ALL SCIENCES ACADEMY

6th International Conference on Modern and Advanced Research ICMAR 2026



August 26-27, 2026 | **Proceeding Book** | Konya, Turkey

All Sciences Academy

PROCEEDING BOOK OF 6TH INTERNATIONAL CONFERENCE ON MODERN AND ADVANCED RESEARCH ICMAR 2026

PROCEEDING BOOK OF 6TH ICMAR 2026:

16-17 August 2026

Konya, Turkey

Publication date: 25.08.2026

Publisher: All Sciences Academy

Proceeding Book of 6th International Conference on Modern and Advanced Research ICMAR 2026

https://as-proceeding.com/index.php/icmar/home

6th International Conference on Modern and Advanced Research ICMAR 2026: edited by All Sciences Academy.

ISBN: 978-625-8839-61-6

As this is a serial book published, it has ISBN. All abstracts were peer reviewed.

ISBN: 978-625-8839-61-6

Printed in the Turkey with ISBN 978-625-8839-61-6

International Conference Website

https://as-proceeding.com/index.php/icmar/home

INTRODUCTION

We had the great honor of 6th International Conference on Modern and Advanced Research ICMAR 2026. It was truly a great pleasure for us to greet a lot of participants from many different countries to attend ICMAR 2026! We firmly believe that the conference will become an important international event in the field of cross-industry discussion about innovations in Academic Studies.

Three cooperating organizations supported the two-day conference. There were 146 papers accepted for presentation at ICMAR 2026, contributed from different countries. We had plenary speeches and several well-known scientists and experts, to give invited talks at differentsessions.

The purpose of ICMAR 2026 was to provide a forum for the participants to report and review innovative ideas, with up-to-date progress and developments, and discuss novel approaches to the application in the field of their own research areas and discuss challenges of doing science.

We sincerely hope that the exchange of ideas on doing research, science and improving education will help the participants, and international cooperation sharing the common interest will be enhanced.

On behalf the Organization Committee of ICMAR 2026, we would like to heartily thank our cooperating organizations for all they have done for the conference. We would also like to thank the authors for theircontribution to the proceedings; the participants and friends of ICMAR 2026, for their interest and efforts in helping us to make the conference possible; and the Editorial boards for their effective work and valuable advice, especially the ICMAR 2026 secretariat and the ICMAR 2026 staff, for their tireless efforts and outstanding services in preparing the conference and publishing the Proceedings.

Publisher: All Sciences Academy

TABLE OF CONTENTS

Introduction II
Table of Contents III
Conference Chairs XI
Scientific Committees XII
Document of Academic Incentive XIII
Document of International Conference XIV

Table List of Content

Content of Abstract Papers

Page No	Title	Authors	Country	Submission ID
1	Neural Network and Image Processing Applications in Engineering Management and Telecommunications: A Convergent Framework	Magda JumaShuayb Albaraesi, Hourie Khalifa Almejrab	Libya, Libya	Mail Submission 2
2	From Classroom to Institution: Evaluating the Reliability of AI-Powered Simultaneous Translation and Deep Learning Writing Tools in Libyan Higher Education and Medium-Size Public Organizations	Rima Subhi Husain Taher, Llahm Ben Dalla, Mohammed A. Abo-Sahal, Tasnem Elsseid, Fakroun, E. Asma Agaal	Libya, Turkey, Libya, Libya, Libya	Mail Submission 5
3	An Adaptive Multi-Factor Authentication Framework for Enhancing User-Centric Security in Dynamic Threat Environments	B. Jimoh, G. George, C. Isah	Nigeria, Nigeria, Nigeria	Mail Submission 6
4	High prevalence of amoxiclav Resistance in Klebsiella pneumoniae isolated from aquatic environment	Vikar Ahmed, Syed Ahmed Rizvi, Qazi Mohd Rizwanul Haq	India, India, India	Mail Submission 7
5	Statistical Techniques for Modern Data Analysis and Technological Innovation	Alina Idrees	Pakistan	Mail Submission 8
6	Folklore as a Catalyst for Sustainable Social Development: An Advanced Research Perspective on Integrating Traditional Wisdom into Modern Frameworks	Habibur Rahman, Mobashir Rahman, Rajiwan Anwar Ratul	Bangladesh, Bangladesh, Bangladesh	Mail Submission 9
7	Experimental investigation of the mode fracture in adhesive	Amin HOUARI, Berivan Yilmazer POLAT, Yassine EL GUERRI, GUEMMOUR Mohamed Boutkhil, Alper POLAT	Turkey, Turkey, Algeria, Algeria, Turkey	Submission 652
8	Numerical modeling of the chemical kinetics of a plasma produced from N2-O2-NO by corona discharge	MEDJAHDI Ines Sarah, BELHOUR Souad, FEROUANI Abdelkarim, ABDELLI Imane, BENMANSOUR Sid Ahmed, BELAROUCI Salim, BENHAMEL Abdelhamid	Algeria, Algeria, Algeria, Algeria, Algeria, Algeria	Submission 655
9	Free Vibration Analysis of Functionally Graded Beams with Power-Law and Sigmoid Material Gradations	Abbes Ouadah	Algeria	Submission 656
10	Physics-Informed Data-Driven Framework for Cutting Force Prediction in Hard Turning of High-Strength Steels	Zaki Abdelfetah ABED	Algeria	Submission 658
11	Multi-Scale Hybrid Modeling of Chip Formation Mechanics and Force Dynamics in Hard Turning of MDN250 Maraging Steel	Zaki Abdelfetah ABED	Algeria	Submission 659
12	Physics-Guided Machine Learning and Finite Element Synergy for Chip Morphology and Force Assessment in MDN250 Turning	Zaki Abdelfetah ABED	Algeria	Submission 660
13	High-Dimensional Kernelized Surrogate Modeling for Predictive Cutting Force Assessment in Hard Machining of MDN250 Steel	Zaki Abdelfetah ABED	Algeria	Submission 661
14	FEA-Driven Machine Learning Surrogates for Real-Time Force Mechanics in Hard Turning of MDN250 Steel	Zaki Abdelfetah ABED	Algeria	Submission 662
15	Integrated Finite Element and Quadratic SVM Modeling of Cutting Forces in the Hard Turning of MDN250 Maraging Steel	Zaki Abdelfetah ABED	Algeria	Submission 663
16	Advanced Precision Cutting Force Prediction Analysis in MDN250 Steel Hard Turning Using Finite Element Modelling and Fine Gaussian Support Vector Machine	Zaki Abdelfetah ABED	Algeria	Submission 664
17	Forecast Modelling of Cutting Forces of MDN250 Steel Hard Turning Employing a Typical Finite Elements Analysis and a	Zaki Abdelfetah ABED	Algeria	Submission 665

