundance of their constituent elements. However, the efficiency of CZTS based solar cells has so far remained below approximately 12 percent, largely due to intrinsic defects and structural disorder that limit performance (Nayak et al., 2019). Researchers continue to address issues related to impurity control and phase segregation in CZTS and related materials. Another promising class of materials includes thin film perovskite derivatives. Lead free perovskites, such as tin based systems, double perovskite structures, and two dimensional or three dimensional hybrid perovskites, are being explored to improve both stability and efficiency (Dada & Popoola, 2023). In addition, fourth generation hybrid cell concepts are being investigated by replacing organic charge transport layers in perovskite devices with more robust inorganic nanomaterials such as carbon nanotubes and graphene derivatives (Kant & Singh, 2022).

Quantum Dot Solar Cells: Quantum dots are nanoscale semiconductor crystals with size tunable band gaps that enable precise control over solar spectrum absorption. Quantum dot solar cells are typically fabricated by depositing colloidal quantum dot solutions as thin films. In theory, quantum dot structures can enhance efficiency through mechanisms such as multijunction effects or hot carrier extraction. In practice, however, the main limitation of quantum dot cells has been inefficient charge transport between nanocrystals, as organic ligands on quantum dot surfaces introduce resistance to carrier flow (Nayak et al., 2019). Nevertheless, recent progress has been achieved through ligand shortening, improved ordering of quantum dots within matrices, and the development of quantum dot sensitized tandem architectures. For example, lead sulfide quantum dot cells reported efficiencies approaching 18 percent in 2021, and hybrid quantum dot–perovskite systems have been shown to approach the 20 percent threshold (Nayak et al., 2019). Quantum dots are also employed in spectral management strategies such as down conversion and up conversion, which aim to utilize portions of the solar spectrum that are otherwise poorly absorbed by conventional cells. Through these approaches, quantum dot

technologies are expected to increasingly integrate with tandem architectures in future PV systems.

Organic and Flexible Solar Cells: Organic solar cells based on polymers and small molecules have gained attention for niche applications due to their solution processability and mechanical flexibility, including wearable electronics and building integrated photovoltaics such as semi transparent windows. Laboratory scale efficiencies of organic PV cells have exceeded 17 percent in recent years. Key advantages of organic technologies include their potential for transparency and color tunability, which facilitates aesthetic integration into buildings. However, limited operational lifetime and stability have constrained broader commercial adoption. Recent advances have brought organic PV modules closer to operational lifetimes exceeding five years. Dye sensitized solar cells, which combine organic dyes with nanoporous titanium dioxide structures, also offer low cost fabrication and strong indoor performance. Despite these advantages, long term stability issues associated with liquid electrolytes have limited their durability. Improved designs employing solid state electrolytes or perovskite sensitizers have achieved efficiencies in the range of 11 to 13 percent along with enhanced stability (Kant & Singh, 2022). In the future, organic PV and dye sensitized technologies are expected to find wider use in specific niche applications that prioritize flexibility, aesthetics, and lightweight design.

Large Scale Manufacturing and Technology Integration: To translate laboratory achievements into real world deployment, scalable manufacturing techniques are being actively developed for emerging PV materials. For perovskite cells, methods such as roll to roll printing, spray coating, and spin coating are being explored for large area film production (Dallaev et al., 2023). Similar scalable printing and coating approaches are suitable for organic and quantum dot based materials. Production on flexible polymer substrates further enables lightweight and portable PV modules. Another important trend involves the integration of new materials into existing silicon manufacturing infrastructure. For example, heterojunction cell technology combines amorphous silicon thin films with crystalline silicon substrates to achieve high efficiency devices. In a similar manner, perovskite top cells can be integrated onto silicon wafers in tandem configurations without fundamentally redesigning production lines. Such hybrid integration strategies are expected to accelerate the commercialization of next generation PV technologies.

In summary, material and design innovations in photovoltaics are advancing at an unprecedented pace. While crystalline silicon technology has demonstrated long term maturity and reliability, thin film and next generation technologies are rapidly approaching competitive levels in terms of both efficiency and cost. When combined with auxiliary efficiency enhancement techniques, the performance

and durability of solar panels continue to improve. Future developments are expected to include wider adoption of tandem architectures, progress in lead free and earth abundant thin film materials, and the emergence of new applications enabled by quantum dot and organic technologies. Together, these trends are likely to increase the share of solar energy in the global energy mix and contribute to a more sustainable energy future.

References

- Al-Ezzi, A. S., & Ansari, M. N. M. (2022). Photovoltaic solar cells: A review. *Applied System Innovation*, 5(4), 67. <https://doi.org/10.3390/asi5040067>
- Chowdhury, M. S., Rahman, K. S., Chowdhury, T., Nuthammachot, N., Techato, K., Akhtaruzzaman, M., ... Amin, N. (2020). An overview of solar photovoltaic panels' end-of-life material recycling. *Energy Strategy Reviews*, 27, 100431. <https://doi.org/10.1016/j.esr.2019.100431>
- Crowell, C. (2021, December 2). Benchmark costs for five solar PV module technologies. Solar Builder. Retrieved from <https://solarbuildermag.com/pv-modules/benchmark-costs-for-five-solar-pv-module-technologies/>
- Dada, M., & Popoola, P. (2023). Recent advances in solar photovoltaic materials and systems for energy storage applications: A review. *Beni-Suef University Journal of Basic and Applied Sciences*, 12, 66. <https://doi.org/10.1186/s43088-023-00405-5>
- Dallaev, R., Pisarenko, T., Papež, N., & Holcman, V. (2023). Overview of the current state of flexible solar panels and photovoltaic materials. *Materials*, 16(17), 5839. <https://doi.org/10.3390/ma16175839>
- Elsafi, A., Aïssa, B., Ilse, K., & Abdallah, A. (2025). Performance and durability of anti-soiling and anti-reflective coatings for photovoltaic systems in desert climates. *Solar Energy*, 293, 113446. <https://doi.org/10.1016/j.solener.2025.113446>
- Fırat, M., Şenocak, Ş. M., Okcu, M., Varol, Y., & Altun, Ş. (2023). Ozone-assisted combustion and emission control in RCCI engines: A comprehensive study. *Energy*, 284, 129247. <https://doi.org/10.1016/j.energy.2023.129247>
- Fraunhofer Institute for Solar Energy Systems ISE. (2025). Photovoltaics report (updated October 31, 2025). Retrieved December 24, 2025, from <https://www.ise.fraunhofer.de/content/dam/ise/de/documents/publications/studies/Photovoltaics-Report.pdf>
- Gürgeç, E., Öztıp, H. F., Yamaç, H. İ., Aksu Canbay, C., Şenocak, Ş. M., Özabacı, M., Gürgeç, T., & Gür, M. (2025a). B₄C-based nanoenhancement on the thermophysical and stability performance of solar salt: A novel approach for high-temperature TES applications. *Case Studies in Thermal Engineering*, 75, 107257. <https://doi.org/10.1016/j.csite.2025.107257>
- Gürgeç, E., Öztıp, H. F., Şenocak, Ş. M., Aktemur, C., Gürgeç, T., Varol, Y., Yamaç, H. İ., & Gür, M. (2025b). Hafnium carbide as a novel nanofiller for RT64HC phase change materials: Enhancing thermal conductivity, heat capacity, and cycling stability. *Thermal Science and Engineering Progress*, 69, 104257. <https://doi.org/10.1016/j.tsep.2025.104257>

- Kant, N., & Singh, P. (2022). Review of next generation photovoltaic solar cell technology and comparative materialistic development. *Materials Today: Proceedings*, 56, 3460–3470. <https://doi.org/10.1016/j.matpr.2021.11.116>
- Kozak-Jagiela, E., Cisek, P., & Ocłoń, P. (2023). Cooling techniques for PV panels: A review. *Scientiae Radices*, 2(1), 47–68. <https://doi.org/10.58332/scirad2023v2i1a03>
- Lira-Cantu, M. (2025). Perovskite solar cells: Technology development in China. *APL Energy*, 3(1), 010401. <https://doi.org/10.1063/5.0263637>
- Nayak, P. K., Mahesh, S., Snaith, H. J., & Cahen, D. (2019). Photovoltaic solar cell technologies: Analysing the state of the art. *Nature Reviews Materials*, 4(4), 269–285. <https://doi.org/10.1038/s41578-019-0097-0>
- Öztürk, G., Şenocak, Ş. M., Şenocak, N., & Fırat, M. (2024). An experimental investigation of 1,2-dimethoxy ethane as a fuel additive in biodiesel-fueled diesel engine. *Journal of the Energy Institute*, 117, 101824. <https://doi.org/10.1016/j.joei.2024.101824>
- Sampaio, P. G. V., & González, M. O. A. (2017). Photovoltaic solar energy: Conceptual framework. *Renewable and Sustainable Energy Reviews*, 74, 590–601. <https://doi.org/10.1016/j.rser.2017.02.081>
- Sharma, R., Sharma, A., Agarwal, S., & Dhaka, M. S. (2022). Stability and efficiency issues, solutions and advancements in perovskite solar cells: A review. *Solar Energy*, 244, 516–535. <https://doi.org/10.1016/j.solener.2022.08.001>
- Svarc, J. (2025, November 9). Most efficient solar panels 2025. Clean Energy Reviews. Retrieved from <https://www.cleanenergyreviews.info/blog/most-efficient-solar-panels>
- Şenocak, Ş. M., Gürgeç, T., Varol, Y., Gürgeç, E., & Öztop, H. F. (2025). Cubic boron nitride enhanced PCM nanocomposites for solar energy storage and thermal management. *Solar Energy*, 303, 114102. <https://doi.org/10.1016/j.solener.2025.114102>
- Tanko, K. T., Tian, Z., Raga, S., Xie, H., Katz, E. A., & Lira-Cantu, M. (2025). Stability and reliability of perovskite photovoltaics: Are we there yet? *MRS Bulletin*, 50(4), 512–525. <https://doi.org/10.1557/s43577-025-00863-5>

CHAPTER 24

Hybrid Laser Aluminum Welding of Characteristics and Industrial Applications

Ferit Artkin¹

¹ Lect. Dr., Kocaeli University, ORCID: 0000-0002-8543-6334

Introduction

The limitations of conventional welding techniques have become clear as sectors like automotive, aerospace, and railway manufacture depend more and more on aluminum to decrease weight and boost efficiency. Because of its strong oxide layer, high heat conductivity, and great reflectivity, aluminum is infamously difficult to weld.

A crucial remedy that closes the gap between accuracy and durability is hybrid laser-arc welding, or HLAW. Hybrid Laser-Arc Welding (HLAW) produces a synergistic effect where the benefits of both sources are maximized and their respective drawbacks are minimized by combining a concentrated laser beam with a Gas Metal Arc Welding (GMAW/MIG) electrical arc in a single process zone.

Hybrid laser aluminum welding is a cutting-edge joining technique that integrates Gas Metal Arc Welding (GMAW/MIG) and Laser Beam Welding (LBW), two separate welding processes, into a single, concurrent operation.

By utilizing the advantages of both a focused laser beam and a wide electric arc, this "hybrid" technique is particularly made to get over the inherent challenges of welding aluminum, such as its high reflectivity and thermal conductivity.

Fundamental Specifications of Hybrid Laser Aluminum Welding

The fundamental specifications of Hybrid Laser-Arc Welding (HLAW) for aluminum are distinct from conventional welding because they must balance two heat sources to maintain a stable keyhole while preventing common aluminum defects like porosity and hot cracking (Casalino, G., et al., 2010).

The most popular use of this technique is for welding industrial 6000-series aluminum (such as 6061-T6 and 6082), and the normal operating requirements are shown below.

In hybrid welding, the choice of filler wire is crucial to avoid "hot cracking" (solidification cracking), which occurs when 6000 series aluminum is welded autogenously (without filler). Aluminum alloys 5XXX (Al-Mg) or 6XXX (Al-Mg-Si) serve as the base material. Wire Filler: The most popular is ER4043 (AlSi₅). greatly lessens the susceptibility to cracks, but the weld strength is reduced.

ER5356 (AlMg₅): Slightly more prone to cracking in the 6XXX family, but used if color matching (anodizing) or greater shear strength is needed.

Gap Tolerance: 0.0 to 1.0 mm (small gaps that a laser alone could not bridge are made possible by the arc component).

The essential operating parameters that must be carefully regulated to produce a stable process and a high-quality weld are referred to as the fundamental specifications of Hybrid Laser Aluminum Welding (HLAW), particularly when handling aluminum's difficult characteristics (high reflectivity, high thermal conductivity, and susceptibility to porosity/cracking) (Bunaziv, I., et al., 2021), (Yan, S., et al., 2020).

Table 1. Fundamental Process Specifications (Typical Ranges) These parameters apply generally to material thicknesses between 2mm and 6mm 6000-series aluminum (e.g., 6061-T6, 6082). (ISO 15615:2013, AWS C7.4 American Welding Society, Eriksson, I., et al., 2013).

Parameter	Specification Range	Function/Note
Laser Power	3.0 – 6.0 kW	Solid-state laser (Fiber or Disk, $\lambda \sim 1 \mu\text{m}$). Power determines penetration depth.
Welding Speed	2.0 – 6.0 m/min	Significantly faster than MIG (typ. 0.8 m/min). High speed reduces heat input.
Arc Process	Pulse MIG or CMT	Pulsed arc is required to break the oxide layer and stabilize droplet transfer.
Arc Current	150 – 280 Amps	Provides fill material and controls bead width.
Arc Voltage	20 – 28 Volts	Controls arc length; must be minimized to reduce heat.
Wire Feed Speed	4.0 – 8.0 m/min	Matched to travel speed to fill the joint without over-reinforcement.
Laser-Arc Distance	2 – 4 mm	Distance between the laser focal point and the arc wire tip.
Defocus Position	0 to -2 mm	Focus is typically on the surface or slightly inside the material.
Shielding Gas	Ar + He (30-50%)	Helium is often added to Argon to increase heat and help gas escape (reducing porosity).

The use of laser hybrid arc welded joints (HLAW) in joints with various Al alloys is growing in popularity, particularly in the contemporary constructions of high-speed trains and cars (F., Nie et al., 2018, Shaohua Y., el al., 2020) Friction stir welding (FSW) is typically used to weld various Al alloys, particularly those in the 7XXX family (U. Donatus, et al., 2018, A. Barbini, et al., 2018, X.H. Zeng, et al., 2019). This is due to the fact that FSW yields stronger joints of 7XXX Al alloys than laser arc welding.

Key Optimization Principle (Priorities): In optimizing the hybrid process for aluminum, research has shown that the factors influencing joint strength and weld

quality have a distinct priority order, as they control the fundamental defects; (Du L., et al., 2023).

Arc Current (I_A): Has the greatest influence on final joint strength, as it governs the amount of filler metal and bead shape.

Laser Power (P_L): The primary influence on penetration and fusion depth (Adisa SB., et al., 2018).

Welding Speed (V): Critical for controlling the overall heat input and minimizing HAZ-related softening (lasertechnologies.co.in, 2025).

Welding Process of Aluminum Materials with Laser Hybrid Welding Method

In addition to welding thin metal sheets, it may be used to link various materials to create a lightweight structural design, as seen in the car sector. A laser riveting procedure is proposed to replace traditional thermal joining processes, with the primary issue being the creation of intermetallic phases from distinct alloying elements (E. Schubert et al. 2001). Figure 1 depicts the laser riveting concept as well as an example of attaching steel to aluminum, which is typically thought to be difficult or impossible to do using conventional thermal techniques. This laser riveting process can also be useful for joining steel and other non-metallic materials such as resin and carbon fiber reinforced plastic (CFRP).

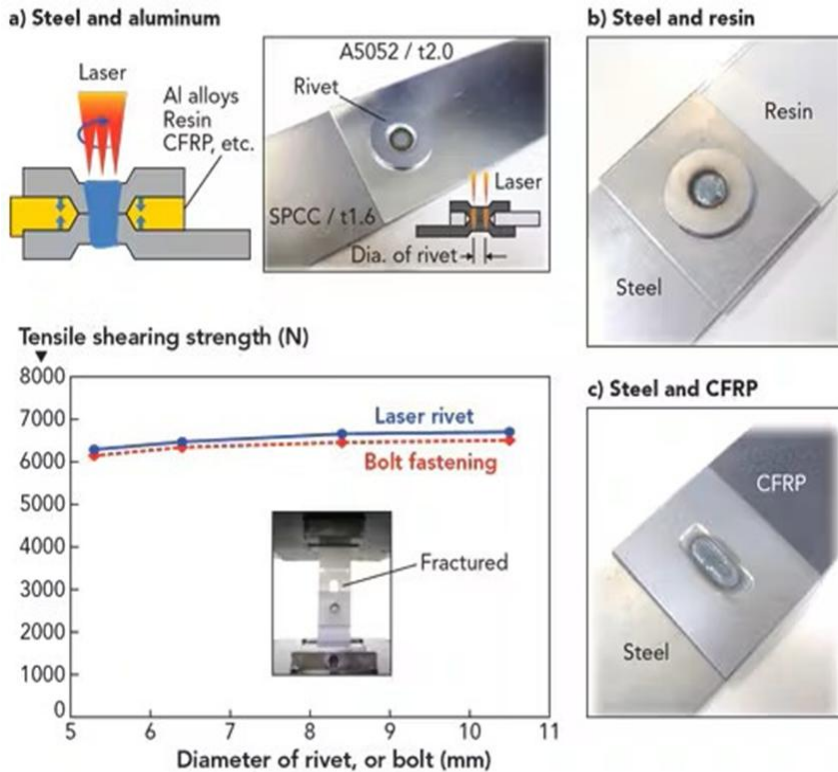


Figure 1. Laser riveting process can join steel to aluminum, resin or CFRP (laserfocusworld.com, 2024).

Laser-MIG hybrid welding is a welding procedure that combines two heat sources: a laser and an arc (Yan, S.H et al. 2014, Su, J., et al. 2023, Vorontsov, A., 2022). The inhibitory impact of the plasma on the laser may be efficiently decreased by the coupling effect between the laser and the arc, lowering laser energy reflection off aluminum alloys. The laser also compresses the arc, assuring arc stability and increasing energy density; as a result, the energy consumption rate of the two heat sources improves concurrently.

The impact of laser-arc hybrid welding settings on the porosity of an 8 mm thick AA6082-T6 aluminum alloy was examined by Zhang et al. (Zhang, C., et al., 2017). The weld porosity was less than 0.5% under optimal welding conditions, and there was a particular correlation between the molten pool's characteristics and the weld porosity. The weld cross-section progressively changed from Y-shaped to V-shaped, and the weld porosity decreased as the arc current or laser power was raised or the welding speed was reduced.

Another research Huang, L.J., et al. was carried out by using laser-MIG hybrid welding of 10 mm thick 5083 aluminum alloy. (Huang et al., 2018). Weld

formation and joint mechanical qualities were examined in relation to the sequence of laser and MIG heat sources. It shown that when the laser heat source was given priority, the weld formation improved and the weld porosity was much reduced. Additionally, the weld metal had better mechanical qualities since the second phase particles were evenly dispersed throughout it. An aluminum alloy was also laser-MIG hybrid welded by (Cai et al., 2019). Porosity flaws, plasma physical characteristics, and weld penetration depth were examined in relation to various He and Ar-He mixed gas volume ratios.

When aluminum alloys are fused with a laser beam, low boiling point metals like magnesium evaporate and burn during the welding process. The Mg element is a component of the Mg_2Si reinforcing phase found in 6XXX aluminum alloys. The burn-off loss of the Mg element reduces the reinforcing phase at the joints, resulting in a drop in the joint's mechanical characteristics. The filler wire is injected into the laser-MIG hybrid welding process and can supplement the alloy elements to some extent. The composition and distribution of alloy components have a significant impact on the mechanical characteristics of welded joints.

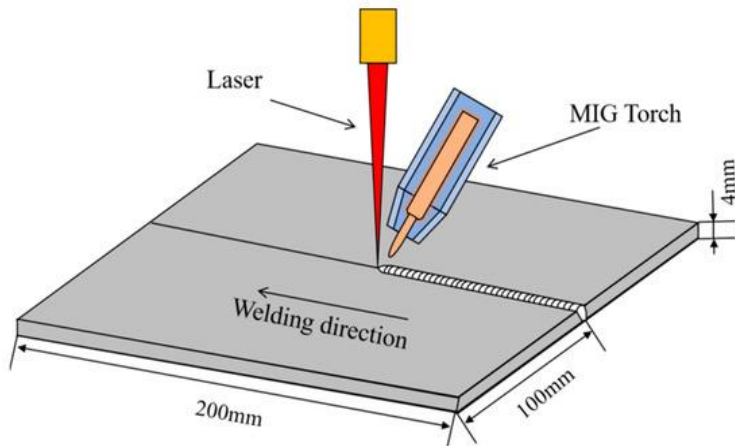


Figure 2. Schematic Diagram of the Laser-MIG Hybrid Aluminum Welding Process (Zhang, Wen, et al., 2024).

The H1 joint has the highest microhardness; the top portion of the FZ has an average hardness of 66 HV, the lower portion has an average hardness of 70 HV, the upper portion of the HAZ has an average hardness of 86 HV, and the bottom portion has an average hardness of 88 HV. The top portion of joint H1 has a little lower hardness than the lower section, as seen in Figure 3b. The MIG heat source in the laser-MIG hybrid welding technique is comparable to a "tempering" treatment for the top portion of the FZ. It causes the microstructure to become homogeneous, the grain to develop, and the mechanical characteristics to

somewhat deteriorate. The microhardness of the middle part of joints L, H1, H2, and H3, as well as the microhardness of the upper and lower parts of joint H1 are tested, as shown in Figure 4.