	Medium Gaussian Support Vector Machine			
18	Improved Predictive Modelling of Cutting Forces in MDN250 Steel Hard Turning Combining ABAQUS Finite Element Simulation and Adaptive Neuro-Fuzzy Inference System (ANFIS)	Zaki Abdelfetah ABED	Algeria	Submission 666
19	Physics-Guided Rational Quadratic Gaussian Process Regression and FE Simulation for Advanced Cutting Mechanics in MDN250 Maraging Steel	Zaki Abdelfetah ABED	Algeria	Submission 667
20	Ultra-Fast FEA-Trained Gaussian Process Surrogates for Real-Time Cutting Force Prediction in MDN250 Orthogonal Machining	Zaki Abdelfetah ABED	Algeria	Submission 668
21	Uncertainty-Aware Thermo-Mechanical Force Modeling in Hard Turning of MDN250 Steel via Matérn 5/2 Matérn 5/2 Gaussian Process Surrogates	Zaki Abdelfetah ABED	Algeria	Submission 669
22	Hybrid Exponential GPR and 2D Finite Element Modeling of Cutting Forces in MDN250 Maraging Steel	Zaki Abdelfetah ABED	Algeria	Submission 670
23	Impact of Volume Fraction on the Dynamics of Decane/Water Microemulsions: A Molecular Dynamics Study	Mohammed Khatouri, Soumia Elkhaoui, Ayoub Arbia, Redouane Elhajjam, Zineb Basbassi, Lamiae Talha, Rachid Ahfir, Mohammed Filali	Morocco, Morocco, Morocco, Morocco, Morocco, Morocco, Morocco, Morocco	Submission 671
24	Assessing AI Applications in Environmental Monitoring: A Review	Erum Aamir	Pakistan	Submission 673
25	Post-Caesarean Stress and Consumer Behaviour: Examining Compulsive Buying in the Shadow of Social Influence	Shahid Mahmood, Raja Irfan Sabir	Pakistan, Pakistan	Submission 677
26	Long-term strain of concrete: Experimental study and finite element analysis	Farid Bouziadi, Nefissa Belkacem, Bensaid Boulekbache, Mostefa Hamrat, Touhami Tahenni, Mohamed Ladjel, Omar Kouider Djelloul	Algeria, Algeria, Algeria, Algeria, Algeria, Algeria, Algeria	Submission 681
27	Presence of Polychlorinated Biphenyls (PCBs) in the Sediments of Lake Ohrid	Hamit Ismaili, Arben Haziri, Ilirjana Osmani, Aurel Nuro	Kosovo, Albania, Albania, Albania	Submission 682
28	Biomimetic Architecture as a Strategy for Sustainable Agricultural Development in Semi-Arid Regions: The Case of an Agricultural Research Eco-Pole in Djelfa	Soumaya KERSENNA, Yakout OUMELKHEIR	Algeria, Algeria	Submission 684
29	Density Function Theory Research, Spectroscopy, Charge Transfer, and Nonlinear Optical Effects of a Thiazolidinone Derivative	Khaled Drim, Youcef Megrouss, Salem Yahiaoui	Algeria, Algeria, Algeria	Submission 685
30	Valorization Of a Biomaterial For Removal Of Cationic Dyes From Wastewater	Rafik El Arslene DRA, Mounya ZENASNI, Malika MEDJAHDI, Nadia RAMDANI	Algeria, Algeria, Algeria, Algeria	Submission 693
31	From the Saharan Atlas to al-Andalus and Back: Population Mobility, Collective Fortification and Architectural Continuity in Amazigh Mountain settlements	Kamal Youcef, Marouane Samir Guedouh	Algeria, Algeria	Submission 695
32	On Spectral Problems with Quasi-Periodic Boundary Conditions	Ibrahim Nabiev, Leyla Mammadova, Hamida Abdulova	Azerbaijan, Azerbaijan, Azerbaijan	Submission 696
33	Synthesis of Carbon Nanotubes from Waste Polyethylene over Talc-Supported Fe–Co Catalysts via Chemical Vapor Deposition	Muhammet Ali Esener, Altay Usta, Birce Pekmezci Karaman, Nuray Oktar	Turkey, Turkey, Turkey, Turkey	Submission 700
34	ANTIOXIDANT AND IMMUNE RESPONSE OF WEST AFRICAN DWARF (WAD) SHEEP DURING ACCLIMATISATION	Oladipupo, O. H., Ewuola, E. O.	Nigeria, Nigeria	Submission 703
35	Structural and Electronic Properties of WSe2/WS2 Heterobilayers	R. Masrour	Morocco	Submission 704
36	DFT Study of the Electronic, Optical, and Photovoltaic Properties of MoS2 Nanoflakes	R. Masrour	Morocco	Submission 705
37	DFT Investigation of the Structural, Electronic, and Electrochemical Properties of WSe2 Nanoflakes for Sustainable Energy Applications	R. Masrour	Morocco	Submission 706
38	Artificial Intelligence and Social Transformations in Light of Legal Protection	Benbouabdallah Mounia, Beddiar Maher	Algeria, Algeria	Submission 708
39	Screening and genomic characterization of a potential biogenic amine-producing lactic acid bacteria strain.	Bitrus B. MAIDAMBE, Kubrat A. OYINLOLA	Nigeria, Nigeria	Submission 709
40	Comparative Cradle-to-Gate Life Cycle Assessment of Burnt Clay and Fly Ash Bricks in Pakistan	Muhammad Hamza Bashir, Muhammad Farhan Javed, Saad Bilal Tabassum, Waqar Shafique	U.K., Pakistan, Germany, Pakistan	Submission 711

41	Comparative efficiency of numerical approaches for optimizing the extraction of alkaloid molecules	LIMANE Wouissem, RIHANI Rachida	Algeria, Algeria	Submission 712
42	Growth Performance of Red Swamp Crayfish (Procambarus clarkii) Fed with Chlorella sp. Supplemented Diets	Olcay ŞİMŞEK, Onur KARADAL	Turkey, Turkey	Submission 713
43	Optimizing FDM Process Parameters for Improved Tensile Properties of PLA Using Taguchi Design and ANOVA	Belkaid khmissi, Guerira Belhi	Algeria, Algeria	Submission 716
44	Parametric Investigation of Electromagnetic Response for Interfacial Defect Evaluation in Hybrid Composite–Metal Structures	Faiza Benlakhdar	Algeria	Submission 718
45	Individual and Household Determinants of Female Labour Force Participation. Micro-Level Evidence from the Western Balkans	Loredana Sulejmani, Elena Polo	Albania, Albania	Submission 723
46	Expression of lncRNAs and circRNAs and Circadian Regulation of GH	Noor N. Abdullateef, Owayes M. Hamed	Iraq, Iraq	Submission 724
47	PEM Elektrolizörlerinde Katalizör Maskelemesinin Gözenek Ölçeğindeki Dinamikleri: Çok Fazlı Lattice Boltzmann Yaklaşımı	Oğuz Ozan YOLCAN	Turkey	Submission 725
48	Eksozomal miRNA'lar: Endometrium–Embriyo İletişiminde İmplantasyon Başarısızlığının Hücresel İmzası mı?	Leyla Bahar	Turkey	Submission 726
49	Endometrial Organoid Modelleri ile İmplantasyon Penceresinin Fonksiyonel Simülasyonu: Kişiselleştirilmiş Üreme Tıbbında Yeni Dönem	Leyla Bahar	Turkey	Submission 727
50	FIRST PRINCIPLES STUDY OF CHARGE TRANSFER PHENOMENON IN ZrV2O7/GRAPHENE AND ZrTi2O6/GRAPHENE HETEROJUNCTIONS	Aqsa Laghari, Asim Mansha	Pakistan, Pakistan	Submission 729
51	Influence of Initial Carbon-to-Nitrogen Ratio on the Composting Performance and Maturity of Date Palm Waste	Abdelouahad Ouali, Fatima Hiouani, Abderrazak Mabrouki, Youcef Ouzani, Houda Boutalbi, Djamila Madani, Chafia Tir	Algeria, Algeria, Algeria, Algeria, Algeria, Algeria, Algeria	Submission 731
52	Dielectric Thickness-Dependent Mobility in Solution-Processed In_2O_3 Thin-Film Transistors: Quantifying the Semiconductor–Dielectric Interface Effect	Djemâa Ben Othmane	Algeria	Submission 734
53	From Traditional Knowledge to Technological Innovation: The Geography of Sustainable Primary Activities in a Changing Climate	OLANIYI, Oladiran Paul	Nigeria	Submission 735
54	Eğitimde Üretken Yapay Zekâ Araştırmaları Yeniden Üretilebilir mi? Raporlama Şeffaflığının İçerik Analizi (2023–2026)	Şener Balat	Turkey	Submission 741
55	Yttrium-Induced Structural and Morphological Modifications in Bi4Ti3O12 Layered Ceramics	Benbrika Chaima, Menasra Hayet, Sbaihi Amira	Algeria, Algeria, Algeria	Submission 742
56	Structural and Morphological Insights into Barium-Doped Bi4Ti3O12 Aurivillius Ceramics	Benbrika Chaima, Menasra Hayet, Sbaihi Amira	Algeria, Algeria, Algeria	Submission 744
57	Effect of Fluorine Doping on the Surface Morphology and Electrical Properties of NiS Thin Films Prepared by a Simplified Spray Pyrolysis Technique	Amira Sbaihi, Said Benramache, Abderraouf Zidane, Chaima Benbrika	Algeria, Algeria, Algeria, Algeria	Submission 745
58	On Strong and complete observability and for rectangular descriptor systems	Abed MAKRELOUFI	Algeria	Submission 746
59	Designing an Islamic Social Capital Market: An Agent-Based Approach	Rashed Jahangir	Turkey	Submission 750
60	Birlikte Dikim Koşullarında Sera Beyazsineğinin Domates ve Patlıcan Arasındaki Tercihi: Patlıcanın Tuzak Bitki Potansiyeli	Gökhan ERDOĞAN	Turkey	Submission 751
61	First-Principles Study of the Structural, Elastic, and Electronic Properties of La1-xYxN ($0 \leq x \leq 1$)	Nour El Imen Zeghoum, Said Lakel	Algeria, Algeria	Submission 752
62	Dekompanse Ve Kompanse Kalp Yetersizliği Hastalarında Nötrofil/Albümin Oranının Karşılaştırılması	Şeyda Şahin	Turkey	Submission 753
63	Isolation and Structural Characterization of Biologically Relevant Secondary Metabolites	Hiouani Khaled El Mehdi, Anede Nada, Habiba Laraoui	Italy, Algeria, Algeria	Submission 755
64	UTILIZATION OF CLOVE ESSENTIAL OIL AGAINST SELECTED FOOD-BORNE FUNGAL ISOLATES	Miracle D. OLAWUNI, Kubrat A. OYINLOLA	Nigeria, Nigeria	Submission 756

65	TOWARDS MITIGATING TRYPTAMINE ACCUMULATION IN FERMENTED CHEESE: SCREENING FOR POTENTIAL BACTERIA AND YEAST STARTERS	Dimeji D. AKINYELE, Kubrat A. OYINLOLA	Nigeria, Nigeria	Submission 759
66	The Legal and Financial Framework for Startups and Self-Employed Enterprises: Banking and Tax Incentives and the Role of Islamic Finance	ZEROUKI Bouzned	Algeria	Submission 760
67	Evaluating the Effect of PCM Integration on Solar Thermal Storage Performance	Littouche Billel, Rebai Billel, Ben Slimaine Hocine	Algeria, Algeria, Algeria	Submission 762
68	Production and Characterization of Hydroxyapatite from Fishery By-Products: Optimization of Calcination and Chemical Treatment	Malek Ibtihel Djerafi, Khatima Kaabeche, Amira Mensour, Nadia Taloub	Algeria, Algeria, Algeria, Algeria	Submission 763
69	MHD Mixed Convection and Entropy Generation in an Octagonal Cavity Containing a Hybrid Nanofluid	Kamel Zitouni	Algeria	Submission 765
70	Exogenous NADH restores pulmonary redox homeostasis and limits cigarette smoke-induced lung remodeling	Nada SLAMA, Amira Bouguedra, Karima Bahria, Amina Abdellatif, Karine Benachour	Algeria, Algeria, Algeria, Algeria, Algeria	Submission 766
71	Biological Treatment of Microcystin-Contaminated Waters: Mechanisms and Effectiveness	Tamer Fatma Zohra, Charef Nacira, Zaidi Hadjer, Larisa Lvova, Nouri Nada, Saada Naziha, Nasri Hichem	Algeria, Algeria, Algeria, Italy, Algeria, Algeria, Algeria, Algeria	Submission 770
72	Neural Operators for Fast Numerical Approximation of Partial Differential Equations	Baibeche Sabah	Algeria	Submission 771
73	Trophic ecology and diet composition of the common sole Solea solea (Linnaeus, 1758) in Oran Bay, Western Algeria	TABECHE Ali, BOUHADIBA Sultana, AOUADJ Sid Ahmed, BELHOUCINE Fatma, ALIOUA BERREBBAH Amel	Algeria, Algeria, Algeria, Algeria, Algeria	Submission 773
74	Relationship between feeding activity and the reproductive cycle in female Solea solea (Linnaeus, 1758) in Oran Bay	TABECHE Ali, BOUHADIBA Sultana, AOUADJ Sid Ahmed, BELHOUCINE Fatma, ALIOUA BERREBBAH Amel	Algeria, Algeria, Algeria, Algeria, Algeria	Submission 774
75	Functionalization of aramid/polypropylene composite membranes with chitosan–magnetite nanocomposites for superior performance	Javaria Kanwal, Sara Musaddiq, Alishba Naz, Sajjad Ahmad, Sadia Iqbal	Pakistan, Pakistan, Pakistan, Pakistan, Pakistan	Submission 779
76	Comparative study of clinicopathological features in triple negative breast cancer and non-triple negative breast cancer in Algeria.	BOUGUEDRA Amira, SLAMA Nada, KOUCHKAR Amal	Algeria, Algeria, Algeria	Submission 781
77	A Novel 4D Hyperchaotic System with No Equilibria: Synchronization and Circuit design	Fareh Hannachi	Algeria	Submission 782
78	Finite Element Analysis and Mechanical Characterization of Functionally Graded Lattice Structures Fabricated via Selective Laser Melting	Bengherbia Nardjes	Algeria	Submission 788
79	Numerical Investigation of Heat Transfer Enhancement and Aerodynamic Drag in Micro-Vortex Generator Integrated Heat Exchangers	Bengherbia Nardjes	Algeria	Submission 790
80	TOWARDS MITIGATING TRYPTAMINE ACCUMULATION IN FERMENTED CHEESE: SCREENING FOR POTENTIAL BACTERIA AND YEAST STARTERS	Dimeji D. AKINYELE, Kubrat A. OYINLOLA	Nigeria, Nigeria	Submission 791
81	EPIDEMIOLOGICAL ASSESSMENT OF HYPERLIPIDEMIA AND DEVELOPMENT OF A RATIONAL HERBAL FORMULATION	Khadija Almas, Wafa Majeed2, Ayesha	Pakistan, Pakistan, Pakistan	Submission 792
82	Linear Interpolation in CNC Turning: A Mathematical Approach to Toolpath Validation	Violeta Krcheva, Stojance Nusev, Miša Tomić	North Macedonia, North Macedonia, Serbia	Submission 793
83	Türkiye'nin Bilgi Güvenliği Açısından Ağ Güvenliği Politikaları: Uluslararası Uygulamalarla Karşılaştırmalı Analiz	Abdulbari Shihab Ahmed CHOLAQ, Ercan Nurcan YILMAZ	Turkey, Turkey	Submission 795
84	From Classroom to Startup: Reimagining Entrepreneurship Education for the Digital Economy	Sakariyyahu Shehu, Semiu Olawale Makinde, Afeez Adebayo Fawale, Abdullahi AbdulRafiu, Ahmed Akeem Apo	Nigeria, Nigeria, Nigeria, Nigeria, Nigeria	Submission 799
85	Algebraic Limit Cycles in Cubic Planar Differential Systems with a Degenerate Singularity and a Focus	Fouad KHIREDDINE, Mohamed GRAZEM	Algeria, Algeria	Submission 808
86	Stability of viscoelastic wave equation with boundary conditions	Ahmed Benzian	Algeria	Submission 809
87	FACTORS ASSOCIATED WITH	JIMOH, Adenike Bilikis	Nigeria	Submission 810

	STUDENTS' ANXIETY TOWARDS LEARNING BIOLOGICAL CONCEPTS IN SELECTED SECONDARY SCHOOLS IN ILORIN METROPOLIS, KWARA STATE			
88	Spatial Analysis of Wildfire Hazard and Climate Change Impacts Using Remote Sensing and Thermal Stress Indicators in Karabük Province, Turkey	Abbas Gulistani, Mehmet Kazim Yetik	Turkey, Turkey	Submission 811
89	Adaptive Quantum-Inspired Swarm Learning (AQISL) for Energy-Efficient Cluster Head Selection in IoT Infrastructure	C. Bala Subramanian, R. Ramkumar, Shanmugapriya S	India, India, India	Submission 672
90	Gender and Instructional Leadership Behaviors	Irfan Bashir, Afshan Naseem	Pakistan, Pakistan	Submission 715

Table List of Content

Content of Full Papers

Page No	Title	Authors	Country	Submission ID
1-7	Molecular Docking Analysis of Compounds Isolated from Marine Sponges Targeting EGFR	Adnan Cetin	Turkey	Mail Submission 1
8-11	Assessment of General Occupational Health and Safety Risks in a Metal Casting Plant Using the Fine-Kinney Method: A Case Study	Zehra Gülten Yalçın	Turkey	Mail Submission 3
12-18	THE ROLE OF SOCIAL MEDIA IN DENTAL TOURISM DECISION-MAKING	Orion Muçaj	Albania	Mail Submission 4
19-23	Anatomical Structure of the Urogygeal Gland (Preen Gland) in the Sparrowhawk, Common Kestrel, Eurasian Falcon (Hobby), Black Kite, and Long-legged buzzard of the Accipitridae (Hawk) Family	Gülseren Kırbaş Doğan	Turkey	Submission 653
24-28	Technical Analysis of a 19 kWp Photovoltaic Installation for Energy Supply: The SADEG Souk-Ahras Case	Naziha Zerari	Algeria	Submission 657
29-33	3D Baskılı PLA/İletken PLA Çekirdek-Kabuk Kompozitlerin Eğilme ve Kırılma Davranışına Isıl İşlemin Etkisi	Dursun Ekmekci, Emrah Kaplan	Turkey, Germany	Submission 674
34-49	Madencilikte Yorgunluk Tespiti ve Yönetimi	Zekeriya Duran, Bülent Erdem	Turkey, Turkey	Submission 676
50-55	Effect of Cutting Parameters on Temperature in Machining of AISI P20+S (DIN 1.2312) Steel	Muhammet Enver GÖKDEMİR, Yunus Kaplan, Tayfun ÇETİN, Mehmet TOPUZ	Turkey, Turkey, Turkey, Turkey	Submission 679
56-63	TEK SAĞLIK ALANINDAKİ ARAŞTIRMA MAKALELERİNİN BİBLİYOMETRİK ANALİZİ	Mustafa Nal, Veli Durmuş, İ. Serden Başak	Turkey, Turkey, Turkey	Submission 680
64-72	Goal-Conditioned Point Navigation of a Differential-Drive Robot Using Deep Deterministic Policy Gradient in MuJoCo	Ufuk Demircioğlu, Halit Bakır	Turkey, Turkey	Submission 683
73-82	Dynamic Stiffness Enhancement of Cross-Laminated Timber Observation Tower through MgO Surface Coating	Sertaç Tuhta	Turkey	Submission 686
83-89	Yeni İnşaat Birim Fiyat Pozlarının Metin Tabanlı Aile Sınıflandırılması: 2020–2026 Zamansal Genelleme Deneyi	Uğur Kafkas	Turkey	Submission 687
90-98	Assessing Long-Term Changes in Precipitation Frequency and Intensity Indices in Pınarbaşı (Kayseri, Türkiye)	Murat Karabulut	Turkey	Submission 688
99-109	Assessment of Modal Behaviour in SiO_2-Coated Glued Laminated Timber Pole Shed via Finite Element Simulation	Sertaç Tuhta	Turkey	Submission 689
110-117	Performance Evaluation of Lightweight 1D-CNN Architectures for Real-Time ECG Anomaly Detection in Edge Devices	Ahmet Akkaya, Mustafa Taşcı	Turkey, Turkey	Submission 690
118-125	Metric Choice Reorders System Rankings in Turkish Offensive-Language Detection: A 25-Configuration Study	Ahmet Akkaya, Muhammet Arucu	Turkey, Turkey	Submission 691
126-132	EFFECTS OF APRICOT KERNEL, BIOCHAR, HUMIC ACID AND POTASSIUM HUMATE ON SOIL C-N DYNAMICS, MICROBIAL BIOMASS AND CO_2 EMISSIONS	Asuman Büyükkılıç Yanardağ, İbrahim Halil Yanardağ, Abdullah Atum, Erdal Sakin, Hakan Yıldırım, Miraç Kılıç	Turkey, Turkey, Turkey, Turkey, Turkey, Turkey	Submission 692
133-138	Comparative Assessment of the Photon Shielding Characteristics of $MgZrO_3$ and $BaZrO_3$ Ceramics	Emine Narmanlı Han	Turkey	Submission 694
139-145	Early Detection of Smoking Habit Using Machine Learning with Biometric Parameters: A Comparative Model Analysis	Halit Çetiner	Turkey	Submission 697
146-157	Capsaicin'in Farmakolojik Özellikleri ve Terapötik Potansiyeli	Serife Tutuncu	Turkey	Submission 698
158-165	HALK SAĞLIĞI POLİTİKALARI KONUSUNU İNCELEYEN ARAŞTIRMA MAKALELERİNİN BİBLİYOMETRİK ANALİZİ	Mustafa Nal, İ. Serden Başak, Veli Durmuş	Turkey, Turkey, Turkey	Submission 707
166-172	Machine and Deep Learning in Mechanics of Materials: The Evaluation Protocol Governs Reported Accuracy in Alloy Yield Strength Prediction	Hamdi Kuleyin, Erkan Caner Ozkat	Turkey, Turkey	Submission 710
173-179	Comparative Performance Analysis of PID and Active Disturbance Rejection Control	Mehmet Latif LEVENT, Halit HÜLAKO, Sertaç YAMAN	Turkey, Turkey, Turkey	Submission 714