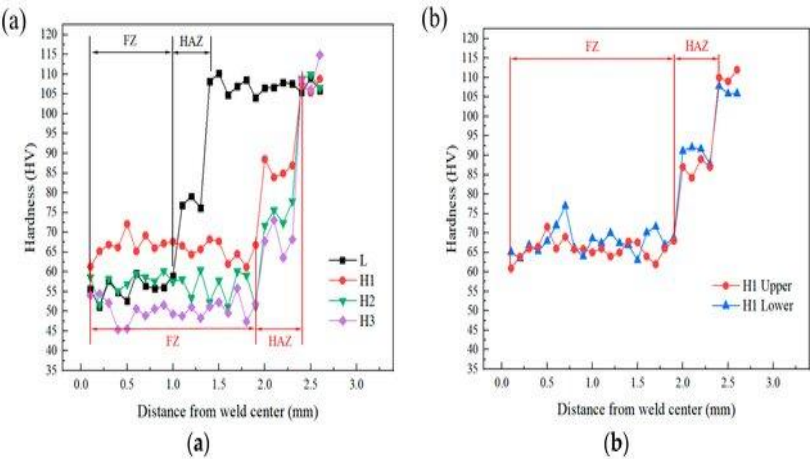
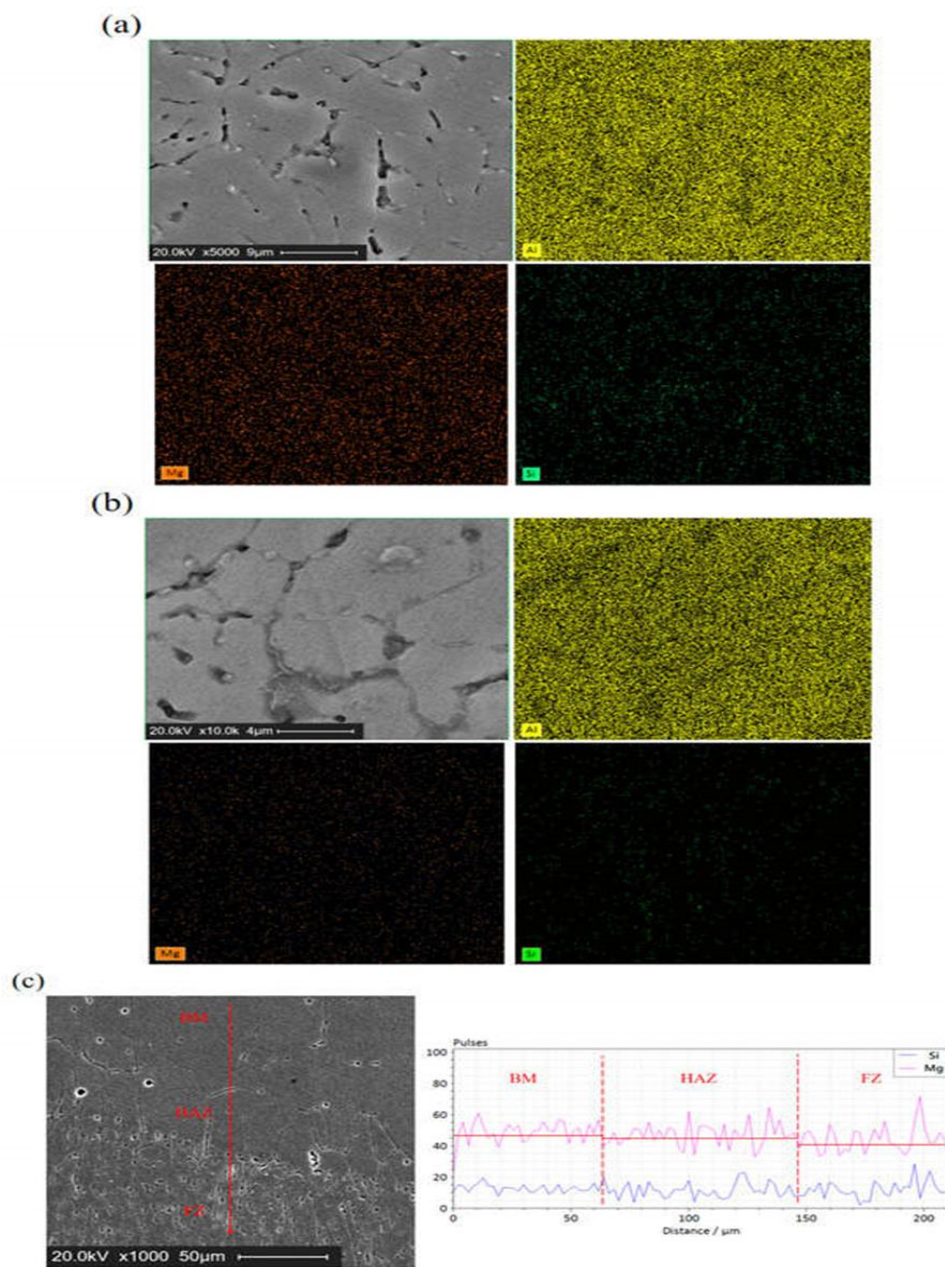


Figure 3. Schematic Diagram of the Laser-MIG Hybrid Aluminum Welding Process (Zhang, Wen, et al., 2024).

The lowest-hardness zone of the LBW joint and the three kinds of laser-MIG hybrid welding joint is the FZ. The hardness of the HAZ is increased to a certain extent while the highesthardness zone is the BM. However, the heat input of the MIG heat source is low; thus, the grain growth is not obvious and the hardness of the upper part of the FZ decreases slightly. For the HAZ, the MIG heat source mainly acts on the upper part of the molten pool and has little effect on the lower part. The upper part of the HAZ suffers a larger weld thermal cycle and a greater degree of “softening” occurs. The average hardness of the FZ of joint L is 55 HV and the average hardness of the HAZ is 77 HV. In the welding process of joint L, the Mg element in the FZ generates obvious evaporation and burning loss. The Mg element is the constituent element of the strengthening phase Mg2Si in Al-Mg-Si alloys, and the burning loss of the Mg element decreases the quantity of the strengthening phase in the FZ, which results in the decrease in strength and hardness of the joint. The average hardness of the FZ of joints H2 and H3 is 54

HV and 48 HV, respectively, and the average hardness of the HAZ is 75 HV and 70 HV, respectively, which is obviously lower than that of joint H1.



The Mg and Si elemental distributions of the L and H1 junctions, respectively, are displayed in the SEM of the FZ of the junctions, as seen in Figure 4. At the weld of the two joints, the magnesium element is evenly distributed and shows little segregation. At the grain boundary, the Si element is mostly concentrated in the white particle phase and exhibits some degree of segregation. The white particles improve the mechanical qualities of the welded connection by increasing the resistance of dislocation migration as it approaches the grain boundary.

Advantages and Disadvantages of Hybrid Laser Aluminum Welding

In the field of advanced manufacturing, Hybrid Laser Aluminum Welding (HLAW)-typically the Laser-MIG (Metal Inert Gas) process-is regarded as a "synergistic" solution that combines the deep penetration of a laser with the gap-bridging capabilities of an arc.

Superior Gap Bridging: Standard laser welding requires extremely tight fit-up tolerances (often < 0.1 mm). HLAW uses the MIG arc and filler wire to bridge significant joint gaps (>0.5 mm), making it more viable for large-scale industrial parts. (Ribic et al., 2009).

Reduced Thermal Distortion: The high travel speed and concentrated energy of the laser result in a lower total heat input compared to traditional MIG. This minimizes the Heat-Affected Zone (HAZ) and prevents warping (Volodymyr, K., et al., 2021).

Higher Welding Productivity: HLAW can increase welding speeds by a factor of 1.8 to 2.6 compared to traditional MIG, allowing for single-pass welding of thick aluminum plates (up to 15 mm) (Costanza G., et al., 2025).

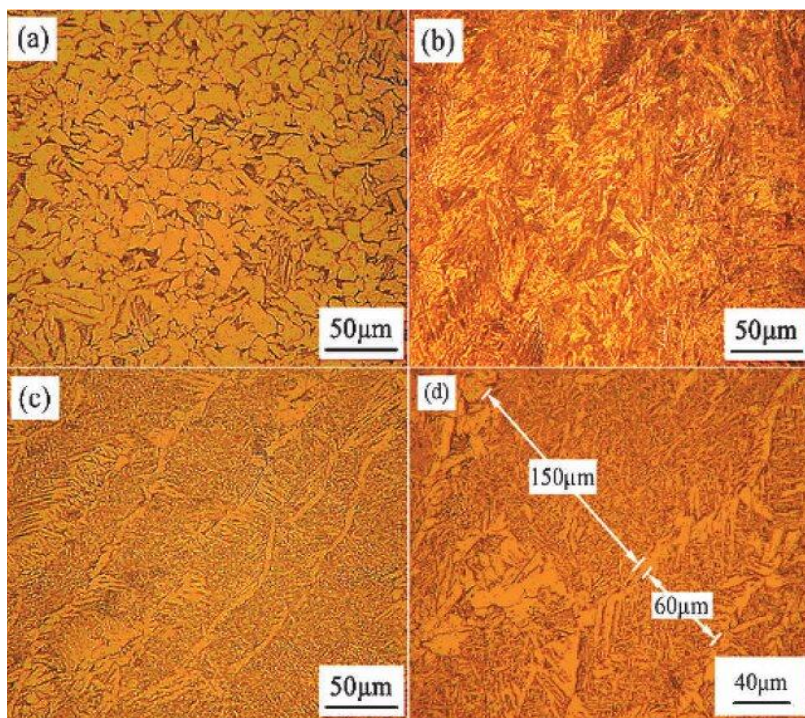


Figure 5.

Micrographs of mild steel a base metal, b laser, c arc and d hybrid weld microstructures. The laser and hybrid weld laser power was 4.5 kW. The GTA and hybrid welds' arc current was 180 A. The welding speed for all welds was 13.33 mm s^{-1} (M.Gao, et al., 2008)(Ribic et al., 2009).

Overcoming High Reflectivity: Aluminum naturally reflects up to 90% of infrared laser energy. The hybrid arc pre-heats the surface and creates a plasma that stabilizes laser absorption (Zhang W., et al., 2024).

Metallurgical Customization: The addition of filler wire allows for the control of the weld pool chemistry, which is essential for preventing solidification cracking in high-strength alloys like the 6000 and 7000 series (ASM Handbook, Volume 06A, 2011).

Disadvantages of HLAW; High Capital Investment (CAPEX): The initial cost is significantly higher than traditional methods, requiring both high-power laser sources and robotic MIG systems, which can limit its use to high-volume industries (Casalino, G., et al., 2010).

Extreme Process Complexity: The synergy depends on over 20 interacting parameters (e.g., laser-arc distance, power ratios, shielding gas composition). This makes the setup technically challenging and sensitive to fluctuations (Wang Z., et al., 2023).

Sensitivity to Keyhole Instability: If parameters are not perfectly optimized, the interaction between the metal vapor and the arc can cause keyhole collapse, leading to significant porosity in the weld (Lindeberg N., B., 2022).

Maintenance of Optics: The MIG arc generates spatter and fumes that can contaminate or damage expensive laser optics, requiring frequent cleaning and the use of protective cross-jets (Allen, C.M., et al., 2007).

Safety and Environmental Requirements: High-power lasers require specialized "Class 1" safety enclosures and advanced fume extraction systems to protect operators from reflected radiation and harmful aluminum oxides (fanucifalcon.com, 2025).

Conclusion

In conclusion, Hybrid Laser Aluminum Welding (HLAW), which usually combines a laser beam with Metal Inert Gas (MIG) welding, is a game-changing technique that addresses the inherent challenges of welding aluminum, according to research findings and industrial studies. The hybrid technique generates a "synergy" that offers the best of both worlds, whereas pure laser welding is too sensitive to joint gaps and standard arc welding is too sluggish and causes deformation. The industry as a whole believes that hybrid laser aluminum welding is the best option for high-productivity manufacturing.

Control: In-situ monitoring is replacing the laborious, traditional post-weld NDT (Radiography, conventional UT) in characterization. HLAW is anticipated to become the norm for bonding structural aluminum in the shift to electric mobility as fiber laser prices continue to decline and AI-driven monitoring systems advance. The evolution of Hybrid Laser Aluminum Welding (HLAW) from a high-performance specialized technology to a clever, widely used production standard will determine its destiny.

The aerospace industry's drive for "fastening-free" structures and the 2025 spike in EV manufacturing are driving technological advancements toward autonomous, self-correcting systems. Simple hardware is giving way to integrated "intelligent" systems in the upcoming generation of HLAW: AI and Machine Learning (ML): AI is currently used in modern systems to keep an eye on the weld pool in real time. The AI may modify laser power or wire feed speed mid-weld to stop typical aluminum flaws like porosity or solidification cracking before they happen by evaluating information from thermal and acoustic sensors.

Blue and Green Laser Integration: Aluminum's high reflectivity has been a significant obstacle. Blue (450 nm) or Green (515 nm) lasers, which aluminum absorbs far more efficiently than conventional infrared lasers and enable more stable welds with less energy, are increasingly being used in future hybrid systems. Manufacturers are employing "Digital Twins" to electronically

replicate the welding of intricate aluminum systems. This enables "prehearsal" of the weld and process optimization, which can reduce physical trial-and-error by as much as 50%. Beam Shaping Technology: To manage the weld's cooling pace, which is essential for preserving the strength of high-performance 6000 and 7000 series aluminum alloys, new "Wobble" and tunable ring-mode lasers are being combined with MIG arcs.

Acknowledge

In the preparation of this study, I contributed to the MAK579 Advanced Welding and 3D Additive Manufacturing lecture notes that I received from Prof. Dr. Mustafa KOÇAK during the Mechanical Engineering graduate course phase. I would like to thank Prof. Dr. Mustafa Koçak. I would also like to thank Dr. Ozan ÇOLAK, Assistant Professor, who assisted in the course.

References

- A. Barbini, J. Carstensen, J.F. dos Santos, Influence of a non-rotating shoulder on heat generation, microstructure and mechanical properties of dissimilar AA2024/AA7050 FSW joints, *J Mater Sci Technol*, 34 (1) (2018), pp. 119-127.
- Adisa SB., Loginova I., Khalil A., Solonin A., Effect of Laser Welding Process Parameters and Filler Metals on the Weldability and the Mechanical Properties of AA7020 Aluminium Alloy. *Journal of Manufacturing and Materials Processing*. 2018; 2(2):33. <https://doi.org/10.3390/jmmp2020033>.
- Allen, C.M., Gerritsen, C., Zhang, Yanhui, Mawella, J. (2007). *Hybrid laser-MAG Welding Procedures and Weld Properties*.
- ASM Handbook, Volume 06A - Welding Fundamentals and Processes, Lienert, Thomas J.; Babu, Sudarsanam Suresh; Siewert, Thomas A.; Acoff, Viola L., Publisher ASM International, 2011 ISBN978-1-61503-133-7.
- AWS C7.4 (American Welding Society): Process Specification and Operator Qualification for Laser Beam Welding. This standard is often adapted for hybrid processes in US aerospace/automotive sectors.
- Bunaziv, I., Akselsen, O. M., et al. (2021). "Laser Beam and Laser-Arc Hybrid Welding of Aluminium Alloys." Published in *Metals* (MDPI).
- Casalino, G., Maso, U., Angelastro, Andrea, Campanelli Sabina Luisa. 2010. Hybrid Laser Welding: A Review. *DAAAM International Scientific Book*. 413-430. [10.2507/daaam.scibook.2010.38](https://doi.org/10.2507/daaam.scibook.2010.38).
- Casalino, G., Maso, U., Angelastro, Andrea, Campanelli, Sabina Luisa. (2010). Hybrid Laser Welding: A Review. *DAAAM International Scientific Book*. 413-430. [10.2507/daaam.scibook.2010.38](https://doi.org/10.2507/daaam.scibook.2010.38).
- Costanza G, Giudice F, Missori S, Scolaro C, Sili A, Tata ME. An Overview of the Working Conditions of Laser–Arc Hybrid Processes and Their Effects on Steel Plate Welding. *Journal of Manufacturing and Materials Processing*. 2025; 9(8):248. <https://doi.org/10.3390/jmmp9080248>.
- Du L., Yang Z., Wang X., Welding Characteristics of Laser-MIG Hybrid Welding of Arc-Welded Aluminum Profiles for High-Speed Trains. *Materials*. 2023; 16(1): 404. <https://doi.org/10.3390/ma16010404>.
- E. Schubert, M. Klassen, I. Zerner, C. Walz, and G. Sepold, *J. Mater. Process. Technol.*, 115, 2–8 (2001).
- Eriksson, I., et al. (2013). "Guidelines in the Choice of Parameters for Hybrid Laser Arc Welding with Fiber Lasers." Published in *Physics Procedia*.

- F., Nie, H. Dong, S. Chen, P. Li, L. Wang, Z. Zhao, et al., Microstructure and mechanical properties of pulse MIG welded 6061/A356 aluminum alloy dissimilar butt joints. *J Mater Sci Technol*, 34 (3) (2018), pp. 551-560.
- <https://fanuci-falcon.com/knowledge-base/a-guide-to-laser-welding-everything-you-need-to-know>, 2025.
- <https://www.lasertechnologies.co.in/laser-welding/fundamentals-of-laser-welding>
Accessed 12 December 2025.
- ISO 15615:2013: Gas welding equipment - Acetylene manifold systems for welding, cutting and allied processes-Safety requirements in high-pressure devices. (While a general safety standard, ISO 12932 pertains specifically to Hybrid Laser-Arc welding qualifications).
- Lindeberg Nilsson, B. (2022). Laser Arc Hybrid Welding of Thicker 6000-series Aluminium Alloy : Laser-MIG butt welding of 8 mm 6082-t6 (Dissertation). Retrieved from <https://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-314915>.
- M. Gao, X. Y. Zeng, J. Yan and Q. W. Hu: *Appl. Surf. Sci.*, 2008,254, 5715–5721.
- Ribic B, Palmer TA, DebRoy T. Problems and issues in laser-arc hybrid welding. *International Materials Reviews*. 2009; 54(4):223-244. doi:10.1179/174328009X411163.
- Shaohua Yan, Chuanping Ma, Hui Chen, Modifying microstructures and mechanical properties of laser-arc welded joints of dissimilar advanced aluminum alloys, *Mater Charact*, 164 (2020), p. 110331.
- U. Donatus , G.E. Thomson, M.I. Momoh, N.B. Maledi, I-L. Tsai, Raphael O. F., Z. Liu, Variations in stir zone and thermomechanically affected zone of dissimilar friction stir weld of AA5083 and AA6082 alloys, *Transactions of Nonferrous Metals Society of China*, Volume 28, Issue 12, 2018, Pages 2410-2418, ISSN 1003-6326, [https://doi.org/10.1016/S1003-6326\(18\)64887-3](https://doi.org/10.1016/S1003-6326(18)64887-3).
- Volodymyr, K., Khaskin, V., Grinyuk, A.A., Shcheretskyi, V. Tymofii, O., Babych, O. (2021). Hybrid Laser-MIG Welding of Aluminum Alloys Al-Mg-Mn, Al-Cu-Mg and Al-Mg-Li Systems. *Environment Technology Resources Proceedings of the International Scientific and Practical Conference*. 125-132.
- Wang Z., Gong M., Zhou L., Gao M., A Review of Numerical Simulation of Laser-Arc Hybrid Welding. *Materials (Basel)*. 2023 May 6;16(9):3561. doi:10.3390/ma16093561.
- X.H. Zeng, P. Xue, L.H. Wu, D.R. Ni, B.L. Xiao, K.S. Wang, et al., Microstructural evolution of aluminum alloy during friction stir welding under different tool

rotation rates and cooling conditions, J Mater Sci Technol, 35 (6) (2019), pp. 972-981.

Yan, S., et al. (2020). "Hybrid laser-metal inert gas welding of Al-Mg-Si alloy joints." Published in Materials & Design.

Zhang W., Wang S., Investigation of Laser-MIG Hybrid Welding of Al-Mg-Si Aluminum Alloy. Metals. 2024; 14(6):729. <https://doi.org/10.3390/met14060729>.

Zhang, Wen, Wang, Shaogang. (2024). Investigation of Laser-MIG Hybrid Welding of Al-Mg-Si Aluminum Alloy. Metals. 14. 729. 10.3390/met14060729.

CHAPTER 25

Synthesis, End-Of-Life, and Economics of Bioplastics

Gamze Özçakır¹

¹ Res. Assist. Dr., Bilecik Şeyh Edebali University, Faculty of Engineering, Department of Chemical Engineering, TR 11100, Bilecik, Türkiye, ORCID: 0000-0003-0357-4176

1. INTRODUCTION

The growing global problem of plastic waste and the push for sustainable resource use have accelerated research into bioplastics. These materials come from renewable biomass sources and are often biodegradable [1].

Bioplastics are broadly categorized based on their source (bio-based) and their end-of-life properties (biodegradable):

- Bio-based and Biodegradable: Examples include Polylactic Acid (PLA) and Polyhydroxyalkanoates (PHAs).
- Bio-based but Non-Biodegradable (Drop-in Bioplastics): Examples include Bio-Polyethylene (Bio-PE) and Bio-Polyethylene Terephthalate (Bio-PET), which are chemically identical to their fossil-fuel counterparts [2].
- The primary renewable resources for bioplastic synthesis are:
 - Starch and Sugarcane (for PLA, Bio-PE, Bio-PET).
 - Plant oils and animal fats.
 - Lignocellulosic biomass (wood powder, agricultural waste).
- Microorganisms like bacteria and microalgae (for PHAs) [3].

This chapter reviews current and emerging synthesis methods for the major types of bioplastics, examines end-of-life options for bioplastics, and considers the possibility of their commercialization.

2. BIOPLASTIC SYNTHESIS

2.1 Major Synthesis Pathways

The current synthesis methods can be grouped into three primary pathways: Chemical Synthesis, Microbial Fermentation, and Direct Extraction/Modification.

2.1.1 Chemical synthesis

This route typically involves converting a biomass-derived sugar into a monomer, which is then chemically polymerized. PLA, Bio-PE, and Bio-PET can be produced by utilizing this route.

Chemical synthesis of Polylactic Acid (PLA) covers fermentation, purification, oligomerization, and polymerization steps. At first, starch in corn or sugarcane is fermented by microorganisms to produce Lactic Acid. Then, Lactic acid is purified and either directly polymerized or converted into a cyclic dimer

called Lactide. In the last step of the process, PLA resin is produced, most commonly through Ring-Opening Polymerization (ROP) of the Lactide monomer [4].

Chemical synthesis of Bio-Polyethylene (Bio-PE) and Bio-Polyethylene Terephthalate (Bio-PET) is the same as in petrochemical-based plastics. At this time, starch in corn or sugarcane is fermented by microorganisms to produce ethanol. Then, ethylene monomer is produced by the dehydration of ethanol. After ethylene polymerization, Bio-PE is synthesized. The scope of synthesis of Bio-PET has differed slightly from that of Bio-PE and PLA. Bio-PET production is composed of the polymerization of bio-derived monoethylene glycol (MEG) and fossil-derived terephthalic acid (TPA) [5].

2.1.2 Microbial fermentation

This route utilizes microorganisms to synthesize and accumulate the polymer inside their cells as an energy reserve, using organic substrates as their carbon source. Polyhydroxyalkanoates (PHAs) can be synthesized using this route.

For PHA synthesis via microbial fermentation, cultivation, accumulation, and extraction and purification steps can be followed. In cultivation, bacteria (e.g., *Cupriavidus necator*) are grown in bioreactors and fed carbon-rich substrates (e.g., sugars, fatty acids, or even food waste). Then, the conditions enable the diversion of the carbon source into PHAs intracellularly in the bacteria. Eventually, the PHA polymer is extracted from the bacterial cells, typically using solvent-based methods, and purified.

Novel technologies focus on utilizing microalgae and cyanobacteria in controlled photobioreactor systems to synthesize polyhydroxyalkanoates (PHAs) more sustainably, utilizing CO₂ as the primary carbon feedstock. Genetic and metabolic engineering are essential methodologies for enhancing PHA yield and fine-tuning the physicochemical properties of the synthesized biopolymer [6].

2.1.3 Direct extraction/modification

This pathway involves directly processing and modifying naturally occurring polymers. In this way, Polysaccharide-Based Bioplastics can be synthesized from starch, cellulose, and chitosan. For example, starch is heated and mixed with glycerol as a plasticizer. On the other hand, cellulose acetate is derived from the acetylation of wood pulp or cotton extracted cellulose. Chitosan is produced by deacetylation of chitin and is converted into a film using the solution-casting method to achieve feasible processability [7].

2.2 Emerging Innovations and Future Trends

Recent innovations aim to address bioplastics' key challenges: cost, mechanical properties (e.g., barrier and fracture strain), and heat resistance. Table 1 briefly reviews current literature. A major trend is combining bioplastics with natural fibres (wood flour, cellulose fibres, agricultural waste) or nanomaterials (cellulose nanocrystals, carbon nanotubes, nanoclays) to enhance structural integrity and lower cost.