	for a Nonlinear Magnetic Levitation System			
180-189	Retrieval-Augmented Forecasting: A Structured Review of Time-Series Foundation Models, Retrieval Strategies, and Application Domains	Ali YILDIZ, Ali AKDAĞLI	Turkey, Turkey	Submission 717
190-196	Liquid–Liquid Extraction of Chicago Sky Blue Using a Diamide–Diether Calix [4] arene Receptor	İbrahim Uyanık	Turkey	Submission 719
197-203	Land Surface Temperature as a Landslide Conditioning Factor: Mechanisms, Applications, and Implications for Susceptibility Assessment	Seda Çellek	Turkey	Submission 720
204-213	Solar Radiation As A Landslide Conditioning Factor: A Literature Review	Seda Çellek	Turkey	Submission 721
214-217	Flow Balancer Machine Design and Manufacturing	Mohamad Ramzı ABOMAGHARA, İlhan ASİLTÜRK	Turkey, Turkey	Submission 722
218-230	Individual and Household Determinants of Female Labour Force Participation. Micro-Level Evidence from the Western Balkans	Loredana Sulejmani, Elena Polo	Albania, Albania	Submission 723
231-243	Application of the Neuman Systems Model in Community Health Services: A Systematic Review	Jonathan Niciza	Burundi	Submission 728
244-250	Provision of Reactive Power Demand Using Solar Photovoltaic Systems	Mert KAYA, Talha Enes GUMUS	Turkey, Turkey	Submission 730
251-255	Sosyal Bilgiler Öğretmenlerinin Dijital Öğretim Yeterlikleri ile Dijital Materyal Tasarımı Yeterlikleri Arasındaki İlişki	Şenol Mail Pala, Hüseyin Bayram	Turkey, Turkey	Submission 732
256-265	Analytical, Bayesian, and Machine-Learning Inference of Radioactive Decay-Chain Parameters: A Reproducible Accuracy–Speed–Uncertainty Benchmark on the Mo-99/Tc-99m System	Cafer Mert Yeşilkanat, Serkan Akkoyun	Turkey, Turkey	Submission 733
266-273	Yakup Kadri Karaosmanoğlu'nun Yaban Romanında Kadın ve Çocuk	Havvaana KARADENİZ	Turkey	Submission 739
274-281	Ortaokul Türkçe Ders Kitaplarındaki Şiirlerin Çocuk Edebiyatı ve Türkçe Öğretimi Açısından İncelenmesi	Havvaana KARADENİZ, Merve MORKOYUNLU	Turkey, Turkey	Submission 740
282-290	Eğitimde Üretken Yapay Zekâ Araştırmaları Yeniden Üretilebilir mi? Raporlama Şeffaflığının İçerik Analizi (2023–2026)	Şener Balat	Turkey	Submission 741
291-298	Transformer Models for Food Safety Inspection Grade Prediction	Ömer Faruk Sarı	Turkey	Submission 743
299-301	A Note on Optimal Control Theory	Muzaffer ATEŞ, Muhammet ATEŞ	Turkey, Turkey	Submission 757
302-305	Dinamik Programlama Yaklaşımı ile Kontrol Problemlerinin Optimizasyonu	Muzaffer Ateş, Hakan Alan	Turkey, Turkey	Submission 758
306-312	Elektroensefalografi Tabanlı Anestezi Derinliği Sınıflandırmasında Zamansal Modellemenin Etkisi	Şule Duyar, Yücel Koçyiğit	Turkey, Turkey	Submission 764
313-320	Agreement Without Convergence? Constructive Ambiguity, Transatlantic Coherence, and Implementation in the Kosovo–Serbia Dialogue	Leonora Bajrami	North Macedonia	Submission 768
321-329	Enhanced Speed Control of DTC based PMSM using Super Twisting Sliding Mode Controller	Mohamed Bashir BIHI, Hilmi AYGÜN, Yunus Emre KARASU	Turkey, Turkey, Turkey	Submission 775
330-333	Osman Nuri Ergin as a Bureaucrat of Transition and the Rationalization of Local Government in Turkey	Mustafa Demirkol	Turkey	Submission 776
334-342	Mardin'de Özel Mülkiyete Ait Bir Bina Örneği Üzerinden Konutta Müşterek Alanların Tanımı Ve Hesabı	Rojat AKSOY IŞIK, Cuvan YILDIRIM	Turkey, Turkey	Submission 778
343-354	Uluslararası Karayolu Taşımacılığında Sınır Geçişlerinden Kaynaklanan Riskler ve Operasyonel Etkileri	Muhammed Turgut	Turkey	Submission 783
355-359	Theoretical investigation and molecular docking analysis of a phenolic Schiff base targeting acetylcholinesterase	Sevgi Kansız	Turkey	Submission 784
360-364	DFT, molecular docking and ADME studies of a Schiff base compound	Sevgi Kansız	Turkey	Submission 785
365-376	Biopolymer Additives for Biological Soil Stabilization in Geotechnical Engineering: A Critical Assessment of MICP Mechanisms, Performance, and Applications	Bayram Ateş	Turkey	Submission 786
377-383	Design of a Notched Stepped Antenna with DGS-Enhanced Bandwidth	Burak Dokmetas	Turkey	Submission 787
384-391	Membranlı Süreçlerin Ağır Metal İçeren Atıksuların Arıtımında Kullanımı: Güncel	Hatice DEMİR BECEREKLİ, Yalçın ALTUNKAYNAK, Mutlu	Turkey, Turkey, Turkey	Submission 789

	Gelişmeler ve Gelecek Perspektifleri	CANPOLAT		
392-403	Biopolymer Additives for Biological Soil Stabilization: Artificial Neural Network Prediction of Unconfined Compressive Strength	Bayram Ateş, Mohammad Azim Eirgash	Turkey, Turkey	Submission 794
404-411	Gyroid ve Split-P TPMS Yapılarının Geometrik Tasarım Kriterlerinin Laminar Akış Karakteristikleri ve Sayısal Analiz Sonuçları Üzerindeki Etkilerinin İncelenmesi	İsmail Tur, Erkan Bahçe	Turkey, Turkey	Submission 797
412-421	Predicting Urban Traffic Congestion from Pre-Event Contextual Features: A Comparative Evaluation of Classical Machine Learning Algorithms	Cafer Yazicioglu, Ammar Aslan	Turkey, Turkey	Submission 802
422-435	Building Extraction from Remote Sensing Imagery: Quantifying Label Efficiency and the Effect of Transfer Learning on Annotation Cost	Ammar Aslan, Cafer Yazicioglu	Turkey, Turkey	Submission 804
436-446	Türkiye'de Elektrik Enerjisi Göstergelerinin Holt-D Modeliyle Tahmini	Agah Başdeğirmen, Ali Şimşek	Turkey, Turkey	Submission 805
447-455	Yapay Zekâ Kavramının Tarihsel ve Kavramsal Dönüşümü	Ali Şimşek, Agah Başdeğirmen	Turkey, Turkey	Submission 806
456-465	Decoupling the Effects of Water to Cement Ratio and Firing Temperature on the Mechanical Performance of Vibration Cast Calcium Aluminate Cement Bodies	Ali KAFALI, Safa POLAT	Turkey, Turkey	Submission 812
466-477	KÜLTÜREL MİRAS ALANLARINDA ZİYARETÇİ DENEYİMLERİ	Ozan Çatır	Turkey	Submission 813

COMMITTEE

ORGANIZING COMMITTEES

Assoc. Prof. Admir Jance, European University of Tirana
Assoc. Prof. Elif Küçük Durur, Atatürk University
Assoc. Prof. Nermin Kişi, Zonguldak Bülent Ecevit University
Assoc. Prof. Ülhak Çimen, Atatürk University
Dr. Krisztina Czakoova, J. Selye University
Dr. Ondrej Takac, J. Selye University
Dr. Ladislav Vegh, J. Selye University

SCIENTIFIC COMMITTES

Prof. Dr. Abdurazzag Ali Aburas, University of Kwazulu-Natal
Prof. Dr. Adnan Sözen, Gazi University
Prof. Dr. Amit Soni, Manıpal University
Prof. Dr. Anis Koubaa, Prince Sultan University
Prof. Dr. Carlos A. Coelho, Universidade Nova de Lisboa
Prof. Dr. Josep Guerrero, Aalborg University
Prof. Dr. Juan Vasquez, Aalborg University
Prof. Dr. Laszlo T. Koczy, Budapest University of Technology and Economics
Prof. Dr. Mohammad Ashraf Gondal, King Fahd University
Prof. Dr. Murat Yakar, Mersin University, Turkey
Prof. Dr. Muhammad Rashid, University of West Florida
Prof. Dr. Saad Mekhilef, Malaya University
Prof. Dr. Salim Ibrir, King Fahd University
Prof. Dr. Subarna Shakya, Tribhuvan University
Prof. Dr. Suzuki Hisashi, Chuo University
Prof. Dr. Usha Mehta, Institute of Technology, Nirma University
Prof. Dr. Vagif Gasimov, Azerbaijan Technical University
Prof. Dr. Weidong Zhu, University of Maryland
Assoc. Prof. Ali Hakan Işık, Burdur Mehmet Akif Ersoy University
Assoc. Prof. Amita Nandal, Manıpal University
Assoc. Prof. Aslan İnan, Yıldız Teknik University
Assoc. Prof. Atif Iqbal, Qatar University
Assoc. Prof. Elif Küçük Durur, Atatürk University
Assoc. Prof. Francesco Cottone, University of Perugia
Assoc. Prof. Farid Melgani, University of Trento
Assoc. Prof. Halit Öztekin, Sakarya Applied Science University
Assoc. Prof. Hamdy Ziedan, Assiut University
Assoc. Prof. Harun Kemal Öztürk, Pamukkale University
Assoc. Prof. Jiri Pela, Brno University of Technology
Assoc. Prof. Merih Palandöken, İzmir Katip Çelebi University
Assoc. Prof. Nermin Kişi, Zonguldak Bülent Ecevit University
Assoc. Prof. Rita Ismailova, Kyrgyz Turkish Manas University
Assoc. Prof. Sadia Samar Ali, King Abdul-Aziz University
Assoc. Prof. Ülhak Çimen, Atatürk University
Assoc. Prof. Tanmay De, National Institute of Technology Durgapur
Assoc. Prof. Zeashan Khan, Air University
Dr. Ali Almisreb, International University of Sarajevo
Dr. Amina Othmani, University of Monastir-Tunisia
Dr. Arti Jain, Jaypee Institute of Information Technology
Dr. Arun Kumar Verma, Malaviya National Institute of Technology Jaipur
Dr. Aziz Shah, University of Glaskow
Dr. Bal Chand Nagar, ECE National Institute of Technology Patna
Dr. Baoze Wei, Aalborg University
Dr. Bohumil Brtník, University of Pardubice
Dr. Diptiranjan Behera, The University of the West Indies at Mona
Dr. Fayçal BENSAALI, Qatar University
Dr. Fugui Qi, Fourth Military Medical University
Dr. Magdi S. Mahmoud, King Fahd University
Dr. Mariya Ouaissa, Moulay Ismail University
Dr. Merdaci Slimane, University of -Sidi Bel Abbés, Algeria
Dr. Moharram Challenger, Antwerp University
Dr. Muhammad Sultan, Zakariya University
Dr. Lesya Anishchenko, Bauman Moscow State Technical University
Dr. R.A. Ilyas, Universiti Teknologi Malaysia
Dr. Sadina Gagula Palalić, International University of Sarajevo
Dr. Stefania Tomasiello, University of Tartu
Dr. Syed Omer Gilani, National University of Sciences and Technology (NUST)
Dr. Tarık Namas, International University of Sarajevo
Dr. Uduak A. Umoh, University of Uyo
Dr. Yajuan Guan, Aalborg University

25.08.2026

16-17 Ağustos 2026 tarihlerinde MEET üzerinden çevrimiçi olarak gerçekleştirilen 6th International Conference on Modern and Advanced Research ICMAR 2026 konferansı akademik teşvik yönetmeliğinin 9. Maddesine istinaden "Tebliğlerin sunulduğu yurt içinde veya yurt dışındaki etkinliğin uluslararası olarak nitelendirilebilmesi için Türkiye dışında en az beş farklı ülkeden sözlü tebliğ sunan konuşmacının katılım sağlaması ve tebliğlerin yarıdan fazlasının Türkiye dışından katılımcılar tarafından sunulması esastır." kriterlerini sağlamaktadır. Toplam 146 adet bildirinin yer aldığı kongre iki gün boyunca çevrimiçi olarak gerçekleştirilmiştir.

Türkiye dışından toplam 17 farklı ülkeden (Fas, Cezayir, Pakistan, Arnavutluk, Birleşik Krallık, Hindistan, Kosova, Almanya, Azerbaycan, Libya, Nijerya, Irak, Burundi, İtalya, Kuzey Makedonya, Sırbistan, Bangladeş) katılım sağlanmış olup, 146 adet bildirinin 81 (%55,47) tanesi yabancı katılımcı tarafından sunulmuştur.

Kongremize ilginiz için teşekkür ederiz.

Saygılarımızla,

25.08.2026

16-17 Ağustos 2026 tarihlerinde MEET üzerinden çevrimiçi olarak gerçekleştirilen 6th International Conference on Modern and Advanced Research ICMAR 2026 Konferansı, 2025 MART DÖNEMİ DOÇENTLİK BAŞVURU ŞARTLARI yönetmeliğinin 2. bölümde yer alan Uluslararası Bilimsel Toplantı tanımlaması "Farklı ülkelerden bilim insanlarının bilim kurulunda bulunduğu ve sunumların bilimsel ön incelemeden geçirilerek kabul edildiği toplantı." şeklindedir.

6TH ICMAR 2026 konferansı bilimsel komitesinin çoğunluğu yabancı bilim insanlarından oluşmaktadır. Aşağıdaki link ile bilimsel komite incelenebilir:

https://as-proceeding.com/index.php/icmar/committe

Ayrıca konferansımızda hakem değerlendirmesi olup hakem değerlendirmesi olumlu sonuçlanmadan ücret talebinde bulunulmamaktadır.

2024 MART DÖNEMİ DOÇENTLİK BAŞVURU ŞARTLARI yönetmeliğinin 8. bölümde yer alan Bilimsel Toplantı tanımlamasının b bendindeki "Diğer Uluslararası/Ulusal bilimsel toplantının düzenlenme komitesinde, kurum/tüzel kişilik/karar organı tarafından resmi olarak görevlendirilmiş üniversite/enstitü/bilimsel kurum/branş derneği akademisyen temsilcisi bulunması zorunludur." şeklindedir.

6TH ICMAR 2026 konferansı düzenleme komitesinde farklı kurumlar tarafından görevlendirilmiş birden fazla üniversiteden akademik temsilciler bulunmaktadır. Aşağıdaki link ile düzenleme komitesindeki akademik temsilcilerin görevlendirme yazıları incelenebilir:

https://drive.google.com/file/d/1eZalxRl8q6-eXATZzwUraH9KOhtQzFho/view

Ayrıca görevlendirme yazıları ilgili konferansın özet ve tam metin kitapçıklarında yer almaktadır.

Kongremize ilginiz için teşekkür ederiz.

Saygılarımızla,



All Sciences Academy

04.09.2023

LETTER OF ASSIGNMENT

We confirm that Ing. Ondrej Takáč, PhD. working at J. Selye University takes part in the organization committee of the conferences organized by of All Sciences Academy.

Confirmation of the Vice Dean of the Faculty of Economics and Informatics

Name of Responsible Person and Signature:

Name and Surname Sign:

PhDr. Imrich Antalík, PhD. _____

Date: 04.09.2023

UNIVERZITA J. SELYEHO
SELYE JÁNOS EGYETEM
FAKULTA EKONÓMIE A INFORMATIKY
DEKANÁT
Bratislavská 167/21, 945 01 Komárno

04.09.2023

LETTER OF ASSIGNMENT

We confirm that PaedDr. Ladislav Végh, PhD. working at J. Selye University takes part in the organization committee of the conferences organized by of All Sciences Academy.