Table 1. Current literature information [8-12].

Innovation	Synthesis Focus	Benefit
Amorphous PHAs (aPHA)	Advanced fermentation/extraction	Softer, rubbery texture, offering flexibility and toughness beyond typical crystalline PHAs.
Lignocellulosic Biopolymers	Deep Eutectic Solvents (DES) processing	Utilizing cheap, abundant wood powder, deconstructed using eco-friendly solvents to create new, strong materials.
Nanocomposites	Blending/Compounding	Incorporating nanofillers (e.g., nanoclays, nanocellulose) into PLA or starch matrices to improve mechanical strength, barrier properties, and heat resistance.
Shrilk	Solution Casting/Blending	Bioplastic made from shrimp shells (chitin) and silk protein, offering high strength, flexibility, and complete biodegradability.
Biorefinery Integration	Process Optimization	Integrating multiple processing steps (pretreatment, fermentation, purification) into a unified, energy-efficient model to enhance the transformation of waste streams into bioplastics.

3. END-OF-LIFE OF BIOPLASTICS

3.1 Biodegradation of Bioplastics

Polymer breakdown occurs in two phases: abiotic (non-living) and biotic (living). This applies to biodegradable polyesters like PLA and PHAs. Abiotic hydrolysis produces oligomers at elevated temperatures, initiating biodegradation. Microorganisms then consume these oligomers, releasing CO₂, H₂O, and CH₄ in the biotic step [13]. Degradation conditions vary across bioplastics and determine the suitability of end-of-life routes (see Table 2). For PLA and PHAs, PHAs degrade enzymatically from the start, while PLA depends on initial abiotic hydrolysis. PHAs also mineralize fully, reducing persistent microplastic risk compared to PLA in nature.

Table 2. Biodegradation of bioplastics [14-17].

Type	Disposal Route	Mechanism & Rate	Outcome
PLA	Industrial Composting	Fastest. High heat (>60°C) accelerates hydrolysis, followed by rapid microbial assimilation by thermophilic bacteria.	Meets ASTM D6400 standards.
	Ambient Soil/Water	Very Slow. Lack of sufficient heat to drive hydrolysis effectively. Degradation is largely limited to surface erosion.	Persistence for years or decades, often resulting in fragmentation (microplastics) before complete mineralization.
	Anaerobic Digestion	Biodegradable, but at a slower rate than compost.	Requires specific conditions and often a high initial moisture content.
PHAs	Soil, Water, Home Compost	Directly Biodegradable. Microorganisms (bacteria and fungi) are readily capable of secreting PHA depolymerase enzymes.	Degradation occurs across varied settings, including soil and marine environments. High rates of mass loss (e.g., 90% in soil within 200 days) are common under optimum conditions.

3.2 Recycling of Drop-in Bioplastics (Bio-PE, Bio-PET)

Drop-in bioplastics like Bio-PE and Bio-PET are designed for circularity, prioritizing recovery via existing recycling systems over biodegradation. The two main recycling methods—mechanical and chemical—are used, with growing interest in chemical processes to produce high-purity, food-grade materials.

3.2.1 Mechanical recycling

Mechanical recycling is the most mature and energy-efficient process for Bio-PE. It can also be appropriate for Bio-PET recovery. The steps for mechanical recycling are given below for Bio-PE:

1) *Collection and Sorting:* Bio-PE packaging is collected and sorted according to resin type, such as HDPE bottles or LDPE films. This is done using established methods, such as Near-Infrared (NIR) spectroscopy. Because Bio-PE is chemically identical to fossil-PE, it is mixed directly into the existing PE recycling stream.

2) *Shredding and Washing:* The plastic is shredded into flakes and thoroughly washed (often with hot water and detergents) to remove contaminants such as adhesives, labels, and food residue.

3) *Drying and Melting (Extrusion)*: Clean flakes are dried and melted in an extruder, filtered, and formed into strands.

4) *Pelletizing*: Strands are cooled and cut into recycled pellets for re-manufacturing [18].

3.2.2 Chemical recycling

Chemical recycling (also called chemolysis) breaks plastics into their original monomers. This process makes new, high-quality plastic. For Bio-PET, three chemical recycling methods are used: hydrolysis, glycolysis, and methanolysis. The reaction solvent may be water, glycol, or methanol [19].

4. COMMERCIALIZATION OF BIOPLASTICS

The commercial success of bioplastics depends on competitive cost, scalability, and reliable performance compared to petrochemical plastics. Even though the market is rapidly growing, high cost remains the main barrier to adoption. Bioplastics are generally more expensive than their conventional counterparts, with costs ranging from 20% to several hundred percent higher. This cost disparity is due to three main factors: feedstock, economies of scale, and process complexity [20].

The comparison of bioplastics and conventional plastics with respect to these parameters is presented in Table 3. The economic hurdles vary significantly depending on the material's synthesis pathway. So, commercialization challenges are investigated by bioplastic type.

Table 3. Cost Comparison: Bioplastics vs. Conventional Plastics [21].

Cost Factor	Bioplastics (PLA, PHA)	Conventional Plastics (PE, PET)
Feedstock	Derived from agricultural biomass (corn, sugar). Price is tied to commodity food markets and agricultural practices.	Derived from inexpensive fossil fuels (crude oil, natural gas).
Economies of Scale	Production capacities are relatively small, resulting in higher fixed costs per unit.	Mature industry with decades of massive scale, driving costs down.
Process Complexity	PLA: Requires multi-stage conversion (fermentation, purification, polymerization). PHA: Requires complex, sensitive microbial fermentation and costly intracellular extraction.	Direct chemical polymerization with mature, highly optimized catalytic processes.

4.1 Polylactic Acid (PLA)

Commercial Status: Most commercially mature of the biodegradable bioplastics due to the established, efficient synthesis route from lactic acid.

Challenges:

1) *Performance Gap:* Low heat resistance and brittleness require costly modifications, such as blending, to match PET and PS.

2) *Composting Infrastructure:* Limited industrial composting facilities restrict end-of-life benefits and create waste stream confusion [22].

4.2 Polyhydroxyalkanoates (PHAs)

Commercial Status: High-value niche market (e.g., medical applications, marine biodegradable products).

Challenges:

1) *High Production Cost:* Complex fermentation, solvent extraction, and strict control make PHAs expensive.

2) *Process Control:* PHAs need a narrow processing window (130-180°C) and are sensitive to heat during melting [23].

4.3 Drop-in Bioplastics (Bio-PE, Bio-PET)

Commercial Status: Strong and Growing. These benefit from drop-in compatibility with existing infrastructure. Bio-PET is a major market player.

Challenges:

1) *Feedstock Reliance:* The cost of the bio-derived monomer (e.g., bio-ethanol for Bio-PE) is often still higher than the fossil-fuel alternative.

2) *Sustainability Credibility:* Since they are not inherently biodegradable, their environmental benefit is tied only to their bio-based origin (lower carbon footprint), which can face skepticism without certified tracing of the feedstock [24].

4.4 Strategies for Cost Reduction and Commercial Viability

Future commercial success hinges on reducing the variable cost of production, primarily the feedstock and purification steps. The strategies that are present in the literature to reach this purpose are presented in Table 4.

Table 4. Cost reduction strategies for bioplastics [25-28].

Strategy	Target Material	Mechanism for Cost Reduction
Utilization of Waste Streams	PHAs (Primary focus)	Using low-cost carbon sources (e.g., food waste, sewage sludge, crude glycerol) to replace expensive pure sugars dramatically reduces feedstock cost.
Genetic/Metabolic Engineering	PLA, PHAs	Engineering microorganisms to increase yield, reduce fermentation time, and simplify the purification process (e.g., by making the polymer easier to extract).
Process Intensification	PLA, Drop-ins	Scaling production plants to achieve greater economies of scale, making production more capital efficient.
Chemical Recycling	Bio-PET	The ability to chemically recycle the polymer back to its high-purity monomer (chemolysis) increases the material's total value and reduces the reliance on virgin bio-feedstock.

5. CONCLUSION

This book chapter reviewed the current synthesis methods, end-of-life options, and commercialization challenges for key bioplastics, materials derived from renewable biomass sources.

Bioplastics are synthesized via two main routes: Chemical Synthesis, used for Polylactic Acid (PLA) and Drop-in Bioplastics (Bio-PE, Bio-PET) and Microbial Fermentation, used for Polyhydroxyalkanoates (PHAs).

The environmental benefit depends heavily on the disposal route, which is determined by the material's chemical structure. For instance, PLA requires high-temperature, high-humidity conditions for initial hydrolysis, meaning it is primarily biodegradable only in industrial composting facilities. It persists for long periods in ambient environments. On the other hand, degradation of PHAs is primarily enzymatic, allowing them to mineralize efficiently in soil, water, and home compost due to the widespread availability of PHA depolymerase enzymes. However, Drop-ins (Bio-PE, Bio-PET) are not biodegradable. Their sustainable value comes from recycling. Advanced technologies like chemolysis (chemical recycling) can break them down into their high-purity monomers for closed-loop reuse.

Cost is the main barrier to widespread bioplastic adoption. Bioplastics are generally more expensive than fossil-based plastics due to smaller economies of scale, volatility in feedstock (agricultural) prices, and the complexity of multi-stage synthesis (especially PHA extraction). So, commercial future viability relies on:

- Reducing feedstock cost by utilizing waste streams (e.g., food waste) for PHA production

- Achieving greater economies of scale for PLA and Bio-plastics.
- Using chemical/enzymatic recycling for polyesters (Bio-PET) to increase their material value and enable circularity.

In summary, bioplastics offer a pathway to reduced reliance on fossil fuels, but their environmental effectiveness is contingent upon optimized synthesis (especially cost reduction for PHAs), and the development of proper end-of-life infrastructure (industrial composting for PLA and advanced chemical recycling for drop-ins).

REFERENCES

1. Haq, F., Kiran, M., Khan, I. A., Mehmood, S., Aziz, T., & Haroon, M. (2025). Exploring the pathways to sustainability: A comprehensive review of biodegradable plastics in the circular economy. *Materials Today Sustainability*, 29, 101067.
2. Lackner, M., Mukherjee, A., & Koller, M. (2023). What are “bioplastics”? Defining renewability, biosynthesis, biodegradability, and biocompatibility. *Polymers*, 15(24), 4695.
3. Coppola, G., Gaudio, M. T., Lopresto, C. G., Calabro, V., Curcio, S., & Chakraborty, S. (2021). Bioplastic from renewable biomass: a facile solution for a greener environment. *Earth systems and environment*, 5(2), 231-251.
4. Khouri, N. G., Bahú, J. O., Blanco-Llamero, C., Severino, P., Concha, V. O., & Souto, E. B. (2024). Polylactic acid (PLA): Properties, synthesis, and biomedical applications—A review of the literature. *Journal of Molecular Structure*, 1309, 138243.
5. Siracusa, V., & Blanco, I. (2020). Bio-polyethylene (Bio-PE), Bio-polypropylene (Bio-PP) and Bio-poly (ethylene terephthalate)(Bio-PET): Recent developments in bio-based polymers analogous to petroleum-derived ones for packaging and engineering applications. *Polymers*, 12(8), 1641.
6. Koller, M. (2018). A review on established and emerging fermentation schemes for microbial production of polyhydroxyalkanoate (PHA) biopolyesters. *Fermentation*, 4(2), 30.
7. Gamage, A., Thiviya, P., Liyanapathirana, A., Wasana, M. D., Jayakodi, Y., Bandara, A., ... & Madhujith, T. (2024). Polysaccharide-Based bioplastics: Eco-Friendly and sustainable solutions for packaging. *Journal of Composites Science*, 8(10), 413.
8. Sedlacek, P., Slaninova, E., Enev, V., Koller, M., Nebesarova, J., Marova, I., ... & Obruca, S. (2019). What keeps polyhydroxyalkanoates in bacterial cells amorphous? A derivation from stress exposure experiments. *Applied Microbiology and Biotechnology*, 103(4), 1905-1917.
9. Tan, Y. T., Chua, A. S. M., & Ngoh, G. C. (2020). Deep eutectic solvent for lignocellulosic biomass fractionation and the subsequent conversion to bio-based products—A review. *Bioresource Technology*, 297, 122522.
10. Trifol, J., Plackett, D., Sillard, C., Szabo, P., Bras, J., & Daugaard, A. E. (2016). Hybrid poly (lactic acid)/nanocellulose/nanoclay composites with synergistically enhanced barrier properties and improved thermomechanical resistance. *Polymer International*, 65(8), 988-995.

11. Uğurlu, E. (2025). Utilizing mussel and shrimp shell waste for chitin and chitosan extraction: a pathway to eco-friendly bioplastics. *Biomass Conversion and Biorefinery*, 1-15.
12. De, D., Sai, M. S. N., Aniya, V., & Satyavathi, B. (2021). Strategic biorefinery platform for green valorization of agro-industrial residues: A sustainable approach towards biodegradable plastics. *Journal of Cleaner Production*, 290, 125184.
13. Tantawi, O., Joo, W., Martin, E. E., Av-Ron, S. H., Bannister, K. Y. R., Prather, K. L., ... & Plata, D. L. (2025). Designing for degradation: the importance of considering biotic and abiotic polymer degradation. *Environmental Science: Processes & Impacts*, 27(5), 1303-1316.
14. Kale, G., Auras, R., Singh, S. P., & Narayan, R. (2007). Biodegradability of polylactide bottles in real and simulated composting conditions. *Polymer testing*, 26(8), 1049-1061.
15. Liao, J., & Chen, Q. (2021). Biodegradable plastics in the air and soil environment: Low degradation rate and high microplastics formation. *Journal of hazardous materials*, 418, 126329.
16. Bracciale, M. P., De Gioannis, G., Falzarano, M., Muntoni, A., Poletti, A., Pomi, R., ... & Zonfa, T. (2023). Anaerobic biodegradation of disposable PLA-based products: Assessing the correlation with physical, chemical and microstructural properties. *Journal of Hazardous Materials*, 452, 1-12.
17. Meereboer, K. W., Misra, M., & Mohanty, A. K. (2020). Review of recent advances in the biodegradability of polyhydroxyalkanoate (PHA) bioplastics and their composites. *Green Chemistry*, 22(17), 5519-5558.
18. Alabi, O. O., Akande, T. O., Gbadeyan, O. J., & Deenadayalu, N. (2025). Advanced technologies for plastic waste recycling: examine recent developments in plastic waste recycling technologies. *RSC advances*, 15(48), 40541-40557.
19. Lamberti, F. M., Román-Ramírez, L. A., & Wood, J. (2020). Recycling of bioplastics: routes and benefits. *Journal of Polymers and the Environment*, 28(10), 2551-2571.
20. Teixeira, L. V., Bomtempo, J. V., Oroski, F. D. A., & Coutinho, P. L. D. A. (2023). The diffusion of bioplastics: What can we learn from poly (lactic acid)?. *Sustainability*, 15(6), 4699.
21. Costa, A., Encarnação, T., Tavares, R., Todo Bom, T., & Mateus, A. (2023). Bioplastics: innovation for green transition. *Polymers*, 15(3), 517.

22. Nampoothiri, K. M., Nair, N. R., & John, R. P. (2010). An overview of the recent developments in polylactide (PLA) research. *Bioresource technology*, 101(22), 8493-8501.
23. Koller, M., & Mukherjee, A. (2022). A new wave of industrialization of PHA biopolyesters. *Bioengineering*, 9(2), 74.
24. McCurdy, C., Dixon, D., Archer, E., Dooher, T., & Edwards, I. (2022). A comparison of the sealing, forming and moisture vapour transmission properties of polylactic acid (PLA), polyethene (PE) and polyethylene terephthalate (PET) coated boards for packaging applications. *Journal of Packaging Technology and Research*, 6(2), 91-100.
25. Rodriguez-Perez, S., Serrano, A., Pantión, A. A., & Alonso-Fariñas, B. (2018). Challenges of scaling-up PHA production from waste streams. A review. *Journal of environmental management*, 205, 215-230.
26. Mezzina, M. P., Manoli, M. T., Prieto, M. A., & Nikel, P. I. (2021). Engineering native and synthetic pathways in *Pseudomonas putida* for the production of tailored polyhydroxyalkanoates. *Biotechnology Journal*, 16(3), 2000165.
27. Vasilakis, P., Pyrgakis, K., Psych, M., Biundo, A., & Kokossis, A. (2023). A novel industrial biotechnology approach to valorize fatty acids to bioplastics: scope for scale-up and process efficiency using an integrated approach. In *Computer Aided Chemical Engineering* (Vol. 52, pp. 2563-2568). Elsevier.
28. Berger, N. J., & Pfeifer, C. (2025). Comparing the financial costs and carbon neutrality of polyester fibres produced from 100% bio-based PET, 100% recycled PET, or in combination. *Biomass Conversion and Biorefinery*, 15(4), 6251-6268.

CHAPTER 26

Natural Coagulants Used in Water and Wastewater Treatment and Current Studies

Hasan Kıvanç Yeşiltaş¹ & Belgin Bayat²

¹ Dr., Çukurova University Faculty of Engineering, Department of Environmental Engineering, ORCID: 0000-0003-3331-3209

² Prof. Dr., Çukurova University Faculty of Engineering, Department of Environmental Engineering, ORCID: 0000-0001-7977-0419

1. WATER AND WASTEWATER TREATMENT

Surface and groundwater, which are renewed through the hydrological cycle, have been diminishing day by day due to excessive consumption and have reached their current critical state. Today, water is supplied from resources for domestic and drinking purposes under the threat of drought driven by high temperatures and seasonal changes (Mishra & Singh, 2010; Trenberth et al., 2014). Water intended for drinking and domestic use is subjected to treatment processes suitable for its purpose and function, rendering it ready for service (Crittenden et al., 2012). However, it is required that wastewaters, particularly those originating from households and industrial activities, are not discharged into receiving environments without treatment due to their large volumes. This emerges as a significant issue on a planet suffering from water scarcity (Shannon et al., 2008; UN-Water, 2021).

The treatment of water and wastewater is carried out using physical, chemical, and biological methods. While physical methods include simple sedimentation, the treatment of water and wastewater can also be achieved through chemical precipitation with the addition of chemicals (Tchobanoglous et al., 2014). Since biological methods utilize living systems, they cannot be suitably employed for the treatment of all types of wastewater. Furthermore, when lower effluent concentrations are desired, chemical treatment methods suitable for the specific pollutant are preferred (Poyatos et al., 2010).

Among chemical treatment methods, coagulation-flocculation and precipitation are widely used in water and wastewater treatment. However, the excessive amount of waste sludge generated and the use of chemical agents for precipitation constitute the disadvantages of this system (Teh et al., 2016; Verma et al., 2012). Electrochemical methods provide a significant advantage in terms of degrading organic matter, specifically resulting in the production of carbon dioxide and water (Martinez-Huitle & Ferro, 2006; Sirés et al., 2014). However, electricity consumption is a disadvantage of the system, and their application is mostly encountered with complex organic materials and high organic loads. Adsorption and ion exchange methods are preferred in water and wastewater treatment because they do not produce sludge, can be synthesized from natural materials, and can be operated without the need for qualified personnel (Crini, 2006; Ali & Gupta, 2007).

2. TREATMENT METHODS USED IN WATER AND WASTEWATER TREATMENT

Industrial activities, carried out to sustain life and meet daily needs, cause the formation of wastes that vary depending on the nature of the process while transforming raw materials into products (Nemerow, 2007). Since the discharge of these industrial wastewaters, which contain pollutants such as heavy metals, organic matter, and solids, without treatment causes significant damage to the ecosystem, the process is strictly controlled by legal regulations (Fu and Wang, 2011). The treatment process for these wastewaters begins with the retention of coarse particles using screens and aeration. Then, particles suspended in the water that do not settle on their own are combined and weighed down with the help of chemical substances through coagulation and flocculation processes. The resulting sediments are separated by settling to the bottom in sedimentation tanks, and immediately afterwards, the clarified water is passed through sand or anthracite filters to completely remove turbidity and fine particles (Bratby, 2006).

The second stage following physical treatment includes adjusting the chemical balance of the water, ensuring its microbiological safety, and advanced treatment techniques. Depending on the wastewater characterization and the targeted quality, odor and organic pollutants can be removed by activated carbon adsorption, while dissolved salts and heavy metals can be removed using membrane technologies such as reverse osmosis (Pendergast and Hoek, 2011). In disinfection, the most critical final step of treatment, pathogenic microorganisms are neutralized using chlorine, ozone, or ultraviolet (UV) rays. Chlorination, in particular, is frequently preferred at this stage because it provides a lasting disinfection effect to prevent the water from becoming contaminated again during reuse or discharge within the facility (Sadiq and Rodriguez, 2004).

When choosing a method in industrial wastewater management, operating conditions and changes in waste load play a decisive role. Although biological treatment methods are common, fluctuations in wastewater composition due to changes in raw materials and processes in industrial production can cause significant performance losses in these systems containing living organisms (Renou et al., 2008). In addition, the need for qualified personnel and high sludge formation are among the disadvantages of biological systems. In contrast, chemical treatment systems are more resistant to variable pollutant loads and provide legal discharge standards more stably. Thus, a more manageable process is obtained that allows the water to be reused in the process or safely disposed of (Kurniawan et al., 2006).

3. COAGULATION-FLOCCULATION AND PRECIPITATION METHOD

Coagulation, flocculation, and sedimentation a chemical treatment method, is effectively used to remove pollution from wastewater by collecting colloidal substances and suspended solids and increasing their volume. The coagulation, flocculation, and sedimentation process consists of three stages. In the first stage (coagulation), a substance called a coagulant is added to the reactor to disrupt the stability of colloidal substances. In the fully mixed environment, the colloidal balance in the water is disturbed by the added coagulant, transforming it into a form that can settle along with suspended solids. In the second stage (flocculation), the clumps ready for collection are gathered more quickly in this unit where slower mixing is continued, making them suitable for pollution removal in the next stage, sedimentation. In the sedimentation unit, the final stage of the SPF process, the flocculation unit removes the clumps that have been prepared for settling by sedimentation without any mixing. The treated water is removed from the sedimentation unit via a collection structure located at the top, while the settled sludge is collected and directed to a sludge dewatering unit for disposal. In this unit, the water content of the resulting treatment sludge is reduced. However, since it will contain pollutants from the wastewater and residues of the coagulant used, it must be disposed of according to the sludge analysis performed. Before applying the coagulation, flocculation, and sedimentation process in the treatment of industrial wastewater, treatability studies are carried out in the laboratory to optimize the process conditions. As a result of the treatability studies conducted in the laboratory, the raw water is characterized, and the contact time of the wastewater in the coagulation, flocculation, and sedimentation units is determined. Technical studies such as which coagulant substance and dosage to add are also carried out. Laboratory analyses, known as jar tests, determine the contact times of the units in the coagulation, flocculation, and sedimentation process, the amount and type of coagulant, and the likely amount of sludge to be formed.

4. NATURAL COAGULANTS IN WATER AND WASTEWATER TREATMENT

The term coagulant is a general name for substances described as clotting or flocculating agents. Coagulants are used in coagulation-flocculation and sedimentation processes to remove pollutants by disrupting their stability and causing them to flocculate. Various chemical coagulants such as aluminum sulfate [Al₂(SO₄)₃], iron(III) chloride (FeCl₃), magnesium chloride (MgCl), and polyaluminum chloride (PAC) are used in the treatment of industrial wastewater. Since the appropriate pH value for each coagulant must be maintained within the reactor, the pH value must also be monitored in the treatment unit following the addition of the coagulant. While research on longer-chain chemical coagulants is currently ongoing, the use of natural materials, either directly or after pre-treatment, in the treatment of industrial wastewater is also among the important studies being conducted today (Yin, 2010). As a result of ongoing research on natural materials as coagulants, materials that would otherwise be disposed of as waste are being re-evaluated, and the formation of treatment sludge with less inorganic salts is achieved instead of chemically-based sludge formation (Oladoja, 2015). Furthermore, by using entirely biological materials instead of synthesizing coagulants from inorganic salts or materials, treatment processes will also be classified within the scope of green policies (Gautam and Saini, 2020). Summaries of various studies in the literature are presented in Table 1.

Table 1. Studies and details investigating natural coagulants in use coagulation-flocculation and sedimentation process

Material	Preparation Method	Parameter(s) Removed	Removal Efficiency (%)	Reference
Jatropha curcas (Seed & Presscake)	Seeds de-shelled, kernels ground to 63–500 µm powder; 5g powder mixed in 100mL distilled water and filtered.	Turbidity (Kaolin suspension)	>96% (Seed), >99.7% (Presscake).	Abidin et al. (2011)
Pinecone Powder (Pinus sp.)	Washed with distilled water, dried at 50°C for 8h, and ground to powder using a grinder.	COD TSS NH ³ -N Fe Ni	COD: 83.3% TSS: 99% Fe: 93.7% Ni: 86.7%.	Abujazar et al. (2022)

Date Stone Powder (<i>Phoenix dactylifera</i>)	Cleaned, dried at 50°C for 8h, and ground to a homogenous powder.	TSS NH^3-N Fe Zn COD	TSS: 99% NH^3-N 92.1% Fe: 97.7% Zn: 94.8%.	Abu Amr et al. (2023)
Eggshell Powder	Ground into powder and prepared as a 0.1 M stock solution in deionized water.	Turbidity Color TSS	Turbidity: 98.89% Color: 98.88% TSS: 98.52%.	Sibiya et al. (2021)
Dragon Fruit Foliage (<i>H. undatus</i>)	Foliage was dried and processed into a fine powder.	Turbidity COD Suspended Solids (SS)	Turbidity: 62.81% COD: 59.57% SS: 57.23%.	Ramlee et al. (2023)
Fenugreek (<i>Trigonella foenum-graecum</i>)	Seeds ground, sieved (1mm), and prepared as a 2% suspension in distilled water.	Turbidity Conductivity COD	Turbidity: 99% (Dairy) 96% (Kaolin); COD: 26%.	Calabkova et al. (2020)
Rosehip Seeds Powder (<i>Rosa canina</i>)	Washed, dried at 50°C for 8h, and ground to obtain powder form.	TSS Ni COD Fe Zn	TSS: 99% Ni: 100% COD: 86.1% Fe: 91.7%.	Abujazar et al. (2022)
Olive Seeds Powder (<i>Olea europaea</i>)	Cleaned, dried at 50°C for 8h, crushed with mortar and pestle, then ground to powder.	TSS Zn COD Fe Ni	TSS: 99% Zn: 92.6% COD: 86.3% Fe: 91.5%.	Karaağaç et al. (2022)
Moringa oleifera (Seeds)	Seeds ground to fine powder and prepared as a 5% (w/v) crude water extract, then filtered.	Turbidity Microorganisms (Bacteria)	Turbidity: 96.8% Bacterial removal: up to 99.9%.	Ndabigengesere & Narasiah (1998)

* COD: Chemical Oxygen Demands

Table 1 shows that the removal of various industrial (steel, textile, paper), municipal and synthetic (kaolin, milk) wastewater solutions has been investigated using plant-based natural materials and biological wastes (Abidin et al., 2011; Abujazar et al., 2022; Abu Amr et al., 2023; Sibiya et al., 2021; Ram-lee et al.,

2023; Calabkova et al., 2020; Ndabigengesere & Na-rasiah, 1998). Furthermore, it is observed that physical methods such as washing the materials, drying them at low temperatures such as 50°C, and mechanically grinding them into powder are more commonly preferred in the preparation of natural coagulants (Abujazar et al., 2022; Abu Amr et al., 2023; Abidin et al., 2011; Ndabigengesere & Narasiah, 1998; Karaağaç et al., 2022; Calabkova et al., 2020). In addition to the aforementioned method, it has been observed that hybridization studies with magnetic nanomaterials (ferromagnetite) using statistical optimization techniques such as Response Surface Methodology (RSM) are also employed, and synthesis can be achieved with organic-based materials such as kernels, seeds, fruit branches, or eggshells (Abujazar et al., 2023; Abu Amr et al., 2023; Sibiya et al., 2021; Ramlee et al., 2023). Furthermore, when the synthesized materials are evaluated in terms of pollutant removal performance, it is understood that Suspended Solids (TSS) and turbidity removal is quite high (reaching levels of 99%), especially when materials such as pine cones, date seeds, or olive seeds are used (Abujazar et al., 2022; Abu Amr et al., 2023; Sibiya et al., 2021; Karaağaç et al., 2022). The studies included in the table generally show that the removal efficiencies of heavy metals such as Iron (Fe), Nickel (Ni), and Zinc (Zn) are quite successful, and many studies exhibit retention capacities above 90% (Abujazar et al., 2022; Abu Amr et al., 2023; Karaağaç et al., 2022).

5. RESULTS

The depletion of water resources due to climate change and excessive consumption necessitates the treatment of high-volume domestic and industrial wastewater before it is discharged into receiving environments. Given the variable characteristics and toxic loads of wastewater, physical and chemical treatment methods are preferred over biological methods, as they yield more stable results. The coagulation-flocculation and sedimentation process, which forms the basis of these methods, relies on the destabilization, flocculation, and sedimentation of suspended particles in water using chemicals. This process typically begins with preliminary physical treatments such as screening and aeration, continues with chemical sedimentation, and is completed with advanced treatment techniques such as membrane technologies, adsorption, and chlorination/UV disinfection, depending on the need, allowing for the reuse or safe discharge of the water. Although the use of traditional chemical coagulants (aluminum sulfate, iron chloride, etc.) is widespread, the high amount of inorganic sludge and environmental disadvantages they produce have focused research on "natural coagulants." Within the framework of the green deal and sustainability principles, the aim is to use agricultural or biological wastes such as pine cones, olive pits, date pits, and eggshells as coagulants after simple processing (washing, drying, grinding). Literature studies show that these natural

materials provide over 90% efficiency in removing turbidity, suspended solids (SS), and heavy metals (Fe, Ni, Zn) from industrial wastewater. This approach both enables the recovery of waste materials and offers an environmentally friendly and low-cost alternative by reducing chemical use.

REFERENCES

- Abidin, Z. Z., Ismail, N., Yunus, R., Ahamad, I. S., & Idris, A. (2011). A preliminary study on *Jatropha curcas* as coagulant in wastewater treatment. *Environmental Technology*, 32(9), 971–977.
- Abu Amr, S. S., Abujazar, M. S. S., Karaağaç, S. U., Mahfud, R., Alazaiza, M. Y. D., & Hamad, R. J. A. (2023). Application of plant-based natural coagulant for sustainable treatment of steel and iron industrial wastewater, Karabuk, Turkey. *Desalination and Water Treatment*, 287, 39–45.
- Abu Amr, S. S., Abujazar, M. S. S., Karaağaç, S. U., Mehdi, K. K., Alazaiza, M. Y. D., Bashir, M. J. K., Fatihah, S., & Vakili, A. H. (2023). Factorial design and optimization of date stone as a natural coagulant for organic and heavy metals removal from industrial wastewater. *Global NEST Journal*.
- Abujazar, M. S. S., Karaağaç, S. U., Abu Amr, S. S., Fatihah, S., Bashir, M. J. K., Alazaiza, M. Y. D., & Ibrahim, E. (2022). The effectiveness of rosehip seeds powder as a plant-based natural coagulant for sustainable treatment of steel industries wastewater. *Desalination and Water Treatment*, 270, 44–51.
- Abujazar, M. S. S., Karaağaç, S. U., Abu Amr, S. S., Fatihah, S., Bashir, M. J. K., Alazaiza, M. Y. D., Yusof, A., & Abou Auda, M. M. (2023). Factorial design and optimization of pinecone seed powder as a natural coagulant for organic and heavy metals removal from industrial wastewater. *Desalination and Water Treatment*, 282, 107–117.
- Abujazar, M. S. S., Karaağaç, S. U., Ramadan, H., Abu Amr, S. S., & Alazaiza, M. Y. D. (2022). Application of pinecones powder as a natural coagulants for sustainable treatment of industrial wastewater. *Desalination and Water Treatment*, 269, 57–64.
- Ali, I., & Gupta, V. K. (2007). Advances in water treatment by adsorption technology. *Nature Protocols*, 2(11), 2661–2667.
- Bratby, J. (2006). *Coagulation and Flocculation in Water and Wastewater Treatment* (2nd ed.). IWA Publishing.
- Calabkova, K., Malíkova, P., Heviankova, S., & Kotalova, I. (2020). Effectiveness of alternative natural coagulants in the treatment of industrial wastewater. *IOP Conference Series: Earth and Environmental Science*, 444, 012007.
- Crini, G. (2006). Non-conventional low-cost adsorbents for dye removal: a review. *Bioresource Technology*, 97(9), 1061–1085.
- Crittenden, J. C., Trussell, R. R., Hand, D. W., Howe, K. J., & Tchobanoglous, G. (2012). *MWH's Water Treatment: Principles and Design*. John Wiley & Sons.

- Fu, F., & Wang, Q. (2011). Removal of heavy metal ions from wastewaters: A review. *Journal of Environmental Management*, 92(3), 407–418.
- Gautam, S., & Saini, G. (2020). Use of natural coagulants for industrial wastewater treatment. *Global Challenges*, 4(9), 1900099.
- Karaağaç, S. U., Abujazar, M. S. S., Kopan, M., Abu Amr, S. S., & Alazaiza, M. Y. D. (2022). The potential use of olive seeds powder as plant-based natural coagulant for sustainable treatment of industrial wastewater. *Desalination and Water Treatment*, 277, 169–176.
- Kurniawan, T. A., Chan, G. Y., Lo, W. H., & Babel, S. (2006). Physico–chemical treatment techniques for wastewater laden with heavy metals. *Chemical Engineering Journal*, 118(1–2), 83–98.
- Martinez-Huitle, C. A., & Ferro, S. (2006). Electrochemical oxidation of organic pollutants for the wastewater treatment: direct and indirect processes. *Chemical Society Reviews*, 35(12), 1324–1340.
- Mishra, A. K., & Singh, V. P. (2010). A review of drought concepts. *Journal of Hydrology*, 391(1–2), 202–216.
- Ndabigengesere, A., & Narasiah, K. S. (1998). Use of Moringa oleifera seeds as a primary coagulant in wastewater treatment. *Environmental Technology*, 19(8), 789–800.
- Nemerow, N. L. (2007). *Industrial Waste Treatment: Contemporary Practice and Vision for the Future*. Elsevier.
- Oladoja, N. A. (2015). Headway on natural polymeric coagulants in water and wastewater treatment operations. *Journal of Water Process Engineering*, 6, 174–192.
- Pendergast, M. M., & Hoek, E. M. (2011). A review of water treatment membrane nanotechnologies. *Energy & Environmental Science*, 4(6), 1946–1971.
- Poyatos, J. M., Muñio, M. M., Almecija, M. C., Torres, J. C., Hontoria, E., & Osorio, F. (2010). Advanced oxidation processes for wastewater treatment: state of the art. *Water, Air, & Soil Pollution*, 205, 187–204.
- Ramlee, A. A., Som, A. M., Puasa, S. W., & Hamid, H. A. (2023). Coagulation–flocculation mechanism and characterisation of *Hylocereus undatus* foliage as a natural coagulant in industrial wastewater treatment. *Chemical Papers*, 77, 6083–6093.
- Renou, S., Givaudan, J. G., Poulain, S., Dirassouyan, F., & Moulin, P. (2008). Landfill leachate treatment: Review and opportunity. *Journal of Hazardous Materials*, 150(3), 468–493.

- Sadiq, R., & Rodriguez, M. J. (2004). Disinfection by-products (DBPs) in drinking water and predictive models for their occurrence: a review. *Science of the Total Environment*, 321(1–3), 21–46.
- Shannon, M. A., Bohn, P. W., Elimelech, M., Georgiadis, J. G., Marinas, B. J., & Mayes, A. M. (2008). Science and technology for water purification in the coming decades. *Nature*, 452(7185), 301–310.
- Sibiya, N. P., Rathilal, S., & Tetteh, E. K. (2021). Coagulation Treatment of Wastewater: Kinetics and Natural Coagulant Evaluation. *Molecules*, 26(3), 698.
- Sirés, I., Brillas, E., Oturan, M. A., Rodrigo, M. A., & Panizza, M. (2014). Electrochemical advanced oxidation processes: today and tomorrow. A review. *Environmental Science and Pollution Research*, 21, 8336–8367.
- Tchobanoglous, G., Stensel, H. D., Tsuchihashi, R., & Burton, F. (2014). *Wastewater Engineering: Treatment and Resource Recovery*. McGraw-Hill Education.
- Teh, C. Y., Budiman, P. M., Shak, K. P. Y., & Wu, T. Y. (2016). Recent advancement of coagulation–flocculation and its application in wastewater treatment. *Industrial & Engineering Chemistry Research*, 55(16), 4363–4389.
- Trenberth, K. E., Dai, A., van der Schrier, G., Jones, P. D., Barichivich, J., Briffa, K. R., & Sheffield, J. (2014). Global warming and changes in drought. *Nature Climate Change*, 4(1), 17–22.
- Verma, A. K., Dash, R. R., & Bhunia, P. (2012). A review on chemical coagulation/flocculation technologies for removal of color from textile wastewaters. *Journal of Environmental Management*, 93(1), 154–168.
- Yin, C. Y. (2010). Emerging usage of plant-based coagulants for water and wastewater treatment. *Process Biochemistry*, 45(10), 1437–1444.

CHAPTER 27

Biomimetric Designs From A Textile Engineering Perspective

Züleyha Değirmenci¹ & Ebru Çoruh²

¹ Gaziantep University, Department of Textile Engineering, Gaziantep, Turkey

² Gaziantep University, Department of Textile and Fashion Design, Gaziantep, Turkey

1. INTRODUCTION

1.1. What is Biomimicry?

Biomimetry, in its simplest form, means inspiration from nature. We can see that the tools we use, many industrial designs, and even the ecosystem itself are inspired by nature. We see around us that a bird is modeled after the design of an airplane wing or that a sea violet's sensor system serves as a model for a light sensor. Analyzing nature and the components of an ongoing system can be groundbreaking for almost any industry. This study will examine the concept of "biomimetry" from a textile engineering perspective, and will bring together academic studies in this field to draw original conclusions.

1.1.1. Fundamental Principles of Biomimicry

Undoubtedly, the most important principle is that nature is the best model when designing a new product. Other principles can be summarized as follows:

- Nature runs on solar energy. (Therefore, using renewable energy in human designs is best for both humans and nature.)
- Nature consumes only what it needs. (This sentence summarizes the concept of sustainability.)
- Nature evolves according to functionality. (A sunflower turning toward the sun to benefit more from the sun is an example.)
- Nature transforms everything. (Plant and animal waste serving as fertilizer for the soil falls into this category.)
- Everything in nature is collaborative. (Plants and animals living in the same land is an example.)
- Nature always economizes to avoid depletion. (Just as desert animals thicken their skin to survive in dry times.) [1]

The principle in nature is both cooperation and competition, just as humans creating a common market and then competing is an example.

1.1.2. Principles of Design and Nature

The fundamental principle in product design can be broadly classified into two categories: artistic and industrial. While the fundamental goal of both designs is satisfaction for humanity, artistic design prioritizes visual harmony, evoking emotions, and presenting distinct features that distinguish it from existing products. In industrial design, more concrete concepts such as functionality, ergonomics, cost, and lifespan are more crucial. Therefore, in artistic design, the designer is the primary element, while in industrial design, the user is the primary

element. A common quest for both designs is the source of inspiration. Can a person design something they cannot see? Or, in other words, do humans only see when awake?

From a neurological or physiological perspective, seeing is defined as "perceiving or selecting the presence of something with the help of the eye." [2]. Therefore, while it is scientifically true that everything viewed objectively through the eye is seen in the same way, it is impossible to interpret the things seen in the same way. The meaning derived from what is seen varies for many reasons, including readiness level, experience, emotional state, age, and gender. Therefore, seeing is not merely a physiological phenomenon. The meanings we ascribe to what we see change by comparing it to what we have seen before. Therefore, dreams, both awake and asleep, are a synthesis of what we see. In his study, Taşgüzen Z. (2010) attempted to analyze the relationship between dreams and design. One of the inferences here relates to surrealism. If we summarize the points that provide insights into design from the Surrealist approach to dreams, the first of these is the striking, original, and inventive qualities of their works. The second implication of Surrealism is the production of a written text about dreams (the Surrealist Manifesto) and the "unifying" and "connective" quality this text provides in a field like art, where individual production is predominant. Another emphasis of Surrealism lies in the importance it places on imaginative experience. As G. Bachelard emphasizes, imaginative experience offers the opportunity to "perceive the world with a different understanding." [3].

In an industrial sense, product design is based on more concrete data. Planning is primarily based on the intended use. Plans must be implemented, tests must be completed, and if the controls are ineffective, precautions must be taken to create new plans (PDCA). At this point, design takes place in the planning phase. While product development efforts based on a previously developed product are quite important, the invention of a product that hasn't been considered before is even more crucial. From this perspective, it can be said that efforts to integrate a product used for a different purpose into a different field are also related to the designer's imagination.

Ultimately, both designers need a model. Living things in nature, with their structures, appearances, or working methods, can always be considered a source of inspiration for humanity. With bioinspiration (inspiration from nature), before the idea is conceived, a specific function performed by a biological system is observed as a source of inspiration, and then an attempt is made to create an imitative system and an innovative design that performs the same function for technological applications [4].

2. Biomimicry in the Textile Industry: Basic Concepts

In biomimicry studies, to gain a deep understanding of organisms and mimic what has been learned, it is necessary to approach the process from three different levels. The first level is the imitation of the natural form. The second, which requires a deeper understanding, is the imitation of natural processes, and the third level is the imitation of the natural ecosystem [4].

2.1. Mimicry of Natural Form

The appearance of some living things we see in nature can be directly transferred to products, both in terms of design and functionality. Figures 1 and 2 are examples of this.



Figure 1. Clothing and fabric design inspired by the red bird (Designed with artificial intelligence)

The red bird in Figure 1 can be used as a model for textile design based on both the combination of colors and the softness of its feathers.



Figure 2. Helicopter Design Inspired by a Dragonfly (Designed with Artificial Intelligence)

The images in Figure 2 clearly demonstrate that the dragonfly's wing system can form the basis for a helicopter design.

The reversible design, seen in both figures, is possible. The key is that the helicopter was designed with dragonflies as inspiration. However, due to the separation from the natural world and the influence of industrial and urban life, it's possible that people today identify helicopters more easily. In this case, they

might compare the insect they see to a helicopter. Similarly, "dragonfly" can also appear as "helicopter insect" in some sources [5].

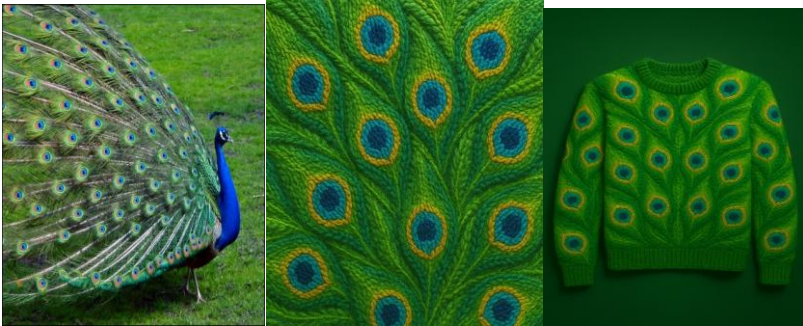


Figure 3. Knitted fabric design and garment application inspired by the peacock (designed using artificial intelligence)

2.2. Level of Imitation of Natural Processes

Natural processes have also guided production processes. Synthetic yarn production, based on the stages of silk production and spider web production methods, is a prime example.

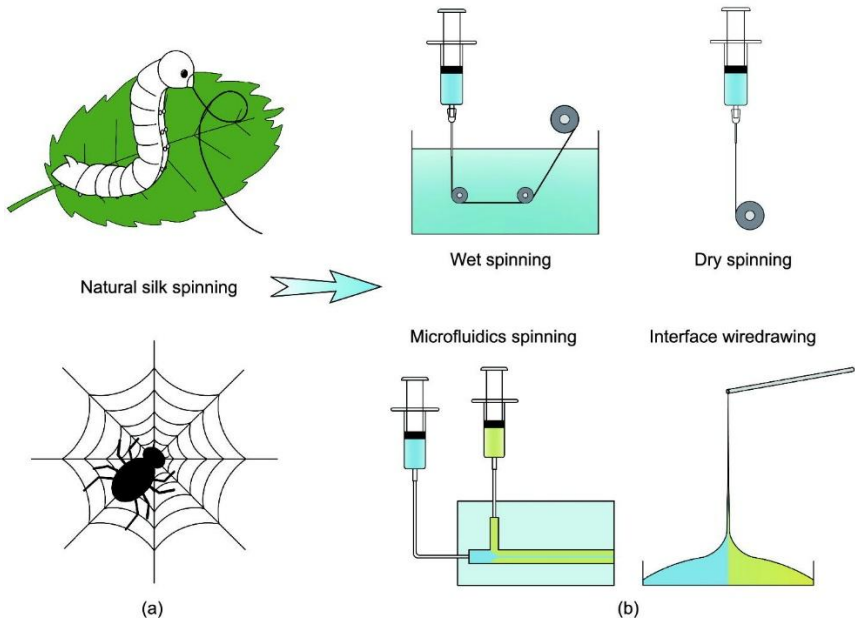


Figure 4. a) Natural silk spinning b) Synthetic fiber production using wet and dry spinning techniques [6]

Using nature as inspiration for design can be easier from an artistic perspective. However, stages such as analyzing functions to mimic the process, considering or inventing appropriate methods, and testing the suitability of the proposed methodology for production are more complex. An interdisciplinary approach is crucial for this type of production. Silk salvia, which is undoubtedly a discovery for someone involved in zoology, has gained meaning from the perspective of natural textile producers. With proper process analysis, silk cultivation is a rather laborious production process consisting of many processes. Transferring this process to synthetic fiber production is possible by analyzing the expected performance of the product, identifying areas for improvement, and identifying existing drawbacks. In this case, inventing the production method can be considered the most challenging step. Revising existing techniques can be much easier.

Ants are highly organized insect societies with specific divisions of labor for systematic operation. There are essentially two types of ants in an ant nest: the queen ant and the worker ants. While the queen ant lays eggs, worker ants either forage for food or defend the nest as soldier ants. Hormones determine which ant plays which role in an ant community. Juvenile hormone is one of these hormones, and ants have a mechanism that determines the activity of juvenile hormone [7]. Workers and managers in a factory can be evaluated in the same way. While the manager comes up with new business ideas or customers, the workers carry out production. The distinguishing factor here is the skills acquired through training or experience.