Confirmation of the Vice Dean of the Faculty of Economics and Informatics

Name of Responsible Person and Signature:

Name and Surname Sign:

PhDr. Imrich Antalík, PhD. _____

Date: 04.09.2023

Universiteti Europian i Tiranës

16.08.2023

LETTER OF ASSIGNMENT

We confirm that Prof. Assoc. Dr. Admir Jançe, who is employed full-time as a lecturer, as well as holding the position of Head of the Imaging Department and Vice Dean of the Faculty of Technical Medical Sciences, takes part in the organization committee of the conferences organized by of All Sciences Academy.

Confirmation of the Rector/Dean/Vice Dean/Head of Department

Name of Responsible Person and Signature:



Rector: Prof. Dr. Tonin GJURAJ

_ _

Date: 16.08.2023

15.09.2023

İlgili Makama

Atatürk Üniversitesi İletişim Fakültesinde görev yapan Doçent Doktor Ülhak Çimen'in All Sciences Academy yayınevi ile beraber düzenlenmesi planlanan konferanslarda ve organizasyonlarda düzenleme veya bilim kurullarında yer alması uygun görülmüştür.

Gereğini rica ederim.

Ad ve Soyad Prof. Dr. Adem YILMAZ Dekan Tarih 15.09.2023	İmza: Prof.Dr. Adem YILMAZ DEKAN

15.09.2023

İlgili Makama

Atatürk Üniversitesi İletişim Fakültesinde görev yapan Doçent Doktor Elif Küçük DURUR'un All Sciences Academy yayınevi ile beraber düzenlenmesi planlanan konferanslarda ve organizasyonlarda düzenleme veya bilim kurullarında yer alması uygun görülmüştür.

Gereğini rica ederim.

Ad ve Soyad Prof. Dr. Adem YILMAZ Dekan Tarih 15.09.2023	İmza: Prof.Dr. Adem YILMAZ DEKAN

04.09.2023

LETTER OF ASSIGNMENT

We confirm that PaedDr. Krisztina Czakóová, PhD. working at J. Selye University takes part in the organization committee of the conferences organized by of All Sciences Academy.

Confirmation of the Vice Dean of the Faculty of Economics and Informatics

Name of Responsible Person and Signature:

Name and Surname Sign: _____

PhDr. Imrich Antalík, PhD.

Date: 04.09.2023

T.C.
ZONGULDAK BÜLENT ECEVİT ÜNİVERSİTESİ REKTÖRLÜĞÜ
Çaycuma Meslek Yüksekokulu Müdürlüğü

Sayı : E-84604848-903.07.04-371012 01.11.2023
Konu : Bilimsel Toplantı Görevlendirme

İLGİLİ MAKAMA

İlgi : Öğretim Üyesi Doç. Dr. Nermin KİŞİ'nin 30.10.2023 tarihli dilekçesi.

Yüksekokulumuz Yönetim ve Organizasyon Bölümü Öğretim Üyesi Doç. Dr. Nermin KİŞİ'nin "All Sciences Academy Yayınevi"nin düzenleyeceği bilimsel toplantıların düzenleme ve bilim kurullarında görev almasında Yüksekokulumuzca bir sakınca bulunmamaktadır. İş bu belge ilgilinin isteği üzerine tanzim edilmiştir.

Bilgilerinizi ve gereğini rica/arz ederim.

Doç. Dr. Hikmet YAZICI
Meslek Yüksekokulu Müdürü

Bu belge, güvenli elektronik imza ile imzalanmıştır.

Adres : BEÜ Çaycuma Meslek Yüksekokulu, 67900, Çaycuma, Zonguldak
Telefon : (0372) 643 66 01 Faks:(0372) 643 66 04
e-Posta : caycumamyo@beun.edu.tr Web : https://cmyo.beun.edu.tr/
Kep Adresi : beun@hs03.kep.tr
Bilgi için : Tülay KEÇECİ
Unvanı : İşçi

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://www.icfarconf.com/

Molecular Docking Analysis of Compounds Isolated from Marine Sponges Targeting EGFR

Adnan Cetin*[1]

[1]*Department of Chemistry, University of Van Yüzüncü Yil, Türkiye*
[*]*(adnancetin@yyu.edu.tr)*

Abstract – Epidermal Growth Factor Receptor (EGFR) is a receptor tyrosine kinase that plays a critical role in the regulation of cell proliferation, survival, and differentiation and is considered an important molecular target in cancer research. Marine sponges are recognized as valuable sources of structurally diverse and biologically active natural products with potential therapeutic properties. In the present study, four compounds isolated from marine sponges were investigated for their potential interaction with EGFR using molecular docking analysis. The three-dimensional structure of EGFR was used as the target protein, and the binding modes and molecular interactions of the selected compounds within the active site were evaluated in detail. Docking results were assessed based on binding affinities and key interactions with amino acid residues in the EGFR binding pocket, including hydrogen bonding and hydrophobic or π-related interactions where applicable. The obtained findings provided insights into the structural basis of ligand–EGFR recognition and identified compounds with promising binding profiles toward the target protein. Overall, the molecular docking results suggest that the investigated marine sponge-derived compounds may represent potential EGFR inhibitor candidates and provide a basis for further experimental evaluation of their biological activities.

Keywords – EGFR; Molecular Docking; Marine Sponges; Natural Products; EGFR Inhibitors; Ligand–Protein Interactions; Cancer Therapy

I. INTRODUCTION

Cancer is one of the major health problems worldwide and is characterized by uncontrolled cell proliferation, resistance to cell death, invasion, and metastatic progression [1, 2]. The development of effective anticancer agents requires the identification of molecular targets that play essential roles in tumor initiation and progression. Among these targets, receptor tyrosine kinases have attracted considerable attention because of their important functions in cellular signaling pathways involved in proliferation, differentiation, migration, and survival. The EGFR, also known as ErbB1/HER1, is a transmembrane receptor tyrosine kinase belonging to the ErbB receptor family. EGFR is activated following the binding of specific ligands, resulting in receptor dimerization and autophosphorylation of intracellular tyrosine residues. Subsequently, several downstream signaling pathways, including the RAS/RAF/MEK/ERK and PI3K/AKT/mTOR pathways, are activated [3, 4]. These signaling cascades regulate essential cellular processes such as proliferation, survival, differentiation, and migration. Aberrant EGFR activation or overexpression has been associated with the development and progression of several human cancers. Therefore, EGFR has become an important target in the discovery and development of anticancer agents. Several small-molecule EGFR inhibitors have been developed to interfere with EGFR-mediated signaling [5]. However, the emergence of resistance mechanisms and alterations in the EGFR signaling network can reduce the therapeutic effectiveness of existing inhibitors. Consequently, the discovery of structurally diverse compounds capable of interacting effectively with EGFR remains an important objective in

medicinal chemistry and drug discovery. In this context, naturally occurring compounds have attracted considerable interest because of their structural diversity and broad range of biological activities [6].

Marine organisms represent a particularly rich source of structurally novel secondary metabolites. Among marine organisms, sponges are considered prolific producers of bioactive natural products, including alkaloids, terpenoids, steroids, peptides, polyketides, and other chemically diverse metabolites [7, 8]. Compounds isolated from marine sponges have demonstrated a wide range of biological activities, including anticancer, antimicrobial, anti-inflammatory, antiviral, and enzyme-inhibitory properties. The unique chemical diversity of marine sponge-derived metabolites therefore provides an attractive starting point for the identification of new bioactive molecules and potential therapeutic agents. Computational approaches have become increasingly important in modern drug discovery.

Molecular docking is a widely used structure-based computational technique that predicts the preferred orientation of a ligand within the binding site of a target protein and provides an estimation of ligand–protein binding affinity [9]. In addition, docking analysis can provide valuable information regarding the molecular interactions responsible for ligand recognition, including hydrogen bonds, hydrophobic interactions, π–π interactions, and other non-covalent contacts. Such information can contribute to the understanding of the potential inhibitory activity of natural products and assist in the prioritization of promising compounds for further experimental investigation. In the present study, four compounds isolated from marine sponges were investigated against EGFR using molecular docking analysis [10]. The binding modes of the selected compounds within the EGFR binding site were examined in detail, and their predicted binding affinities and interactions with key amino acid residues were evaluated. The aim of this study was to explore the potential of these marine sponge-derived Seragamid B (SerB), Seragamid C (SerC), Seragamid D (SerD), and Seragamid E (SerE) compounds as EGFR inhibitor candidates and to provide structural insights into their interactions with the target protein in Fig 1. The findings may contribute to the identification of promising natural-product scaffolds for further investigation in EGFR-targeted anticancer drug discovery.