2.3. Mimicking the Natural Ecosystem

For the textile industry, "biomimetics" or biomimicry allows for the human-made redesign of mechanisms found in nature for performance and sustainability. Natural selection, which has persisted for millions of years, has always served as a role model for humanity. This process, discovered by the systematic production of thread by the silkworm, is now prominently featured in many innovative studies. Nature is undoubtedly a laboratory of high-performance materials. This laboratory contains endless information for humanity about the most efficient production process with the least resources. Some examples from the textile perspective can be listed as follows:

1.Spider silk: This fiber boasts superior properties in terms of lightness, flexibility, and durability. Spider silk, a natural alternative to synthetic fibers, has a strength ranging from 1 to 4 GPa, a value comparable to steel. However, spider silk clearly has a significant advantage in terms of lightness. The material's ability

to stretch around 35% means the fiber also has a high impact absorption capacity. This high energy absorption capacity is also a result of the fiber's high resistance to breakage. This value is reported in the literature as having a toughness rate of around 160 MJ m^{-3} . The strength of the protein-based (spidroin-based) fiber is due to highly organized crystalline regions composed of polyalanine repeats; Its flexibility is provided by glycine-rich amorphous regions.

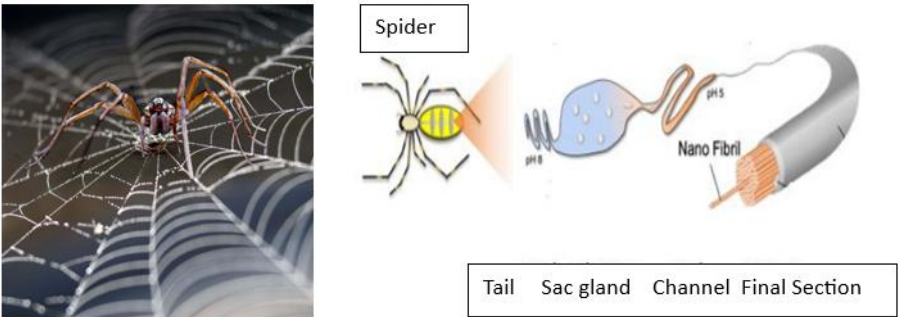


Figure 5. Transforming spider silk into fiber [8]

Spider silk fiber has a unique molecular structure and excellent properties. After being converted into textiles, it is much more comfortable to wear. It is soft, breathable, dense, wrinkle-resistant, non-yellowing, and resistant to ultraviolet rays. With the continuous advancement of science and technology, this superior property of spider silk fiber is constantly being researched and developed by the textile industry and applied in the fashion field. For example, British textile designers Simon Peers and Nicholas Godley teamed up to employ more than 70 workers to design a small, hand-cranked machine for extracting spider silk. They searched for the golden spider in Madagascar and then collected and extracted spider silk from over 1 million golden spiders. After five years of research and exploration, the team collected 1.2 million strands of spider silk and ultimately woven a magnificent golden spider silk measuring 335 cm long and approximately 122 cm wide. In conclusion, spider silk is a natural textile material with superior properties. With the urgent need for environmental protection in modern society and the widespread adoption of environmental protection concepts, spider silk fiber will be continuously extracted and expanded in various fields under the encouragement of science and technology [9].

2. Lotus plant: The surface of this plant's leaf is quite smooth. Because no material adheres to the fiber, dirt formation is not observed. The term "lotus effect" is used in the literature instead of self-cleaning. The leaf's surface is referred to as "superhydrophobic." This phenomenon is attributed to the inability of water droplets to adhere to the fiber's surface and the resulting roll-off of dirt particles. Inspiring nanotextile studies, a level of smoothness minimizes adhesion and allows all types of particles on the surface to be removed by gravity.

The self-cleaning abilities of surfaces have attracted significant attention due to their potential applications in environmental sustainability, healthcare, and advanced materials. These effects are driven primarily by two mechanisms based on surface properties: hydrophobicity (water repellency) and hydrophilicity (water attraction): Hydrophobic surfaces exhibit high water contact angles ($\theta > 90^\circ$), which allow water droplets to bead up and roll off, transporting dirt and contaminants. Superhydrophobicity, with contact angles exceeding 150° , is achieved with hierarchical surface structures and low-surface energy materials such as fluorinated silanes or polytetrafluoroethylene (PTFE). The "lotus effect," inspired by lotus leaves, is an example of this mechanism, where micro- and nanoscale roughness traps air and minimizes water-surface contact. This principle is used in self-cleaning glass, which reduces maintenance costs by up to 40%, and in engineered textiles that increase durability and stain resistance. This provides a groundbreaking example for the engineering applications of biomimicry. [10].

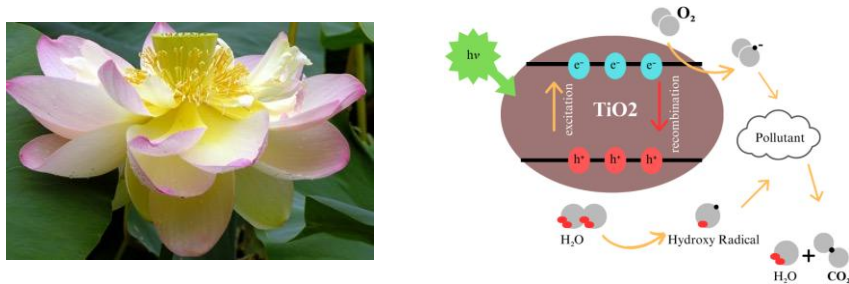


Figure 6. a) Lotus plant [11] b) Self-cleaning mechanism inspired by the lotus plant [10]

3. Polar bear: The fur-like hairs covering the bodies of polar bears absorb solar radiation and transmit it to the skin, generating heat in the region. This mechanism has become a role model for thermal insulation studies. Such a composite product design can be particularly effective for providing both flexible and transparent thermal insulation. Yue et al. (2020) developed a structure that mimics the polar bear's fur, which allows it to cope with the extreme cold of the

Arctic. Polar bears absorb heat from the sun through their fur and reflect the heat generated in their bodies back through their fur, maintaining body temperature. For this purpose, they created a porous Ag/cellulose/carbon nanotube (CNT) structure. One side of the cellulose base membrane was coated with CNTs using a foam finish, and the other side was coated with Ag using a magnetron scattering method. The formed Ag/cellulose/CNT laminate (ACCNL) shows high absorbance on the CNT side under the entire solar spectrum and high reflectivity on the Ag side in the FIR region under human body radiation. In addition to this passive heating property, since Ag is highly conductive, it can also gain an additional function with an external source. The mechanism is given in Figure 7 [12, 13].

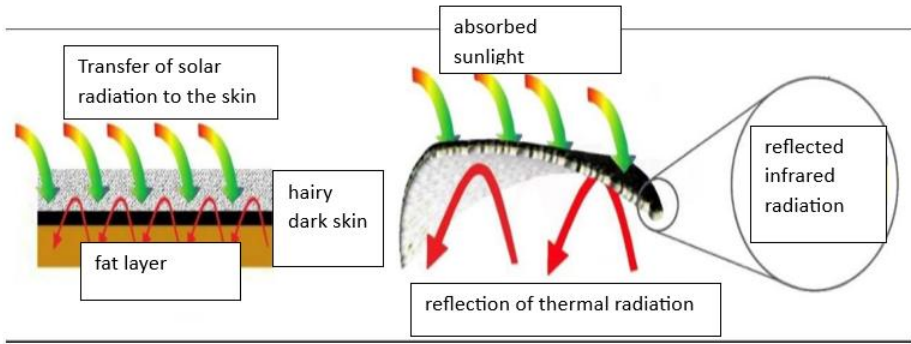


Figure 7. Example design of the polar bear's insulation mechanism in textile materials [13]

4. Peacock and Morpho butterfly: The colors of these animals are dazzling. These colors are not based on pigments. Nanostructures on the surface of the feathers interact with light when exposed to light, turning them blue-green. Considering the environmental damage caused by dyeing textiles, the importance of imitating such a mechanism for sustainability becomes clearer. Another example of this ecological mechanism is the Morpho butterfly. The Morpho butterfly's vibrant, luminous blue colors are due to microscopic scales on the back of its wings that reflect light. The underside of the Morpho's wings is a dull brown with numerous eyespots, which provide camouflage against predators such as birds and insects when its wings are closed. When the Morpho flies, the contrasting bright blue and dull brown colors glow, giving the impression that the Morpho is appearing and disappearing [14]. This is called natural coloring by reflection of light.



Figure 8 a) Blue peacock [15] b) Blue morpho butterfly [16]

5. Pine cone: Pine cone scales have a hygroscopic mechanism that can open and close in response to air humidity. Imitating this mechanism allows the development of breathable textile materials that can adapt to ambient temperature. The design of thermal comfort clothing for textiles has been the subject of much research. Conventional fibers swell when they absorb moisture, reducing the number of pores in their textile structures. The new-generation pine cone-effect fibers and yarns developed by MMT Textile Ltd. increase the number of pores when they absorb moisture. In dry conditions, their structures open like a pine cone, reducing air permeation and enhancing insulation properties.

In a study where this mechanism was applied to a textile prototype, the face was made of tightly woven polyester fabric, while the back was coated with a non-porous hydrophilic polyether-ester block copolymer membrane known as Sympatex™ (Figure 9). The two materials are bonded together using an industrial lamination process that uses hydrophilic polymeric adhesives. Such laminates allow moisture vapor to pass through the textile but prevent external water droplets from penetrating the clothing system. Polyester fibers swell slightly when exposed to moisture, while the polyether component swells significantly. "U"-shaped die cuts positioned along the plane of the double-layer fabric open (Figure 9 b) and close (Figure 9 a) as humidity increases and decreases, thus making the layer porous [17].

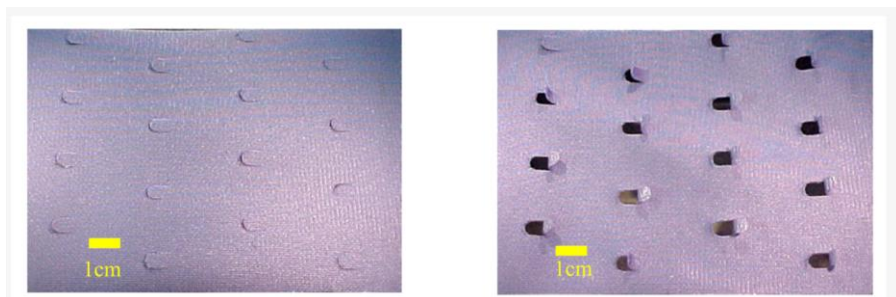


Figure 9. Photograph of the original pinecone-effect textile in dry (a) and moist (b) conditions. The sample size is 10×5 cm. The pores open when the hygroscopic polyether layer swells upon exposure to moisture vapor.

A commercial company (Nike) developed this double-layer textile structure as a knitted fabric using a machine with a double needle set to produce two parallel fabric layers joined by interlocking stitches. The active textile system was applied in terms of yarn fiber composition. The front side of the textile (top, outward-facing layer) used yarns containing hydrophobic polyester fibers. The second layer, or back side of the textile, consisted of yarns containing hydrophilic polyamide fibers. The polyamide fibers expanded when wet [18]. The resulting system was applied to a tennis dress. Sequins resembling U-shaped wings in a pine cone-effect textile were placed on the back of the dress; the resulting technology was marketed under the trade name Macro React (Figure 10).



Figure 10. Detail of the active double-fabric system of Nike Micro React sportswear in humid conditions. Source: Getty Images

However, mimicking these engineering marvels of nature necessitates the development of complex and precise technological platforms capable of reproducing the molecular blueprints of biology on an industrial scale.

3. From Biology to Factory: Production Technologies of Biomimetic Textiles

Commercializing biological designs and producing them on an industrial scale is not easy. In this process, calculations and production parameters must be established to mass-produce designs produced at a microscale in the laboratory by mimicking nature.

3.1 Biopolymer Synthesis: Recombinant DNA Technology

Spider silk is undoubtedly the best example of this technology. This material is a high-performance protein-based fiber, but its industrial-scale production is quite difficult and costly. In this case, a more controllable production process is created by first synthesizing the protein and then transferring it to another suitable organism. This allows for production on a larger scale. This technology is called recombinant DNA. The process can be summarized simply as gene cloning followed by host organism selection. In this case, the correct selection of the host system (bacterial, yeast, plant, or animal cultures) can affect the lifespan, cost, and yield of the synthesized protein. For example, bacteria (*E. coli*) are fast-growing, low-cost, and simple media, but their lack of secretion ability is a disadvantage. Yeast (*P. pastoris*) has high culture density but low yield. Plants are suitable for large-scale production and have a low cost. However, the risk of genetic variability and a long production cycle are disadvantages. Mammalian cell cultures are actually the closest to human protein, but the cost is high. Considering transgenic animals, large-scale production is possible, but ethical debates continue.

3.2. Purification and Fiber Production

To convert a raw protein produced by host organisms into fiber, the protein must first be purified. The most important factor here is to separate components such as host cell DNA, lipids, and other unwanted proteins. This is the most important factor affecting performance. Filtration, centrifugation, and chromatography are standard methods used in industrial biotechnology. They are a critical cost and quality control point because they directly affect the fiber's mechanical integrity and biocompatibility. Process efficiency is also a key step in product commercialization. Various spinning technologies are used to obtain continuous fibers from purified and concentrated protein solutions. These technologies mimic the natural spinning process, allowing protein molecules to achieve a regular structure. Wet spinning and electrospinning are examples of this. In wet spinning, the critical factor is the chemicals in the coagulation bath, while in electrospinning, the voltage level in the electric field. Fibers obtained

with this method generally form non-woven, porous networks with high surface area.

4.0 Market Applications and Commercial Potential

Commercially, biomimetic technologies have begun to leave the theoretical and experimental stages behind and transform into tangible, value-added products in the textile market. Inspired by nature's millions of years of evolutionary optimization, these innovations push the performance limits of traditional materials while also addressing critical market demands such as sustainability and functionality. This section analyzes the key applications of biomimetic textiles in various sectors and the commercial advantages these applications offer.

4.1 High-Performance and Protective Textiles

Spider silk's exceptional toughness (energy absorption) makes it an ideal candidate for next-generation anti-ballistic equipment for the military and defense industries. It holds the potential to develop lighter and more flexible ballistic vests compared to traditional protective vests. The biocompatibility and biodegradability of these fibers are revolutionary in the medical field. Applications such as dissolvable surgical sutures, high-strength artificial ligaments, and tissue engineering scaffolds that serve as a scaffold for cell growth both increase patient comfort and accelerate healing processes. The high strength, lightness, and environmentally friendly solutions offered by technical textiles used in industrial settings highlight the fibers' potential.

4.2 Functional Surface Technologies

Surface modifications inspired by nature are giving textile materials new and valuable functions they never had before. An example of this is textiles based on the self-cleaning properties of the lotus plant. Its intended use is to provide solutions for difficult-to-clean areas such as roofs, awnings, and shower curtains. The velcro strips used in apparel (Figure 11 b) were inspired by the millions of micro-hairs on the gecko's foot structure (Figure 11 a).

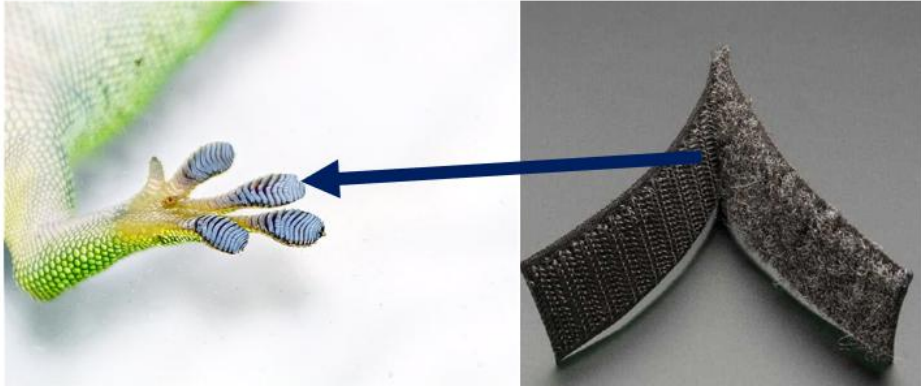


Figure 11. a) Lizard's sticky hand [<https://www.flickr.com/photos/vanshy/5746363829>]
b) Adhesive tape <https://www.robotistan.com/iletken-velkro-bant-cirt-cirt>

4.1 Smart and Sustainable Clothing

For clothing to be considered smart, it must also possess functionality beyond fashion. Nature is a good teacher in this regard. Pine cones open and close based on humidity levels, and textile systems developed based on this principle increase air permeability in response to the wearer's perspiration. This principle has inspired the development of adaptive sportswear, which incorporates microstructures that open when the wearer's body temperature and humidity increase, thus providing active ventilation and supporting performance. Fiber technologies, inspired by the wing structure of the Morpho butterfly, offer a radical solution to the environmental problems of the textile dyeing industry. Experimental studies have shown that artificial furs designed using biomimetic principles can offer better thermal insulation, higher water vapor permeability (breathability), and a warmer feel compared to natural furs.

4.4 Sustainable Composites and Lightweight Structures

Composites are generally known for their lightness and strength. Inspired by the hollow, graded, and fiber-reinforced structures of plants such as horsetail (*Equisetum hyemale*) and giant spear (*Arundo donax*), this concept aims to achieve maximum structural efficiency with minimal material. This production method is both environmentally friendly and sustainable.

5. Future Perspective and Strategic Implications

Biomimetics, which has become increasingly popular in recent years, offers the potential to produce value-added products for the future. Analyzing problems in nature followed by observing emerging solutions plays a key role in the production of innovative functional textile products. However, even with accurate analysis of problems and solutions, integrating this concept into textile production is not easy. While this requires extensive R&D expertise, production conditions must also be suitable for this purpose.

5.1 Emerging Trends and Expectations

Analyzing the sector reveals potential products that could be developed, including self-healing materials (biological structures that can repair themselves when damaged, such as the vine (*Aristolochia macrophylla*)). This could be a particularly important innovation in terms of sustainability, as it prevents products from becoming unusable after even the slightest wear. Examples of biodegradable synthetics include biopolymers such as polylactic acid (PLA), produced from corn starch, and polyhydroxyalkanoates (PHA), synthesized by genetically modified microorganisms. These products offer sustainable and circular alternatives to petroleum-based synthetics. By integrating "smart" functions (sensors, data transmission, adaptive properties) into the fabric's fiber structure, they will have the ability to dynamically adapt to the environment and the user, like a living organism. This approach will integrate wearable technology into daily life without compromising aesthetics or requiring additional hardware.

5.2 Challenges and Opportunities

While there are significant obstacles to the widespread adoption of biomimetic textiles, each of these obstacles also brings with it strategic opportunities. Cost, the technology's rarity, the impossibility of standardization, low durability, and the lack of interdisciplinary understanding can be summarized as obstacles to the development of biomimetic. However, this technique is an important technique for producing value-added products in specific areas. This method has the potential to produce products where functionality, rather than cost, is particularly important for specialized fields such as the defense industry.

5.3 Conclusion

Biomimetry can be defined as drawing inspiration from or imitating nature. In the textile industry, the first step is to analyze methods for solving problems in nature and then to develop solution-oriented textile products by matching these problems with the textile industry. Accurately understanding user expectations is

crucial for designing innovative products. Collaboration between disciplines in this area will facilitate easier solutions. Biologists' knowledge of the structural characteristics of living things in nature can only be translated into production methods through an engineer's perspective. Just as it's impossible for a civil engineer to understand the techniques an ant uses to build its nest, it's impossible to transform spider secretions into fiber without a polymer engineer. When it comes to design, mastering the details of design methods is as crucial as being a keen observer. Proper selection of production methods, and the commercial viability of products are crucial.

7.REFERENCES

1. www.Biolearn.Eu
2. <https://Tr.Wiktionary.Org/Wiki/G%C3%B6rmek>
3. Taşgüzen, Zamzam. Dream And Design: An Inquiry Into A Fundamental Formation. 2010.
4. Degerli Geyik, N. "Innovative Design Approach In Textiles: Biomimicry." *İdil Journal* 68 (2020): 676-680.
5. Aydın, L., & Selçuk, Ö. Less Important Harmful Arthropods (Arthropods) Found In Honeybees Chapter 1: Insects.
6. Xu, Z., Wu, M., Ye, Q., Chen, D., Liu, K., & Bai, H. (2022). Spinning From Nature: Engineered Preparation And Application Of High-Performance Bio-Based Fibers. *Engineering*, 14, 100-112.
7. <https://www.Bilimup.Com/Bilim-İnsanlari-Karincalarin-Hiyerarsik-Duzeyini-Degistirmayı-Basardi>
8. <https://www.Biomedya.Com/Orumcek-İpegi>
9. Wu, X., & Chen, L. (2022, April). New Textile Material—Spider Silk Fiber. In *Journal Of Physics: Conference Series* (Vol. 2256, No. 1, P. 012003). Iop Publishing
10. Ghosh, J., Rupanty, N. S., Noor, T., Asif, T. R., Islam, T., & Reukov, V. (2025). Functional Coatings For Textiles: Advancements In Flame Resistance, Antimicrobial Defense, And Self-Cleaning Performance. *Rsc Advances*, 15(14), 10984-11022
11. Lotus
12. Yu, W. L., Wang, Z., And Jin, L. (2012). The Experimental Study On Infrared Radiation Spectrum Of The Human Body. *Proceedings - Ieee-Embs International Conference On Biomedical And Health Informatics: Global Grand Challenge Of Health Informatics, Bhi 2012*, 25(Bhi), 752–754. <https://Doi.Org/10.1109/Bhi.2012.6211692>,
13. Fidan, E. (2024). Development Of Far Infrared Reflective Pigment Coated Fabric Structures
14. <https://www.Rainforest-Alliance.Org/Species/Blue-Morpho-Butterfly/>
15. https://tr.wikipedia.Org/Wiki/Tavus_Kusu
16. <https://www.Rainforest-Alliance.Org/Species/Blue-Morpho-Butterfly/>
17. Kapsali, V., & Vincent, J. (2020). From A Pinecone To The Design Of An Active Textile. *Biomimetics*, 5(4), 52.

18. Baron, M. R., Carter, K. E., & Dua, B. (2010). Base. Patent No. 7,754,626.
Washington, Dc: U.S. Patent And Trademark Office

CHAPTER 28

Calcium-Based Borohydride Systems for Hydrogen Storage: Mechanochemical Synthesis and Catalytic Dehydrogenation

Ahmet Fırat Karabulut³

³ Department of Chemistry, Engineering and Natural Sciences Faculty, Osmaniye Korkut Ata University, Türkiye

1. Introduction

Hydrogen is widely recognized as a promising energy carrier for future sustainable energy systems due to its high gravimetric energy density and environmentally benign combustion products. However, the practical deployment of hydrogen technologies is strongly dependent on the development of safe, efficient, and compact hydrogen storage methods. Among the various storage strategies, solid-state hydrogen storage using complex metal hydrides has attracted significant interest because of its advantages in safety, volumetric density, and reversibility compared to compressed or liquefied hydrogen systems [1–3].

Within this class of materials, boron-based hydrides—particularly metal borohydrides—stand out due to their exceptionally high hydrogen contents. Calcium borohydride ($\text{Ca}(\text{BH}_4)_2$) is of particular interest, as it combines a high theoretical hydrogen capacity (~11.5 wt%) with relatively low material cost and abundance of calcium-containing precursors [4,5]. Despite these advantages, $\text{Ca}(\text{BH}_4)_2$ suffers from several critical limitations, including high dehydrogenation temperatures, slow hydrogen release kinetics, and limited reversibility after hydrogen desorption [6]. These challenges have motivated extensive research efforts aimed at developing alternative synthesis routes and advanced catalytic strategies.

In recent years, mechanochemical processing has emerged as a powerful and sustainable approach for the synthesis of complex hydrides. By utilizing mechanical energy instead of high temperatures or solvents, mechanochemical methods offer unique advantages in terms of energy efficiency, scalability, and environmental compatibility.

2. Mechanochemical Synthesis of Borohydrides

Mechanochemical synthesis involves the initiation of chemical reactions through mechanical energy, typically introduced via high-energy ball milling. During milling, repeated fracturing, cold welding, and plastic deformation generate fresh reactive surfaces, lattice defects, and enhanced atomic diffusion pathways, enabling solid-state reactions that are otherwise inaccessible under conventional thermal conditions [7–9].

Early mechanochemical studies demonstrated that complex hydrides could be synthesized through solid-state metathesis reactions driven by intensive mechanical activation [10]. Compared to solution-based or melt-based synthesis routes, mechanochemical methods eliminate the need for solvents, reduce processing temperatures, and significantly shorten reaction times. As a result, mechanochemistry has become an important tool in the synthesis of hydrogen storage materials.

2.1. Mechanochemical Parameters and Reaction Pathways

The efficiency of mechanochemical synthesis is governed by a complex interplay of milling parameters, including milling speed, milling time, ball-to-powder ratio (BPR), vial and ball materials, and milling atmosphere. These parameters directly influence phase evolution, reaction kinetics, and product purity. High-energy milling promotes defect formation and localized transient heating at impact sites, which facilitate diffusion-controlled reactions at ambient temperatures [11–13].

For calcium borohydride systems derived from borate precursors, optimization of milling conditions is particularly critical. Insufficient milling energy may result in incomplete conversion, whereas excessive milling can induce undesirable side reactions or partial decomposition of reaction intermediates. Studies have shown that optimized BPR values, typically in the range of 10:1 to 20:1, combined with controlled milling durations, enable efficient solid-state metathesis without compromising phase purity [14].

Reaction pathways in mechanochemical synthesis differ fundamentally from those observed in conventional thermal processes. Rather than being governed by equilibrium thermodynamics, mechanochemical reactions proceed through defect-mediated activation and localized non-equilibrium conditions. This allows direct synthesis of complex hydrides from mineral-derived precursors under ambient conditions, significantly reducing energy input and processing complexity [15].

2.2. Practical Advantages, Limitations, and Scale-Up Considerations

Mechanochemical synthesis is attractive for borohydride production because it can combine mixing, activation, and reaction in a single solid-state step, while avoiding large solvent volumes and salt-containing waste streams typical of solution metathesis routes. In addition to lower process complexity, mechanochemical routes allow the direct use of mineral-derived precursors, supporting cost-effective sourcing and potential regional supply chains for calcium- and boron-rich feedstocks [7–9,15,16].

In practical implementation, three limitations are repeatedly highlighted in the literature. First, borohydrides and activated borate intermediates can be highly sensitive to moisture and oxygen, requiring controlled atmospheres during powder handling, milling, and sampling. Second, local temperature spikes at impact sites and overall heat build-up in the vial may lead to side reactions or partial decomposition if not managed via intermittent milling, external cooling, or optimized charge size. Third, scale-up is not a simple linear extrapolation: milling energy density, ball dynamics, and mixing efficiency change with reactor

volume, and continuous or semi-continuous milling concepts may be needed for industrial throughput [10–12,14].

From an engineering perspective, reporting and controlling the milling atmosphere (Ar/N₂), sealing quality, and heat dissipation strategy are essential not only for product quality but also for safety. In a review context, synthesis studies can be compared most meaningfully when they provide complete milling parameter sets (rpm, time, BPR, vial/ball material) along with post-synthesis handling steps and characterization protocols [11–13].

2.3. Key Characterization Metrics for Mechanochemically Prepared Hydrides

For review and comparison purposes, mechanochemical borohydride studies are commonly benchmarked using: (i) phase purity and crystallite size (XRD with Rietveld refinement when feasible), (ii) hydrogen release onset temperature, peak temperature(s), and total released hydrogen (TGA/DSC coupled with MS or volumetric methods), (iii) apparent activation energy (e.g., Kissinger or isoconversional analyses), and (iv) cycling stability under repeated dehydrogenation/regeneration conditions [6,19,20].

Because mechanochemical processing can introduce both beneficial nanostructuring and unintentional contamination, complementary analyses are useful: SEM/EDS for morphology and elemental mapping, BET for surface area (particularly in composite/catalyst systems), and IR/Raman for monitoring B–H related vibrational signatures when applicable. Clear reporting of sample mass, heating rate, carrier gas, and baseline corrections is important for fair comparison of hydrogen release profiles across different laboratories [6,15,19].

Table 1. Recommended reporting parameters for mechanochemical borohydride synthesis studies.

Category	Parameter	Typical reporting	Rationale
Milling	Speed / time / BPR	rpm, hours, ball-to-powder ratio	Defines energy input and reaction extent
Environment	Atmosphere / sealing	Ar/N ₂ , glovebox transfer, leak check	Prevents hydrolysis/oxidation; impacts safety
Materials	Vial/ball material	Steel, WC, ZrO ₂	May introduce contamination or catalytic effects
Outputs	Hydrogen metrics	onset/peak T, wt% H ₂ , kinetics model	Enables comparison across studies

3. Mechanochemical Synthesis of Calcium Borohydride from Colemanite-Derived Precursors

A major advancement in sustainable borohydride synthesis has been achieved through the mechanochemical production of $\text{Ca}(\text{BH}_4)_2$ using calcium-containing borate minerals, such as synthetic colemanite, as starting materials. This approach eliminates the reliance on expensive and moisture-sensitive calcium halides traditionally used in chemical synthesis routes [16].

In this method, high-energy ball milling activates the borate structure and promotes solid-state reactions between calcium-containing precursors and reducing agents. The resulting mechanochemical reduction pathway enables direct formation of $\text{Ca}(\text{BH}_4)_2$ under ambient conditions. Importantly, this strategy allows the utilization of abundant and low-cost boron minerals, offering a promising route toward scalable and environmentally friendly production of calcium borohydride.

Initial investigations into these mechanochemical pathways were reported through conference-level studies, which provided early insights into reaction feasibility and mechanistic trends [17]. Subsequent peer-reviewed journal publications expanded upon these findings by presenting detailed structural characterization, thermal analysis, and hydrogen release behavior [16,18]. Together, these studies demonstrate that mechanochemical synthesis from colemanite-derived precursors is a viable alternative to conventional chemical routes.

3.1. Conceptual Reaction Scheme and Process Flow

A simplified process flow for colemanite-derived $\text{Ca}(\text{BH}_4)_2$ synthesis can be described as: (i) preparation of a calcium borate precursor (e.g., synthetic colemanite or calcium metaborate), (ii) intimate mixing with a suitable reducing/borohydride-forming agent under inert atmosphere, (iii) high-energy milling to activate the borate network and enable solid-state metathesis, and (iv) post-milling conditioning (controlled heating, evacuation, or solvent-free separation) to remove volatile byproducts and stabilize the borohydride phase [16,18].

In the reviewed studies, mechanochemical activation plays a dual role: it increases the reactivity of the borate lattice by generating defects and it improves reagent contact by reducing particle size and promoting microstructural refinement. These effects can shorten reaction times and reduce temperature requirements, which is especially beneficial when mineral-derived feedstocks are employed. However, controlling side-phase formation (e.g., residual borates, oxides, or borides) remains essential because these phases can influence hydrogen release pathways and regeneration feasibility [6,19,20].

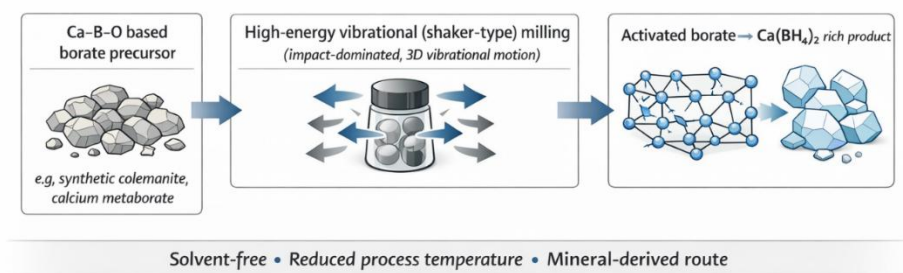


Figure 1. Schematic process flow for Ca–B–O based borate precursors via mechanochemical activation, with colemanite-derived systems representing a key example, highlighting key inputs (precursor, inert atmosphere, milling parameters) and outputs (borohydride phase, byproducts). Schematic redrawn by the author based on Refs. [16–18].

3.2. Summary of Representative Outcomes and Practical Benchmarks

To facilitate comparison across studies, practical benchmarks include the achievable borohydride fraction (phase purity), the onset temperature for hydrogen release, and the cumulative hydrogen desorbed under standardized conditions. In addition, reproducibility and sensitivity to handling are important for mineral-derived routes. In a review setting, these benchmarks can be summarized in a compact table that connects synthesis conditions with performance metrics and dominant byproduct phases [16,18,19,27].

Table 2. Example benchmarking descriptors for calcium borohydride synthesis and dehydrogenation studies (illustrative).

Study type	Route additive	Key conditions	Dehydrogenation trend	Notes
Mechanochemical synthesis	Colemanite-derived route	Inert milling; optimized BPR/time	Onset typically >350 C (uncatalyzed)	Mineral feedstock; purity depends on milling control
Catalysis	Metal fluoride halide additives	Low wt% catalyst; dispersion critical	Reduced onset; faster kinetics	Catalyst choice and distribution dominate
Hybrid architecture	Carbon-supported or hydrogel-based	High surface area support / matrix	Further kinetic enhancement	Improved contact; potential cycling benefits

4. Dehydrogenation Behavior of Calcium Borohydride

Despite its successful synthesis, the practical application of $\text{Ca}(\text{BH}_4)_2$ as a hydrogen storage material is limited by its unfavorable dehydrogenation behavior. Thermal decomposition of $\text{Ca}(\text{BH}_4)_2$ typically occurs above 350 °C and involves multiple reaction steps, leading to the formation of stable boron-containing phases (e.g., borides and borates), depending on reaction conditions [6,19].

4.1. Thermodynamic and Kinetic Considerations

From a thermodynamic perspective, the strong B–H bonds in $\text{Ca}(\text{BH}_4)_2$ favor hydrogen retention, while kinetically, hydrogen diffusion and phase transformation processes are severely constrained in the solid state. As a result, the mismatch between high theoretical hydrogen capacity and poor hydrogen release performance represents a major bottleneck for practical applications [20].

Mechanochemical processing partially alleviates these limitations by producing nanostructured materials with high defect densities and reduced diffusion lengths. However, milling alone is insufficient to fully overcome the kinetic barriers associated with hydrogen desorption. This limitation has driven the development of catalytic strategies aimed at lowering dehydrogenation temperatures and enhancing hydrogen release kinetics.

4.2. Reaction Pathways, Byproducts, and Implications for Regeneration

Dehydrogenation of $\text{Ca}(\text{BH}_4)_2$ is often described as a multi-step process involving intermediate phases and irreversible side products, depending on temperature, atmosphere, and additives. Formation of stable calcium borate phases (e.g., calcium metaborate) is frequently implicated in loss of reversibility. For a review chapter, it is useful to distinguish (i) primary hydrogen release steps from (ii) secondary transformations that lock boron into thermodynamically stable frameworks. The latter processes are the main cause of capacity fade and explain why simple rehydrogenation is generally ineffective for calcium borohydride systems [6,19,20].

Catalysts can alter these pathways by promoting lower-temperature routes, modifying intermediate stability, or enabling alternative product distributions that are more amenable to rehydrogenation. Consequently, dehydrogenation studies should be interpreted together with post-reaction phase analysis (XRD/SEM-EDS) to identify dominant byproducts and to assess whether regeneration is plausible. This synthesis-decomposition linkage is central for designing closed-loop calcium borohydride cycles [21-27].

5. Catalytic Enhancement of Dehydrogenation

Catalytic modification has proven to be an effective strategy for improving the hydrogen release behavior of calcium borohydride-based systems. A wide range of catalysts, including transition metal fluorides, chlorides, and oxides, have been investigated for this purpose [21–23].

5.1. Transition Metal-Based Catalysts

Transition metal additives influence dehydrogenation pathways by destabilizing B–H bonds, promoting intermediate phase formation, and enhancing hydrogen diffusion. Among these, cobalt- and titanium-based compounds have shown particularly strong catalytic effects, often reducing dehydrogenation temperatures by 50–100 °C compared to pristine $\text{Ca}(\text{BH}_4)_2$ [24].

5.2. Carbon-Supported and Hydrogel-Based Catalytic Systems

Carbon-supported catalysts introduce additional benefits by providing high surface area and facilitating uniform dispersion of active species. Activated carbon-supported metal fluorides, for example, exhibit synergistic effects that improve hydrogen release kinetics and cycling stability [25].

More recently, hydrogel-based catalytic architectures have emerged as a novel platform for borohydride dehydrogenation. Hydrogels provide a flexible polymeric matrix that enables homogeneous distribution of catalytic sites while maintaining intimate contact with the hydride phase. This structural flexibility enhances hydrogen diffusion, suppresses particle agglomeration, and improves overall catalytic efficiency [26].

5.3. Mechanistic Considerations and Catalyst Dispersion

Although catalytic effects are often reported as shifts in onset temperature and faster desorption rates, the underlying mechanisms can differ substantially between additive classes. Metal halides/fluorides may participate in transient species formation, alter the stability of borohydride intermediates, or act as nucleation sites for new phases that reduce the apparent activation barrier. In supported systems, dispersion and interfacial contact frequently dominate performance; the same nominal catalyst loading can yield different results depending on how uniformly the active phase is distributed over the borohydride particles and support.

For hydrogel-based systems, the polymeric matrix can improve the spatial distribution of active centers and maintain intimate contact with the borohydride particles during gas evolution. This may suppress agglomeration and preserve reactive surface area, which is especially relevant under repeated heating cycles. A practical recommendation emerging from the literature is that catalyst studies

should report the dispersion strategy (co-milling, impregnation, or in situ formation), support characteristics (surface area/pore structure), and post-dehydrogenation phase analysis to clarify the dominant catalytic pathway [21-26].

5.4. Practical Comparison of Catalytic Strategies

Table 3 summarizes the main catalytic strategy families discussed in this chapter and highlights typical advantages and limitations when applied to $\text{Ca}(\text{BH}_4)_2$. In the context of a mineral-derived borohydride route, the most attractive solutions are those that (i) reduce onset temperature without producing strongly irreversible byproducts and (ii) can be implemented without introducing complex purification steps. Hybrid approaches (e.g., co-milling with a supported catalyst) are particularly compatible with mechanochemical processing.

Table 3. Qualitative comparison of catalytic strategies for $\text{Ca}(\text{BH}_4)_2$ dehydrogenation.

Strategy	Typical benefit	Typical limitation	Notes for implementation
Transition metal halides/fluorides	Lower onset temperature; faster kinetics	May form stable byproducts; sensitivity to dispersion	Prefer low loadings; report phase analysis post-run
Carbon-supported catalysts	Enhanced dispersion; improved mass transport	Support variability; possible adsorption effects	Report support area/pore structure; control moisture uptake
Hydrogel-based architectures	Uniform active-site distribution; reduced agglomeration	Matrix stability under heat; reproducibility	Describe hydrogel chemistry, loading protocol, and stability

Transition metal additives influence the dehydrogenation behavior of calcium borohydride primarily by destabilizing B–H bonds, promoting the formation of reactive intermediates, and facilitating hydrogen diffusion through the solid matrix. Unlike simple particle size reduction effects induced by milling, transition metal species can actively modify reaction pathways by altering intermediate stability and lowering the apparent activation energy for hydrogen release.

Among the investigated systems, titanium- and cobalt-based compounds have consistently demonstrated pronounced catalytic effects, often leading to significant reductions in hydrogen release onset temperatures and accelerated desorption kinetics. These effects are commonly attributed to the formation of transient metal–boron or metal–hydrogen interactions, which weaken B–H bonds and promote alternative, lower-energy decomposition routes. In addition, finely

dispersed metal species can act as heterogeneous nucleation sites for intermediate phases, further enhancing reaction rates.

The effectiveness of transition metal catalysts is strongly dependent on dispersion and interfacial contact with the borohydride phase. Mechanochemical processing offers a particularly effective strategy for achieving homogeneous catalyst distribution, as co-milling enables intimate contact between catalyst particles and $\text{Ca}(\text{BH}_4)_2$ while simultaneously introducing lattice defects that synergistically enhance catalytic activity. However, excessive catalyst loading or poor dispersion can lead to the formation of stable secondary phases that do not contribute to hydrogen release and may even hinder reversibility.

From a practical perspective, catalytic loadings are typically kept low to balance kinetic enhancement against phase stability and material cost. Consequently, comparative evaluation of transition metal-based catalysts should consider not only reductions in onset temperature and improvements in hydrogen release rates, but also the nature of post-dehydrogenation phases and the feasibility of subsequent regeneration steps.

6. Recovery and Re-Synthesis of Calcium Borohydride

One of the most critical challenges associated with $\text{Ca}(\text{BH}_4)_2$ systems is the limited reversibility after hydrogen release. Dehydrogenation typically results in stable calcium borate phases that are difficult to reconvert into borohydrides. Nevertheless, recent studies have demonstrated that mechanochemical re-synthesis of $\text{Ca}(\text{BH}_4)_2$ from calcium metaborate is feasible, enabling partial recovery of hydrogen storage capacity [27].

This closed-loop approach, which combines hydrogen release and mechanochemical regeneration, represents a significant step toward sustainable hydrogen storage cycles based on calcium borohydride systems.

6.1. Regeneration Pathways and Closed-Loop Concepts

Regeneration of borohydrides after hydrogen release is a central hurdle for practical storage. For calcium borohydride systems, irreversible conversion to stable borides/borates can prevent direct rehydrogenation under mild conditions. Mechanochemical regeneration concepts aim to bypass high-pressure hydrogenation by reconvertng stable calcium borate phases (e.g., calcium metaborate) to borohydride-containing products through solid-state activation with appropriate reagents. The literature indicates that partial recovery is feasible, but the achievable yield depends strongly on phase purity and the byproduct spectrum after dehydrogenation [27].

In a closed-loop perspective, regeneration should be evaluated not only by the recovered hydrogen capacity but also by reagent availability, process safety, and

waste minimization. Therefore, future work should emphasize process-integrated regeneration strategies that leverage mineral feedstocks, reduce reagent intensity, and enable repeated cycling with minimal performance drift.

6.2. Practical Constraints: Safety, Handling, and Reproducibility

Borohydrides are typically moisture sensitive and can generate hydrogen rapidly upon contact with water; accordingly, safe handling under inert atmosphere and robust containment are essential for both laboratory and scale-up workflows. Mechanochemical processing introduces additional hazards (pressure build-up, heat generation), so documentation of vial sealing, gas release management, and post-milling conditioning is important. From a review standpoint, reproducibility is improved when studies report both the intended composition and the measured post-milling composition (including trace impurities), because small differences can strongly affect dehydrogenation kinetics and apparent reversibility.

7. Outlook and Future Perspectives

The integration of mechanochemical synthesis with advanced catalytic architectures offers a promising strategy for addressing the long-standing challenges associated with calcium borohydride-based hydrogen storage materials. Mineral-derived synthesis routes, particularly those utilizing colemanite-based precursors, provide a sustainable and cost-effective pathway for large-scale production. At the same time, catalytic strategies involving carbon-supported and hydrogel-based systems enable substantial improvements in hydrogen release behavior.

Future research efforts should focus on optimizing milling–catalyst synergy, understanding phase evolution during cycling, and developing scalable processing technologies. Such advances will be essential for translating calcium borohydride systems from laboratory-scale materials into practical hydrogen storage solutions.

7.1. Conclusions

Calcium-based borohydride systems remain compelling for hydrogen storage because of their high theoretical hydrogen content and the prospect of mineral-derived synthesis routes. Mechanochemical processing offers a practical pathway to activate borate precursors and produce borohydride phases with reduced process complexity. However, kinetic limitations during dehydrogenation necessitate catalytic and composite strategies, and long-term viability depends on regeneration concepts that minimize irreversible byproduct formation. Consistent reporting of milling parameters, handling protocols, and hydrogen measurement conditions will accelerate cross-study comparisons and guide scale-up.

References

- [1] Schlapbach L, Züttel A. Hydrogen-storage materials for mobile applications. *Nature* 414:353–358 (2001).
- [2] Grochala W, Edwards PP. Thermal decomposition of the non-interstitial hydrides for the storage and production of hydrogen. *Chem Rev* 104:1283–1315 (2004).
- [3] Orimo S, Nakamori Y, Eliseo JR, Züttel A, Jensen CM. Complex hydrides for hydrogen storage. *Chem Rev* 107:4111–4132 (2007).
- [4] Jensen CM, Gross KJ. Development of catalyzed borohydrides for hydrogen storage. *J Alloys Compd* 446–447:663–667 (2007).
- [5] Nakamori Y, Miwa K, Ninomiya A, et al. Correlation between thermodynamical stabilities of metal borohydrides and their structural properties. *Phys Rev B* 74:045126 (2006).
- [6] Rude LH, Filsø U, D’Anna V, et al. Tailoring properties of borohydrides for hydrogen storage by anion substitution. *J Phys Chem C* 115:7768–7777 (2011).
- [7] Baláz P. Mechanochemistry in nanoscience and minerals engineering. *Chem Soc Rev* 42:7571–7637 (2013).
- [8] James SL, Adams CJ, Bolm C, et al. Mechanochemistry: opportunities for new and cleaner synthesis. *Chem Soc Rev* 41:413–447 (2012).
- [9] Boldyrev VV. Mechanochemistry and mechanical activation of solids. *Russ Chem Rev* 75:177–189 (2006).
- [10] Huot J, Liang G, Schulz R. Mechanically alloyed metal hydride systems. *J Alloys Compd* 293–295:495–500 (1999).
- [11] Varin RA, Zbroniec L. Hydrogen desorption kinetics in mechanically milled hydrides. *J Alloys Compd* 504:S478–S482 (2010).
- [12] Boynuegri TA, Karabulut AF, Gürü M. Synthesis of Borohydride and Catalytic Dehydrogenation by Hydrogel Based Catalyst. *J Electron Mater* 45(8):3949–3956 (2016). <https://doi.org/10.1007/s11664-016-4620-4>
- [13] Takacs L. Self-sustaining reactions induced by ball milling. *Prog Mater Sci* 47:355–414 (2002).
- [14] Suryanarayana C. Mechanical alloying and milling. *Prog Mater Sci* 46:1–184 (2001).
- [15] Milanese C, Garroni S, Gennari F, et al. Hydrogen storage properties of borohydrides. *Int J Hydrogen Energy* 39:16521–16530 (2014).