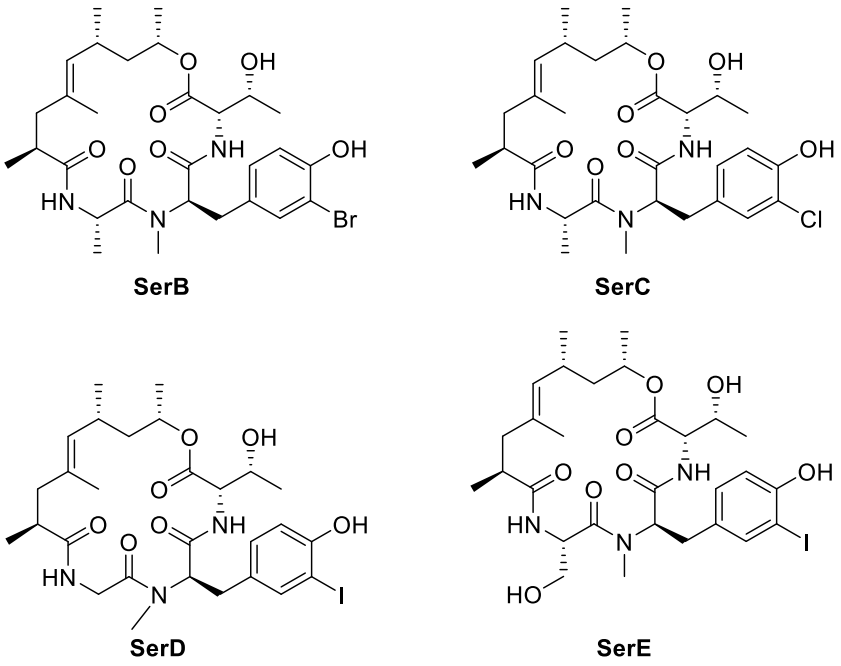


Fig. 1 Selected Marines Sponges

II. Materials and Method

Molecular docking simulations were carried out using AutoDockTools version 1.5.7 [11] to investigate the binding characteristics of the selected compounds toward the target protein. The resulting ligand–protein complexes were subsequently analyzed using BIOVIA Discovery Studio [12] to characterize the

binding poses, intermolecular interaction patterns, and distances between the ligand and surrounding amino acid residues. The crystal structure of the CDK8/Cyclin C complex, co-crystallized with N-(4-chlorobenzyl)isoquinolin-4-amine, was retrieved from the Protein Data Bank (PDB ID: 6T41; resolution: 2.45 Å) and employed as the receptor model [13]. Before docking calculations, the protein structure was prepared by eliminating crystallographic water molecules and the co-crystallized ligand to prevent unwanted interactions during the docking procedure. Structural deficiencies, including missing atoms and incomplete side-chain residues, were corrected using appropriate structural modeling procedures. The resulting protein model was further optimized by assigning suitable protonation states and performing energy minimization to obtain a stable and docking-compatible receptor conformation. The docking protocol was subsequently validated by comparing the experimentally observed orientation of the reference ligand within the binding site with the corresponding pose predicted by the docking procedure. Structural superposition of the experimental and predicted conformations was used to evaluate the reliability of the selected docking parameters and the ability of the protocol to reproduce the experimentally determined binding mode. For the docking calculations, a grid box covering the complete protein structure was generated to allow the ligands to explore the accessible binding regions of the receptor. This comprehensive search strategy enabled the identification of potential binding pockets and the evaluation of the preferred ligand orientations within the target protein. The resulting docking poses were examined in detail to identify relevant non-covalent interactions, including hydrogen bonds, hydrophobic contacts, and aromatic interactions, which may contribute to ligand stabilization within the binding site.

III. RESULTS AND DISCUSSION

According to the molecular docking results shown in Fig 2, SerB binds to the active site of the EGFR protein with a binding energy of −8.8 kcal/mol. This binding energy indicates a good level of binding affinity for small-molecule–protein interactions and demonstrates that SerB can form a stable complex in the EGFR binding pocket. In general, binding energies below −7.0 kcal/mol are considered moderate, while values of −8.0 kcal/mol and lower are regarded as indicative of a strong binding tendency. Therefore, the obtained value of −8.8 kcal/mol is a significant finding supporting the compound's potential as an EGFR inhibitor. The docking analysis reveals that SerB primarily forms π-π stacking, π-cation, π-alkyl, and van der Waals interactions within the binding pocket. In particular, the π-π stacking interaction observed between the aromatic ring system and the Phe100 amino acid significantly contributes to the stability of the complex.

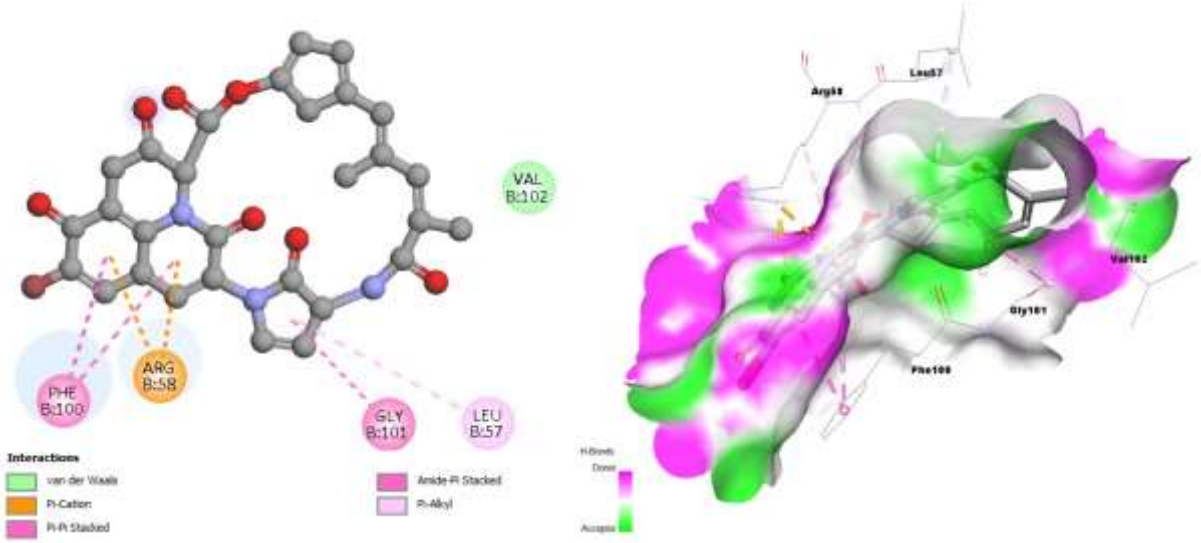


Fig. 2 Molecular docking interaction of SerB within the binding site of the target protein

Similarly, the π-cation interaction formed with the Arg58 residue enhances binding affinity due to the electrostatic attraction between the aromatic electron cloud and the positively charged guanidinium group. Furthermore, the occurrence of amide-π stacking with Gly101 and π-alkyl interactions with Leu57 indicates that the ligand is appropriately positioned in the hydrophobic region. In addition, the van der Waals interactions formed with the Val102 residue demonstrate that the ligand fits well geometrically into the binding pocket and supports the stability of the complex. Although no hydrogen bonds are observed in the interaction map, the presence of numerous hydrophobic and aromatic interactions together helps maintain the complex's stability. In particular, π-π stacking and π-cation interactions are among the types of interactions that play a significant role in the binding of many tyrosine kinase inhibitors to EGFR [14]. The fact that SerB exhibits a similar interaction profile suggests that the compound can effectively occupy the active site of EGFR and may have the potential to inhibit the receptor's catalytic function.

According to the results of the molecular docking analysis as shown in Fig 3, SerC binds to the active binding pocket of the EGFR protein with a binding energy of −8.8 kcal/mol. This binding energy indicates that a complex with high binding affinity and thermodynamic stability is formed between the ligand and the target protein. Generally, binding energies of −8.0 kcal/mol and lower indicate that small molecules tend to bind strongly to target proteins; therefore, it can be concluded that SerC may exhibit potential inhibitory properties against EGFR. An analysis of the interaction network shows that SerC is stabilized in the EGFR active site through numerous distinct interaction mechanisms. The compound forms a conventional hydrogen bond with the Asp2 amino acid, thereby enhancing the specificity of the binding and the stability of the complex. In addition, the carbon-hydrogen bond observed between one of the carbonyl oxygens and the Lys1 amino acid, while not as strong as a hydrogen bond, contributes to maintaining the ligand's position within the active site.