- [16] Karabulut AF, Gürü M, Boynueğri TA, Aydın MY. Synthesis of $\text{Ca}(\text{BH}_4)_2$ from Synthetic Colemanite Used in Hydrogen Storage by Mechanochemical Reaction. *J Electron Mater* 45(8):3957–3963 (2016). <https://doi.org/10.1007/s11664-016-4550-1>.
- [17] Karabulut AF, Gürü M, Akkaş Boynueğri T, Aydın MY. Synthesis of $\text{CaBH}_4(2)$ from synthetic colemanite by mechanochemical reaction. Proceedings of the III European Conference on Renewable Energy Systems (ECRES), Antalya, Turkey (2015).
- [18] Ar İ, Usta Güler Ö, Gürü M. Synthesis and characterization of sodium borohydride and a novel catalyst for its dehydrogenation. *Int J Hydrogen Energy* 43(44):20214–20233 (2018). <https://doi.org/10.1016/j.ijhydene.2018.06.184>.
- [19] Dornheim M. Thermodynamics and kinetics of complex hydrides. *J Alloys Compd* 404–406:150–153 (2008).
- [20] Züttel A. Materials for hydrogen storage. *Mater Today* 6:24–33 (2003).
- [21] Barkhordarian G, Klassen T, Bormann R. Catalytic mechanism of transition-metal compounds on MgH_2 . *J Phys Chem B* 108:11070–11076 (2004).
- [22] Vajo JJ. Influence of nanostructuring on hydrogen storage properties. *Curr Opin Solid State Mater Sci* 15:52–61 (2011).
- [23] Liang G, Huot J, Boily S, Van Neste A, Schulz R. Hydrogen storage properties of mechanically milled MgH_2 . *J Alloys Compd* 292:247–252 (1999).
- [24] Bogdanović B, Schwickardi M. Ti-doped alkali metal aluminium hydrides as potential novel reversible hydrogen storage materials. *J Alloys Compd* 253–254:1–9 (1997).
- [25] Çakanyıldırım Ç, Gürü M. Decomposition of NaBH_4 with self-regeneration of carbon-supported CoCl_2 catalyst. *Int J Green Energy* 14(12):1005–1010 (2017). <https://doi.org/10.1080/15435075.2017.1354297>.
- [26] Boynueğri TA, Gürü M. Catalytic dehydrogenation of calcium borohydride by using hydrogel catalyst. *Int J Hydrogen Energy* 42(28):17869–17873 (2017). <https://doi.org/10.1016/j.ijhydene.2017.02.208>.
- [27] Karabulut AF, Gürü M, et al. Recovery of calcium metaborate and re-synthesis of calcium borohydride by mechanochemical reaction. Proceedings of the European Conference on Renewable Energy Systems (ECRES), Sarajevo, Bosnia and Herzegovina (2017).

CHAPTER 29

Particleboard and Composite Materials from Agricultural Wastes

Ahmet Fırat Karabulut¹

¹ Department of Chemistry, Engineering and Natural Sciences Faculty, Osmaniye Korkut Ata University, Türkiye

1. Introduction

Particleboard is one of the most widely used engineered wood products due to its low cost, ease of manufacturing, and suitability for interior applications. However, the rapid growth of furniture and building sectors has resulted in increasing pressure on forest resources and rising environmental concerns [1,2]. Conventional particleboard production relies heavily on virgin wood fibers, contributing to deforestation and increased carbon emissions.

Agricultural wastes constitute an abundant and renewable source of lignocellulosic materials. Residues such as nut shells, fruit stones, crop husks, and food-processing wastes are generated in large quantities worldwide and are often underutilized or disposed of through environmentally unfavorable practices such as open burning or landfilling [3–5]. Many of these wastes possess chemical compositions comparable to wood, containing cellulose, hemicellulose, and lignin, making them suitable for particleboard and composite production [6,7].

Figure 1 schematically illustrates the main classes of agricultural wastes used in particleboard production, including nut shells, fruit stones, crop residues, and food-processing wastes. This classification highlights the diversity of agro-waste resources available for composite panel manufacturing.

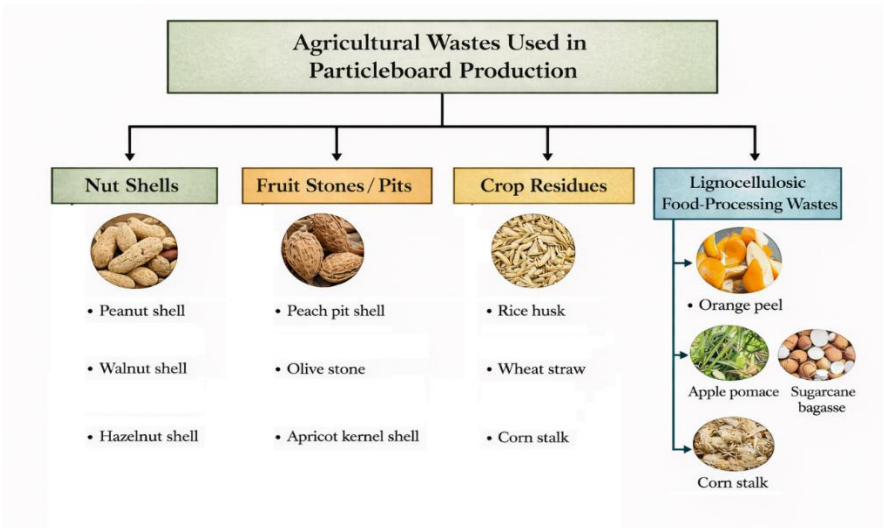


Figure 1. Classification of agricultural wastes directly used in particleboard production, including nut shells, fruit stones, crop residues, and lignocellulosic food-processing wastes

2. Raw Materials and Processing Routes

Agricultural wastes used for particleboard production vary significantly in terms of particle size, density, ash content, and chemical composition. Nut shells and fruit stones (e.g., peanut shell and peach pit shell) are generally denser and harder, whereas crop residues such as rice husk exhibit high ash and silica content [8,9].

Typical processing routes include drying, grinding, particle size classification, binder mixing, mat formation, and hot pressing. Processing parameters such as pressing temperature (100–180 °C), pressure (2–10 MPa), and pressing time strongly affect panel density, interfacial bonding, and final mechanical performance.

Figure 2 presents a schematic representation of the general manufacturing route for agro-waste-based particleboards.



Figure 2. Schematic illustration of particleboard manufacturing from agricultural wastes

At the end of this section, typical processing parameters together with the corresponding test standards are summarized in Table 1.

Table 1. Typical processing parameters used for agro-waste-based particleboards

Parameter	Typical range
Pressing temperature (°C)	100–180
Pressing pressure (MPa)	2–10
Pressing time (min)	5–15
Target board density (kg m ⁻³)	600–900
Binder type	UF, PF, epoxy, polymeric binders
Binder content (wt%)	8–20
Particle size range (mm)	0.25–4
Mat moisture content (%)	6–12

Mechanical and physical properties discussed in this chapter were evaluated according to relevant standards (EN 310 for bending strength, EN 317 for water absorption, EN 312 for product specifications, and ISO 4589 for limiting oxygen index).

3. Peanut Shell–Based Particleboards

Peanut shell (*Arachis hypogaea*) is a porous, low-density agricultural residue generated in large quantities during food processing. Due to its limited intrinsic strength, peanut shell has mainly been investigated for applications where fire resistance and moisture repellency are more critical than structural performance.

Non-flammable and water-repellent particleboards manufactured from peanut shell using polymer-modified binder systems and mineral additives have been reported previously [14]. The developed panels exhibited a bending strength of approximately 2.1 MPa, which is lower than that of conventional wood-based particleboards but sufficient for non-load-bearing interior applications. Fire performance was significantly improved, with limiting oxygen index (LOI) values reaching 30 %, indicating self-extinguishing behavior. Water absorption was reduced from 37 % to 21 %, demonstrating improved dimensional stability.

4. Peach Pit Shell–Based Particleboards Reinforced with Glass Powder

Peach pit shell (*Prunus persica*) is a dense agricultural waste characterized by high hardness and thermal stability. Şahin et al. [15] investigated particleboards produced from waste peach pit shell reinforced with glass powder. Under optimized processing conditions, bending strength values of 34–34.5 MPa were achieved, representing one of the highest mechanical performances reported for agro-waste-based particleboards.

Glass powder incorporation significantly enhanced fire resistance, yielding LOI values up to 44.5 %. Water absorption values ranged between 0.15 and 0.32 g/g, depending on molding pressure. These results clearly demonstrate the potential of peach pit shell–glass powder composites for high-performance and fire-sensitive interior applications.

5. Palm Kernel–Based Hybrid Composites Reinforced with Mint Fiber

Palm kernel waste is generated in large quantities during palm oil production. Taşdemir et al. [16] reported palm kernel-based composites reinforced with mint fiber. Mechanical strength values of approximately 6.2 MPa were obtained, while LOI values reached 50 %, indicating exceptional flame retardancy. The enhanced performance was attributed to improved stress transfer and char formation induced by fibrous reinforcement.

6. Orange Peel–Based Composite Particleboards

Orange peel is a food-processing waste rich in cellulose, hemicellulose, and pectin. Taşdemir et al. [17] reported bending strength values up to 15 MPa in orange peel-based particleboards. Under optimized polymer/orange peel ratios, LOI values as high as 55 % were reported. Water absorption decreased from 0.35 to 0.16 g/g with increasing binder content, emphasizing the importance of matrix selection for moisture-sensitive agro-waste systems.

7. Fireproof and High-Temperature Durable Particleboards from Rice Husk

Rice husk is distinguished by its high silica content, which imparts superior fire resistance and thermal stability. Gürü et al. [18] reported rice husk particleboards with bending strength values of 11.4 MPa, LOI values around 40 %, and a fire point of 493 K. Silica-rich char structures formed in situ throughout the composite matrix during combustion act as effective barriers against heat transfer.

Figure 3 schematically illustrates the in-situ, bulk fire-retardant mechanism associated with silica-rich rice husk particleboards. It should be emphasized that the fire-retardant behavior does not originate from a surface coating or a distinct

protective layer, but from an intrinsic bulk mechanism governed by the homogeneous distribution of silica-rich rice husk particles within the composite.

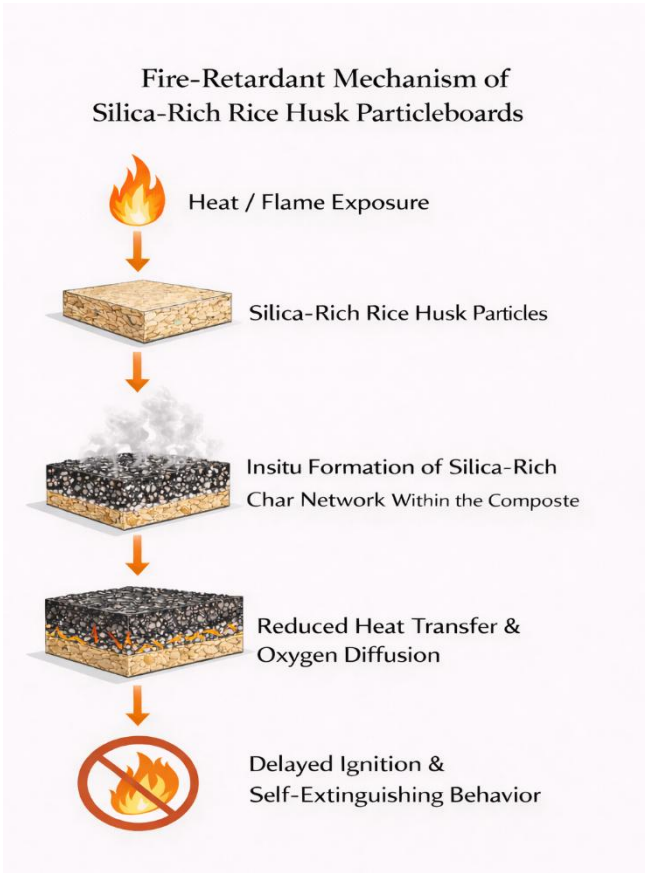


Figure 3. Schematic illustration of the *in-situ* fire-retardant mechanism in silica-rich rice husk particleboards, where silica-containing particles are homogeneously dispersed within the composite matrix, leading to bulk char network formation rather than a discrete surface layer under thermal exposure.

8. Comparative Discussion

A comparative evaluation of agro-waste-based particleboards reveals that performance strongly depends on raw material type and composite design. Peach pit shell–glass powder systems provide the highest mechanical strength (34–34.5 MPa) combined with excellent fire resistance (LOI 44.5 %). Palm kernel–mint fiber composites exhibit outstanding flame retardancy (LOI 50 %) with moderate mechanical strength, whereas rice husk particleboards offer a balanced combination of durability and fire resistance (LOI 40 %).

These trends are summarized in Table 2, which compares the mechanical and fire performance of different agro-waste-based particleboards.

Table 2. Comparative mechanical, fire, and water resistance performance of agro-waste-based particleboards

Agro-waste	Binder reinforcement	/ Bending strength (MPa)	LOI (%)	Water absorption (%)	Ref.
Peanut shell	Polymer + mineral filler	2.1	30	21	[14]
Peach pit shell	Glass powder	34.5	44.5	15	[15]
Palm kernel	Mint fiber	6.2	50	12	[16]
Orange peel	Polymeric binder	15	55	16	[17]
Rice husk	UF binder	11.4	40	12	[18]

Water absorption values originally reported as g g^{-1} were converted to percentage for consistency. “–” indicates data not reported in the original studies.

Figure 4 provides a comparative visualization of mechanical strength and fire resistance for the investigated agro-waste systems.

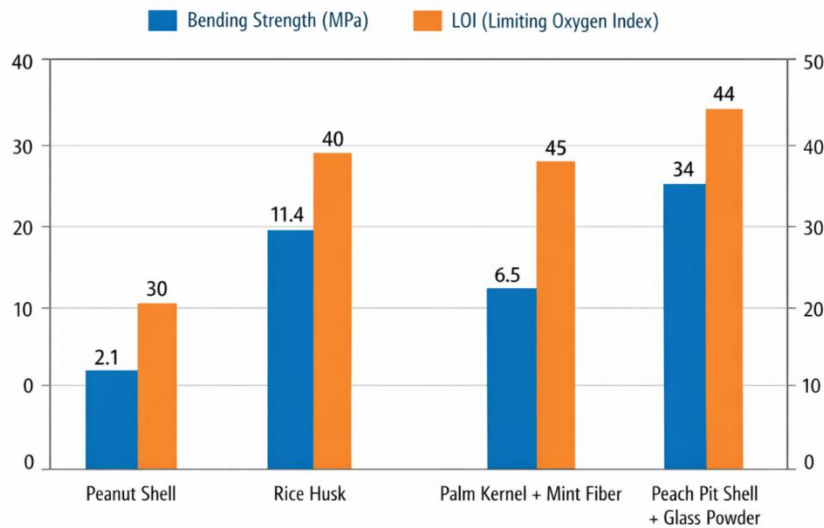


Figure 4. Comparative mechanical strength and fire resistance (LOI) of recommended agro-waste-based particleboard systems.

9. Conclusions and Future Perspectives

Agricultural wastes can be effectively converted into functional particleboard materials with competitive mechanical and fire performance. By employing suitable binder systems, mineral additives, and fibrous reinforcements, agro-waste-based particleboards can serve as sustainable alternatives to conventional wood-based panels. Future research should focus on hybrid composite designs, long-term durability assessment, and industrial scalability to facilitate commercial implementation.

References

1. Rowell, R.M., *Handbook of Wood Chemistry and Wood Composites*, CRC Press, Boca Raton, FL, USA, 2012.
2. Papadopoulos, A.N., *Polymers* **12** (2020) 2079.
3. Ashori, A., *Composites Part B* **39** (2008) 84–95.
4. Faruk, O., Bledzki, A.K., Fink, H.-P., Sain, M., *Progress in Polymer Science* **37** (2012) 1552–1596.
5. Fiore, V., Scalici, T., Valenza, A., *Composites Part B* **68** (2015) 235–242.
6. Pickering, K.L., Efendy, M.G.A., Le, T.M., *Composites Part A* **83** (2016) 98–112.
7. Bektas, I., Guler, C., Kalaycioglu, H., *BioResources* **15** (2020) 4321–4345.
8. Malanit, P., Laemsak, N., Ounsakul, M., *Construction and Building Materials* **25** (2011) 438–444.
9. Yang, H., Yan, R., Chen, H., Lee, D.H., Zheng, C., *Fuel* **86** (2007) 1781–1788.
10. Ayrlimis, N., *Industrial Crops and Products* **52** (2014) 416–423.
11. Mohanty, A.K., Misra, M., Drzal, L.T., *Natural Fibers, Biopolymers, and Biocomposites*, CRC Press, Boca Raton, FL, USA, 2005.
12. EN 312, *Particleboards – Specifications*, European Committee for Standardization (CEN), Brussels, Belgium.
13. ISO 4589, *Plastics – Determination of burning behavior by oxygen index*, International Organization for Standardization (ISO), Geneva, Switzerland.
14. Karabulut, A.F., Production of non-flammable, water-repellent polymeric composite particleboard from peanut shell, *Journal of Polytechnic (Politeknik Dergisi)* **23** (2020) 983–990.
15. Şahin, A., Taşdemir, H.M., Karabulut, A.F., Gürü, M., *Arabian Journal for Science and Engineering* **42** (2017) 4857–4866.
16. Taşdemir, H.M., Şahin, A., Karabulut, A.F., Gürü, M., Investigation of the properties of composite material produced from mint fiber added waste palm kernel, *Journal of the Faculty of Engineering and Architecture of Gazi University* **34** (1) (2019) 409–420.
17. Taşdemir, H.M., Şahin, A., Karabulut, A.F., Gürü, M., *Cellulose Chemistry and Technology* **53** (2019) 189–198.
18. Gürü, M., Karabulut, A.F., Aydın, M.Y., Bilici, İ., *High Temperature Materials and Processes* **34** (2015) 865–872.

CHAPTER 30

From Neural ODEs to Liquid Time-Constant Networks and Closed-form Continuous-time Policies: A Control-Oriented Review

Asaf Sayıl¹

¹ Sivas University of Science and Technology, Department of Astronautical Engineering

Why Continuous Time Re-entered Learning for Robotics

Robotic control problems are fundamentally continuous-time: a physical system evolves according to differential equations, while sensors and controllers interact with it through sampled measurements and digital computation. The resurgence of continuous-time learning—including Neural Ordinary Differential Equations (Neural ODEs) (Chen et al., 2018), liquid neural models such as Liquid Time-Constant (LTC) networks (Hasani et al., 2021), and closed-form continuous-time recurrent cells (CfC) (Hasani et al., 2022)—can be read as a response to recurring friction points in robotics and control-centric machine learning. In parallel, architectural priors such as Neural Circuit Policies (NCP) (Lechner et al., 2020) have emerged to impose structure and sparsity, motivated by auditability and deployment constraints. In parallel, reinforcement learning (RL) has also been adopted in some control and robotics settings as a policy optimization approach, often discussed together with stability and performance considerations (Demircioğlu & Bakır, 2024, 2025).

(P1) Time is not an implementation detail.

Many learned policies are trained and evaluated under a fixed sampling cadence (e.g., simulator step size), while real robots experience time-varying computation, communication delays, and sensing rates. Frame drops, jitter, and multi-rate sensing (camera vs. IMU vs. telemetry) introduce variability in the effective discretization of the closed loop. Controllers that implicitly assume a fixed step size can fail under deployment conditions that induce a different effective Δt .

(P2) Discretization couples modeling and numerical stability.

Even when an underlying continuous-time closed loop would be stable, its discrete-time implementation may become unstable if the effective time step is too large or the numerical integration is ill-suited (e.g., stiffness, solver tolerances). In learning-based pipelines, discretization is often hidden inside the architecture (e.g., an RNN update) or inside an ODE solver whose step-size control is not designed with closed-loop stability in mind (Hairer et al., 1993).

(P3) Memory and timescales are task-dependent.

Robotic tasks require accumulating context (partial observability, drift, noise) while reacting quickly to disturbances. Classical RNNs implement memory in discrete time but do not explicitly represent time or multiple timescales. Liquid models were explicitly designed to let effective time constants change with the current state and input, producing context-dependent memory horizons (Hasani et al., 2021).

A unifying view: policies as dynamical systems

A useful starting point is to view both the plant and the controller as dynamical systems. Consider a plant with state $s(t) \in \mathbb{R}^{n_s}$ and control $u(t) \in \mathbb{R}^{n_u}$:

$$\dot{s}(t) = f(s(t), u(t), t), \quad y(t) = g(s(t)), \quad (1)$$

where $y(t)$ denotes measured outputs. A learned controller with internal state $h(t) \in \mathbb{R}^{n_h}$ can be written as

$$\dot{h}(t) = F_\theta(h(t), o(t), t), \quad u(t) = \pi_\theta(h(t), o(t)), \quad (2)$$

where $o(t)$ is the observation stream (e.g., $o(t) = y(t)$ or image features).

In real systems, both plant and controller are observed/updated at discrete times $\{t_k\}_{k \geq 0}$ with gaps $\Delta t_k = t_{k+1} - t_k$. Deployment therefore induces a discrete-time interconnection:

$$\begin{aligned} s_{k+1} &= \Phi(s_k, u_k; \Delta t_k), \\ h_{k+1} &= \Psi_\theta(h_k, o_k; \Delta t_k), \quad u_k = \pi_\theta(h_k, o_k), \end{aligned} \quad (3)$$

where Φ and Ψ_θ depend on how continuous dynamics are discretized (explicitly or implicitly). The central question motivating modern continuous-time learning methods is:

Can we design and train controllers whose discrete-time realization $\Psi_\theta(\cdot; \Delta t)$ remains sensible, stable, and consistent across variable Δt and across deployment conditions?

A brief historical arc (problem $\rightarrow$ method $\rightarrow$ trade-off)

Neural ODEs (Chen et al., 2018): learn a continuous-time vector field and rely on numerical integration. Promise: time-consistency. Cost: solver dependence and potential stiffness/latency variability (Hairer et al., 1993).

Liquid models (LTC/LNN) (Hasani et al., 2021): engineer recurrent dynamics with input/state-dependent time constants to obtain adaptive timescales. Cost: naive implementations still need ODE integration.

CfC (Hasani et al., 2022): retain explicit Δt dependence via a closed-form-inspired update, avoiding per-step ODE solvers. Trade-off: a structured approximation of continuous dynamics.

NCP and CfC-NCP (Lechner et al., 2020): impose wiring priors (structured sparsity) to improve efficiency and inductive bias, particularly in embedded control.

Neural Ordinary Differential Equations: Learning Vector Fields

Neural ODEs (Chen et al., 2018) propose to parameterize dynamics in continuous time and treat inference as ODE integration. In robotics, this aligns learning models with continuous-time physics and makes time irregularity explicit rather than implicit.

Model definition

A basic Neural ODE defines latent dynamics

$$\dot{z}(t) = f_{\theta}(z(t), t), \quad (4)$$

optionally with an exogenous input $x(t)$,

$$\dot{z}(t) = f_{\theta}(z(t), x(t), t). \quad (5)$$

Observations or outputs are obtained through a readout:

$$y(t) = g_{\theta}(z(t)). \quad (6)$$

Given samples at times $t_0 < t_1 < \dots < t_K$, inference requires solving the initial value problem on each interval:

$$z(t_{k+1}) = \text{ODESolve}(f_{\theta}, z(t_k), [t_k, t_{k+1}]). \quad (7)$$

A loss is defined over measured times:

$$\mathcal{L}(\theta) = \sum_{k=0}^K \ell(y(t_k), \hat{y}(t_k)), \quad (8)$$

where $\hat{y}(t_k)$ is obtained by integrating eqn. 4 – eqn. 6.

What problem did Neural ODEs address?

From a robotics viewpoint, Neural ODEs address a mismatch: conventional RNNs define a discrete recurrence whose behavior can change unpredictably if the effective step size changes. In contrast, a continuous-time definition makes sampling variability explicit: changing sampling times corresponds to integrating for a different Δt .

Solver dependence, stiffness, and real-time implications

Neural ODEs shift complexity to numerical analysis. Inference and training depend on:

the ODE solver (Euler, RK schemes, adaptive methods),

local error tolerances and step-size control,

stiffness properties of f_{θ} (which may emerge during training),

and these choices influence both accuracy and compute variability (Hairer et al., 1993). For control, variable compute can feed back into variable Δt , turning the solver into part of the closed-loop system.

A control-relevant stability reminder (discretization can destabilize)

Consider the stable scalar ODE

$$\dot{z} = -\lambda z, \quad \lambda > 0. \quad (9)$$

Forward Euler yields

$$z_{k+1} = (1 - \lambda \Delta t) z_k, \quad (10)$$

which is stable only if $|1 - \lambda \Delta t| < 1$, i.e. $0 < \Delta t < 2/\lambda$ (Hairer et al., 1993). This illustrates a broader point: even stable continuous-time dynamics can be destabilized by a poor discrete realization.

Training: differentiating through ODE solutions

Neural ODE training requires gradients through [eqn. 9](#). Two main routes are used:

Differentiate-through-solver (unrolled backprop): treat solver steps as a computation graph (memory-heavy).

Adjoint sensitivity method: compute gradients by solving a backward-in-time adjoint ODE (Chen et al., 2018).

Adjoint derivation (step-by-step, publication-grade)

Assume

$$\dot{z}(t) = f_\theta(z(t), t), \quad z(t_0) = z_0, \quad \mathcal{L} = \ell(z(T)). \quad (11)$$

Step 1: Lagrangian for constrained optimization.

Introduce $a(t) \in \mathbb{R}^n$ and define

$$\mathcal{J}(\theta, z, a) = \ell(z(T)) + \int_{t_0}^T a(t)^\top (\dot{z}(t) - f_\theta(z(t), t)) dt. \quad (12)$$

Step 2: Variation w.r.t. z .

Let $z(t) \mapsto z(t) + \epsilon \delta z(t)$ with $\delta z(t_0) = 0$:

$$\delta \mathcal{J} = \frac{\partial \ell}{\partial z}(z(T))^\top \delta z(T) + \int_{t_0}^T a(t)^\top (\delta \dot{z}(t) - \frac{\partial f_\theta}{\partial z}(z(t), t) \delta z(t)) dt. \quad (13)$$

Step 3: Integration by parts.

Using $\int a^\top \delta \dot{z} = a^\top \delta z|_{t_0}^T - \int \dot{a}^\top \delta z$:

$$\begin{aligned} \delta \mathcal{J} = & \frac{\partial \ell}{\partial \mathbf{z}}(\mathbf{z}(T))^{\top} \delta \mathbf{z}(T) + \mathbf{a}(T)^{\top} \delta \mathbf{z}(T) - \int_{t_0}^T \dot{\mathbf{a}}(t)^{\top} \delta \mathbf{z}(t) dt \\ & - \int_{t_0}^T \mathbf{a}(t)^{\top} \frac{\partial f_{\theta}}{\partial \mathbf{z}}(\mathbf{z}(t), t) \delta \mathbf{z}(t) dt. \end{aligned} \quad (14)$$

Step 4: Choose $\mathbf{a}(t)$ so $\delta \mathcal{J} = 0$ for all $\delta \mathbf{z}(t)$.

This holds if

$$\begin{aligned} \dot{\mathbf{a}}(t) &= - \left(\frac{\partial f_{\theta}}{\partial \mathbf{z}}(\mathbf{z}(t), t) \right)^{\top} \mathbf{a}(t), \\ \mathbf{a}(T) &= - \left(\frac{\partial \ell}{\partial \mathbf{z}}(\mathbf{z}(T)) \right). \end{aligned} \quad (15)$$

Step 5: Gradient w.r.t. θ .

Differentiating [eqn. 15](#) w.r.t. θ gives

$$\frac{d\mathcal{L}}{d\theta} = - \int_{t_0}^T \mathbf{a}(t)^{\top} \frac{\partial f_{\theta}}{\partial \theta}(\mathbf{z}(t), t) dt. \quad (16)$$

Control interpretation.

The adjoint propagates sensitivity backward through time, akin to continuous-time BPTT. For robotics, numerical robustness of backward integration (and stiffness) is not a footnote but a deployment concern ([Hairer et al., 1993](#)).

Where did Neural ODEs lead?

Neural ODEs established the design principle that time should be explicit. However, solver dependence and variable compute can be problematic for real-time control. This motivated approaches that bake time-handling and stability-friendly structure into the recurrent cell itself—the perspective taken by liquid networks ([Hasani et al., 2021](#)) and CfC ([Hasani et al., 2022](#)).

Liquid Dynamics for Control: LTC/LNN and the Rise of Closed-form Continuous-time Cells

Liquid Time-Constant (LTC) networks ([Hasani et al., 2021](#)) (often discussed under the “Liquid Neural Networks” umbrella) were designed to address a key deficiency of conventional RNNs in robotics: fixed, implicit timescales. The central idea is to endow each unit with an effective time constant that depends on context, enabling adaptive memory horizons without online parameter updates.

From leaky integrators to time-aware recurrence

A canonical continuous-time memory unit is the leaky integrator:

$$\tau \dot{\mathbf{h}}(t) = -\mathbf{h}(t) + \boldsymbol{\eta}(t), \quad (17)$$

with time constant $\tau > 0$. If $\eta(t)$ is approximately constant over $[t_k, t_{k+1}]$, the solution has the form

$$h_{k+1} \approx e^{-\Delta t_k/\tau} h_k + (1 - e^{-\Delta t_k/\tau}) \eta_k. \quad (18)$$

This is an explicit Δt -dependent mixing between memory (h_k) and drive (η_k).

LTC/LNN: making the time constant a learned function

LTC generalizes [eqn. 18](#) by allowing the effective decay rate to depend on state and input. A useful abstraction is

$$\dot{h}(t) = -\alpha_\theta(h(t), x(t)) h(t) + \alpha_\theta(h(t), x(t)) A_\theta(h(t), x(t)), \quad (19)$$

with $\alpha_\theta(\cdot) \geq 0$. The effective time constant is

$$\tau_{\text{eff}}(h, x) = \frac{1}{\alpha_\theta(h, x)}. \quad (20)$$

Thus, a single trained model can behave as a “slow integrator” in some regimes and a “fast reactor” in others ([Hasani et al., 2021](#)).

Dynamic adaptivity vs. adaptive control.

Equation [eqn. 20](#) yields dynamic adaptivity (context-dependent dynamics) with fixed parameters θ . Classical adaptive control updates parameters online via an adaptation law ([Khalil, 2002](#)); the two notions are conceptually distinct (though they can be combined).

Vector form and coupling

For $h(t) \in \mathbb{R}^{n_h}$:

$$\dot{h}(t) = -\Lambda_\theta(h(t), x(t)) h(t) + \beta_\theta(h(t), x(t)), \quad (21)$$

where Λ_θ is often diagonal or sparse and constrained to be nonnegative (dissipative), while β_θ aggregates recurrent and feedforward drives.

Control-friendly sufficient conditions (boundedness template)

Lemma (ultimate boundedness template).

Assume $\Lambda_\theta(h, x) \succcurlyeq \lambda I$ uniformly and $\|\beta_\theta(h, x)\| \leq B$ for all encountered (h, x) . Then solutions satisfy

$$\|h(t)\| \leq e^{-\lambda(t-t_0)} \|h(t_0)\| + \frac{B}{\lambda} \left(1 - e^{-\lambda(t-t_0)}\right). \quad (22)$$

Proof sketch.

Apply a comparison argument to $\frac{d}{dt} \|h\| \leq -\lambda \|h\| + B$.

The remaining friction: solver dependence

If above equation is implemented as a true ODE, inference requires numerical integration per step. This reintroduces solver cost and step-size complications (Hairer et al., 1993), motivating a solver-free continuous-time-inspired update.

CfC: closed-form-inspired continuous-time recurrence (review perspective)

CfC as a learned exponential integrator.

CfC implements a bounded, learnable approximation of a time-aware update via a gate that depends explicitly on Δt_k :

$$h_{k+1} = a_\theta(h_k, x_k, \Delta t_k) \odot g_\theta(h_k, x_k) + (1 - a_\theta(h_k, x_k, \Delta t_k)) \odot r_\theta(h_k, x_k), \quad (23)$$

with a typical gate parameterization

$$a_\theta(h, x, \Delta t) = \sigma(-f_\theta(h, x) \Delta t),$$

where f_θ acts as a learned rate (Hasani et al., 2022).

Architectural priors: NCP and CfC-NCP (briefly)

Neural Circuit Policies (NCP) (Lechner et al., 2020) introduce a wiring prior: structured sparse connectivity to improve efficiency and auditability. CfC-NCP combines CfC recurrence with NCP wiring constraints, yielding compact Δt -aware controllers.

Stability, Robustness, and Control-Oriented Design Principles

Incremental stability / contraction (sufficient condition)

A stronger sufficient condition is contraction (incremental stability) (Lohmiller & Slotine, 1998). Suppose the controller update is contractive in h for fixed $(o, \Delta t)$:

$$\|\Psi_\theta(h_1, o; \Delta t) - \Psi_\theta(h_2, o; \Delta t)\| \leq c(\Delta t) \|h_1 - h_2\|, \quad c(\Delta t) < 1. \quad (24)$$

Small-gain intuition for robustness

Small-gain arguments provide intuition for interconnections (Khalil, 2002) and motivate penalizing high-frequency action changes and limiting recurrent Jacobian norms.

Time-step variation and why explicit Δt helps

CfC explicitly conditions on Δt , so the model can learn time-robust policies by design (Hasani et al., 2022). The key point is not that CfC guarantees stability, but that it exposes Δt as a first-class variable.

References

- Chen, R. T. Q., Rubanova, Y., Bettencourt, J., & Duvenaud, D. K. (2018). Neural ordinary differential equations. *Advances in Neural Information Processing Systems (NeurIPS)*.
- Hairer, E., Nørsett, S. P., & Wanner, G. (1993). *Solving ordinary differential equations I: Nonstiff problems*. Springer.
- Hasani, R., Lechner, M., Amini, A., Rus, D., & Grosu, R. (2021). Liquid time-constant networks. *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*, 7657–7666.
- Hasani, R., Lechner, M., Amini, A., Liebenwein, L., Ray, A., Tschaikowski, M., Teschl, G., & Rus, D. (2022). Closed-form continuous-time neural networks. *Nature Machine Intelligence*, 4(11), 992–1003. <https://doi.org/10.1038/s42256-022-00556-7>
- Khalil, H. K. (2002). *Nonlinear systems* (3rd ed.). Prentice Hall.
- Lechner, M., Hasani, R., Amini, A., Henzinger, T. A., Rus, D., & Grosu, R. (2020). Neural circuit policies enabling auditable autonomy. *Nature Machine Intelligence*, 2, 642–652. <https://doi.org/10.1038/s42256-020-00237-3>
- Lohmiller, W., & Slotine, J.-J. E. (1998). On contraction analysis for nonlinear systems. *Automatica*, 34(6), 683–696. [https://doi.org/10.1016/S0005-1098\(98\)00019-3](https://doi.org/10.1016/S0005-1098(98)00019-3)
- Demircioğlu, U., & Bakır, H. (2025). Reinforcement learning–driven proportional–integral–derivative controller tuning for mass–spring systems: Stability, performance, and hyperparameter analysis. *Engineering Applications of Artificial Intelligence*, 162, 112692.
- Demircioğlu, U., & Bakır, H. (2024). Artificial intelligence-based position control: reinforcement learning approach in spring mass damper systems. *Physica Scripta*, 99(4), 046003.

CHAPTER 31

From Discrete Neural Network Graphs To Liquid Time-Constant Networks

Asaf Sayıl¹

¹ Sivas University of Science and Technology, Department of Astronautical Engineering

1. Introduction: why continuous time, and why structure?

Deep neural networks are traditionally specified in discrete layers and trained to approximate static maps. However, in control and robotics, the *clock is part of the system*: sensors arrive asynchronously, actuators saturate, computation budgets fluctuate, and closed-loop stability depends on temporal consistency. Classical discrete-time RNNs can model sequences, but their implicit time-scale is tied to the sampling rate and training rollouts, making them sensitive to *rate mismatch* and long-horizon integration errors. Continuous-time modeling addresses this by learning dynamics in the form of differential equations, enabling principled handling of irregular sampling and allowing explicit reasoning about time scales. In parallel, reinforcement learning (RL) has been used as a practical control design route in related benchmark dynamical systems, including RL-based position control and RL-driven PID tuning studies that explicitly examine stability and performance (Demircioğlu & Bakır, 2024, 2025)

Yet *generic* neural ODEs often define the derivative by an unconstrained network, raising questions about stability, boundedness, and computational cost (numerical integration at inference). LTC networks were introduced as a middle ground: they preserve the continuous-time formulation while embedding a specific first-order dynamical structure that supports bounded behavior and interpretable time constants, at the cost of still requiring an ODE solver at runtime.

2. Neural Network Graphs (NNGs): discrete computation as a directed graph

2.1. Feedforward NNGs

A standard feedforward network can be formalized as a directed acyclic graph (DAG) whose nodes implement affine maps and nonlinearities. Let $x \in \mathbb{R}^{d_0}$ be the input and define layerwise states

$$h^{(\ell+1)} = \phi^{(\ell)}(W^{(\ell)}h^{(\ell)} + b^{(\ell)}), \quad \ell = 0, \dots, L-1, \quad (1)$$

with $h^{(0)} = x$ and output $y = h^{(L)}$. The computational graph viewpoint is useful for (i) automatic differentiation, (ii) modular composition, and (iii) interpreting depth as iterative refinement.

2.2. Expressivity and approximation

For suitable nonlinearities, multilayer feedforward networks are universal approximators of continuous functions on compact sets. From a control perspective, this motivates using NNGs to approximate feedback laws, value functions, or disturbance maps. However, static approximation does not automatically yield robust closed-loop behavior: stability depends on the interconnection of the learned map with the plant (Hornik et al., 1989).

3. Discrete-time recurrence: RNNs as dynamical systems

3.1. Vanilla RNNs

A discrete-time RNN defines a state recursion

$$h_{k+1} = \sigma(W_h h_k + W_x x_k + b), \quad y_k = W_y h_k, \quad (2)$$

which is a nonlinear discrete dynamical system driven by inputs x_k . Backpropagation through time (BPTT) differentiates through the unrolled graph.

3.2. Vanishing/exploding gradients as Jacobian products

Let $J_k = \partial h_{k+1} / \partial h_k$. Gradients through time contain products $\prod_{j=0}^k J_j$, so their norm is governed by the spectral properties of the Jacobians. This yields the classical vanishing/exploding gradient pathology that motivated gated architectures.

3.3. Gated recurrence (LSTM/GRU) and implicit time scales

LSTM (Hochreiter & Schmidhuber, 1997) and GRU (Cho et al., 2014) introduce gates that regulate information flow. While effective, their time scale remains implicit: the “step” is the sampling period. If the deployment sampling differs from training (common in embedded systems with jitter), the learned recurrence may not transfer gracefully without careful augmentation.

4. Continuous-time recurrence and Neural ODEs

4.1. CT-RNNs

A canonical continuous-time recurrent model is the leaky, input-driven ODE

$$\dot{x}(t) = -\frac{1}{\tau} x(t) + f(x(t), u(t), t; \theta), \quad (3)$$

where τ sets a baseline time constant and f is a parameterized nonlinearity. Such models can approximate a broad class of dynamical systems under mild conditions (Funahashi & Nakamura, 1993).

4.2. Neural ODEs

Neural ODEs further generalize by defining the vector field directly:

$$\dot{x}(t) = f(x(t), I(t), t; \theta), \quad (4)$$

and computing $x(t)$ via a numerical ODE solver (Chen et al., 2018). Training can differentiate through the solver or treat it as a black box and apply adjoint-based gradients (Chen et al., 2018; Gholami et al., 2019; Dupont et al., 2019). In many robotics settings, the central trade-off is: expressive continuous-time dynamics versus runtime cost and numerical sensitivity.

4.3. Why “unstructured” neural ODEs can be problematic in control

From a control viewpoint, an unconstrained f can learn stiff dynamics, saturating nonlinearities, or unstable modes that (i) demand small solver steps, (ii) amplify discretization error, or (iii) degrade robustness under model mismatch. This does not imply neural ODEs are unusable; rather, it motivates injecting *structure* that biases learning toward well-behaved dynamics.

5. Liquid Time-Constant (LTC) networks

5.1. Design principle

LTC networks define a continuous-time RNN by coupling *linear first-order dynamics* with *nonlinear interlinked gates* so that the effective time constant becomes state-dependent (“liquid”). Conceptually, rather than learning $\dot{x} = f(\cdot)$ directly, LTCs learn how gates modulate decay and input injection terms.

5.2. Single-neuron dynamics

A representative scalar LTC form (as used in the LTC/CfC literature) can be written as

$$\dot{x}(t) = -(w_\tau + f(x(t), I(t)))x(t) + A f(x(t), I(t)), \quad (5)$$

where w_τ is a baseline decay parameter, A is an affine target level, and f is a bounded nonlinearity (often sigmoid-like gating). This equation makes explicit two stabilizing mechanisms: (i) dissipation through the $-w_\tau x(t)$ term, (ii) bounded drive through $A f(x(t), I(t))$ when f is bounded (Hasani et al., 2021; Hasani et al., 2022).

5.3. Network dynamics

For a network state $x(t) \in \mathbb{R}^n$, one writes a coupled system

$$\dot{x}(t) = -(D_\tau + G(x(t), I(t)))x(t) + A(x(t), I(t)), \quad (6)$$

where D_τ is typically diagonal with positive entries, and $G(\cdot)$ encodes gate-modulated coupling across neurons. The precise parameterization varies (dense, sparse, or circuit-wired), but the key idea is preserved: *first-order stable building blocks + bounded gates*.

5.4. Training and inference

Unlike discrete RNNs, LTCs require integrating (5) (or its vector counterpart) over time, typically with standard ODE solvers. This makes LTCs attractive for irregular sampling but introduces a compute bottleneck in high-rate control loops (Hasani et al., 2021).

6. Mathematical analysis: existence, boundedness mechanisms, and stability-relevant structure

6.1. Well-posedness

Assume $f(\cdot)$ is locally Lipschitz in x and measurable in t (standard in ODE theory). Then for any initial condition $x(t_0) = x_0$, there exists a unique maximal solution on an interval. Most practical LTC parameterizations use smooth activations, ensuring local Lipschitz continuity.

6.2. Why bounded gates matter

If $f(\cdot)$ is bounded (e.g., sigmoid with $0 \leq f \leq 1$), then the input injection term Af is bounded whenever A is bounded. Moreover, the decay coefficient $w_\tau + f$ stays positive, providing dissipativity. These two properties enable explicit bounds on trajectories under suitable parameter constraints.

6.3. Liquid time constants as adaptive dissipation

Rewrite (5) as

$$\dot{x}(t) = -\frac{1}{\tau_{\text{sys}}(t)} x(t) + \frac{1}{\tau_{\text{sys}}(t)} x_\infty(t), \quad (7)$$

where the *effective* time constant and target are

$$\tau_{\text{sys}}(t) \triangleq \frac{1}{w_\tau + f(x(t), I(t))}, \quad x_\infty(t) \triangleq \frac{A f(x(t), I(t))}{w_\tau + f(x(t), I(t))}. \quad (8)$$

Hence the dynamics continuously interpolate between different time scales as the gate changes with state and input. This is precisely the “liquid” property: the network can slow down or speed up its internal response depending on context.

6.4. Control-relevant interpretation

In feedback control, dissipation corresponds to stabilizing energy decay. LTCs embed a tunable dissipation term $-(w_\tau + f)x$ whose strength changes online. This can be interpreted as a learned, state-dependent “leak” that regularizes memory without requiring explicit spectral normalization tricks often used in discrete RNN stabilization.

7. Strengths and shortcomings of LTCs

7.1. Strengths

- Continuous-time semantics: naturally handles irregular sampling and time stamps (Hasani et al., 2021).
- Structured dynamics: first-order stable backbone with bounded gating encourages well-behaved trajectories (Hasani et al., 2021).
- **Interpretability hooks:** time constants and gates can be inspected as internal adaptation signals.

7.2. Shortcomings

- **Numerical integration at inference:** high-rate control can be limited by solver cost and stiffness sensitivity.
- **Solver-induced variability:** step-size choices introduce another layer of engineering that interacts with deployment hardware.

These limitations motivated the later development of closed-form variants (treated independently in a later chapter).

8. Conclusion

NNGs provide a natural computational-graph language for modern learning systems, but discrete recurrence ties models to specific sampling assumptions. Continuous-time approaches (CT-RNNs and neural ODEs) restore time as a first-class variable, and LTC networks further inject stabilizing first-order structure through liquid time constants and bounded gates. This structured continuous-time view is particularly aligned with robotics, where timing, stability margins, and deployment constraints are inseparable.

9. References

- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323, 533–536.
- Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359–366.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780.
- Cho, K., van Merriënboer, B., Bahdanau, D., & Bengio, Y. (2014). On the properties of neural machine translation: Encoder–decoder approaches. In *Proceedings of the Workshop on Syntax, Semantics and Structure in Statistical Translation*.
- Funahashi, K.-i., & Nakamura, Y. (1993). Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Networks*, 6(6), 801–806.
- Chen, R. T. Q., Rubanova, Y., Bettencourt, J., & Duvenaud, D. K. (2018). Neural ordinary differential equations. In *Advances in Neural Information Processing Systems (NeurIPS)*.
- Dupont, E., Doucet, A., & Teh, Y. W. (2019). Augmented neural ODEs. In *Advances in Neural Information Processing Systems (NeurIPS)*.
- Gholami, A., Keutzer, K., & Biros, G. (2019). ANODE: Unconditionally accurate memory-efficient gradients for neural ODEs. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*.
- Hasani, R., Lechner, M., Amini, A., Rus, D., & Grosu, R. (2021). Liquid time-constant networks. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 35, 7657–7666.
- Hasani, R., Lechner, M., Amini, A., Liebenwein, L., Ray, A., Tschaikowski, M., Teschl, G., & Rus, D. (2022). Closed-form continuous-time neural networks. *Nature Machine Intelligence*, 4(11), 992–1003. <https://doi.org/10.1038/s42256-022-00556-7>