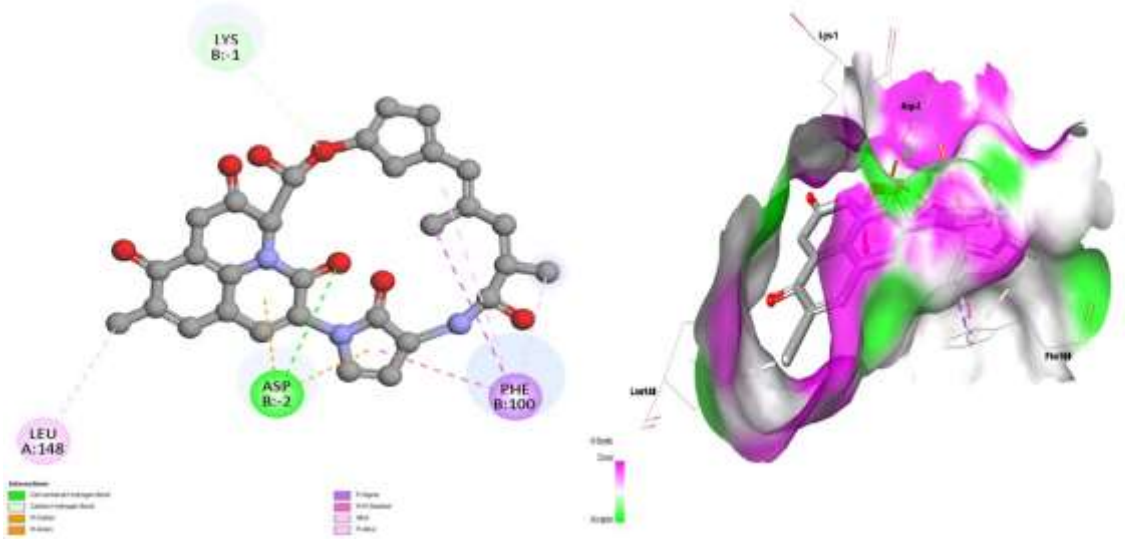


Fig. 3 Molecular docking interaction of SerC within the binding site of the target protein

The aromatic structure of SerC also gives rise to significant electrostatic and hydrophobic interactions. The π-anion interaction with Asp2 enhances binding affinity through the attraction between the negatively charged aspartate residue and the ligand's π-electron system. In addition, the π-cation interaction observed with Asp2 reinforces the network of electrostatic interactions in the active site. The presence of these two types of interactions around the same amino acid indicates that the ligand adopts a geometric arrangement suitable for the binding pocket [15]. Phe100, one of the aromatic amino acids in the binding pocket, forms multiple hydrophobic interactions with SerC. In particular, π-π stacking and π-sigma interactions significantly contribute to the stability of the complex through electron cloud interactions between the aromatic rings. Furthermore, the π-alkyl interactions formed with Phe100 and the alkyl interaction observed with the Leu148 amino acid indicate that the ligand is strongly retained in the hydrophobic region. These

hydrophobic contacts, along with hydrogen bonds, are key factors supporting the overall stability of the complex. It is observed that SerC forms a rich network of interactions in the EGFR active site, consisting of hydrogen bonds, carbon-hydrogen bonds, π-anion, π-cation, π-π stacking, π-sigma, π-alkyl, and alkyl interactions. This multifaceted interaction profile indicates the formation of a stable complex that is not based solely on hydrophobic contacts but is also reinforced by electrostatic and hydrogen-bond contributions. In particular, the interactions formed with the Asp2 and Phe100 residues play a decisive role in the strong binding of SerC to the EGFR active site.

According to the results of the molecular docking analysis as shown in Fig 4, SerD binds to the active binding site of the EGFR protein with a binding energy of −8.9 kcal/mol. This binding energy indicates that a high binding affinity exists between the ligand and the protein, resulting in a stable complex. The fact that SerD's binding energy is slightly lower (more negative) than that of Ser B and C suggests that the compound is capable of forming a stronger interaction with the EGFR active site. This indicates that SerD is a promising candidate molecule that could be evaluated as an EGFR inhibitor. An analysis of the docking results shows that SerD is stabilized in the EGFR active site by forming both hydrogen bonds and electrostatic interactions. One of the ligand's carbonyl oxygens forms a conventional hydrogen bond with the Leu179 amino acid, while the other carbonyl group forms a conventional hydrogen bond with the Ser62 amino acid. These hydrogen bonds help maintain the ligand's correct orientation within the active site, thereby enhancing binding specificity and the stability of the complex. It is well known that hydrogen bonds formed with serine residues play a significant role in the binding of tyrosine kinase inhibitors. In addition, SerD forms two π-cation interactions with the Arg178 amino acid in the aromatic ring system. This electrostatic attraction between the positively charged arginine side chain and the ligand's π-electron system is one of the interactions that significantly contributes to the improvement in binding energy [16]. Furthermore, the π-alkyl interaction observed in the hydrophobic region of the molecule indicates that the binding model is consistent with the hydrophobic character of the active site. These hydrophobic contacts, along with hydrogen bonds and π-cation interactions, are complementary interactions that support the stability of the complex. It appears that SerD forms two conventional hydrogen bonds, two π-cation interactions, and one π-alkyl interaction in the EGFR binding pocket. The hydrogen bonds determine the ligand's binding orientation, while the π-cation interactions enhance binding affinity by providing an electrostatic contribution, ensuring that SerD is stably retained in the active site. This interaction profile can be considered one of the primary reasons why the compound exhibits high binding energy despite making contact with a relatively small number of amino acids.

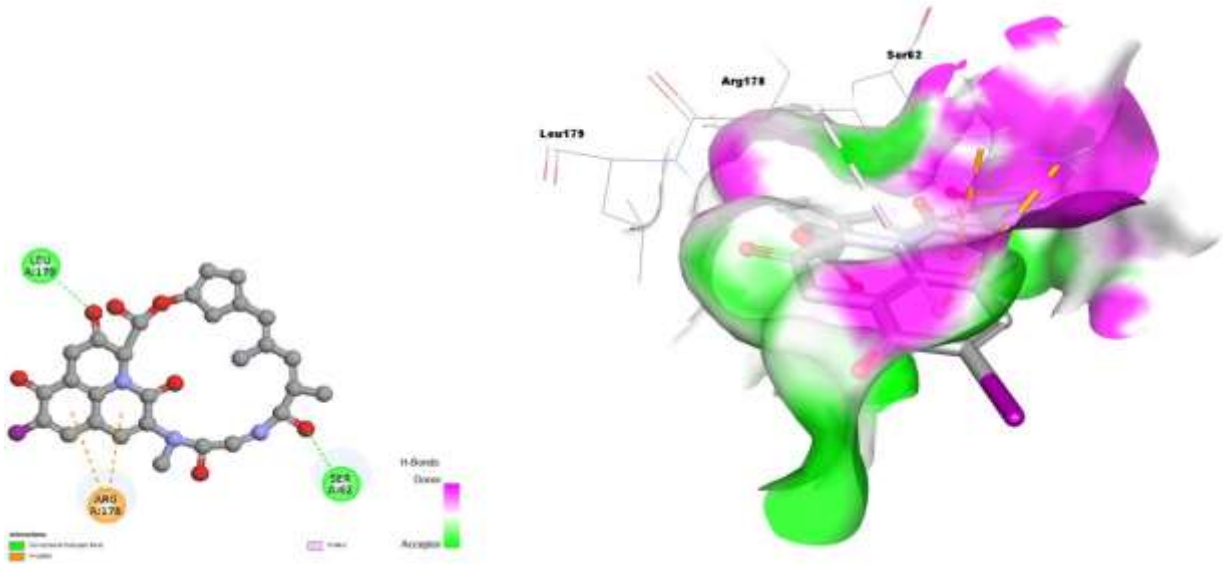


Fig. 4 Molecular docking interaction of SerD within the binding site of the target protein

SerE binds to the active binding pocket of the EGFR protein with a binding energy of −9.2 kcal/mol. This value represents the lowest (most negative) binding energy among the Seragamid derivatives studied (B, C, and D), indicating that SerE exhibits the highest binding affinity for the EGFR active site. In general, more negative binding energies indicate that the ligand-protein complex is more stable and that the binding process is thermodynamically more favorable. Therefore, it can be said that SerE is one of the most promising candidates for evaluation as an EGFR inhibitor. It appears that SerE forms a versatile and strong network of interactions within the EGFR active site as shown in Fig 5. The ligand forms conventional hydrogen bonds with the Asp2 and Phe100 amino acids via its carbonyl oxygens. In addition, the carbon-hydrogen bond formed between one of the ester oxygen atoms and the Lys1 residue, although weaker than a hydrogen bond, helps maintain the ligand's position within the binding pocket. This network of hydrogen bonds plays a key role in stabilizing SerE in the correct orientation within the active site. In addition, the aromatic system of SerE forms a π-anion interaction with the Asp2 amino acid. This electrostatic interaction between the negatively charged aspartate residue and the ligand's π-electron system is one of the key factors that increases binding affinity. Furthermore, the π-sigma interaction and two π-alkyl interactions formed with the Phe100 amino acid indicate that the ligand is strongly stabilized within the hydrophobic binding pocket. In particular, these multiple hydrophobic contacts established with the aromatic phenylalanine residue significantly contribute to the improvement of binding energy by maintaining the structural integrity of the complex. SerE forms two conventional hydrogen bonds, one carbon-hydrogen bond, one π-anion interaction, one π-sigma interaction, and two π-alkyl interactions in the active site of EGFR. While hydrogen bonds determine the ligand's binding orientation, electrostatic and hydrophobic interactions enhance the stability of the complex. In particular, the interactions formed with the amino acids Phe100, Asp2, and Lys1 stand out as the key molecular interactions that enable SerE to bind strongly and stably to the active site.

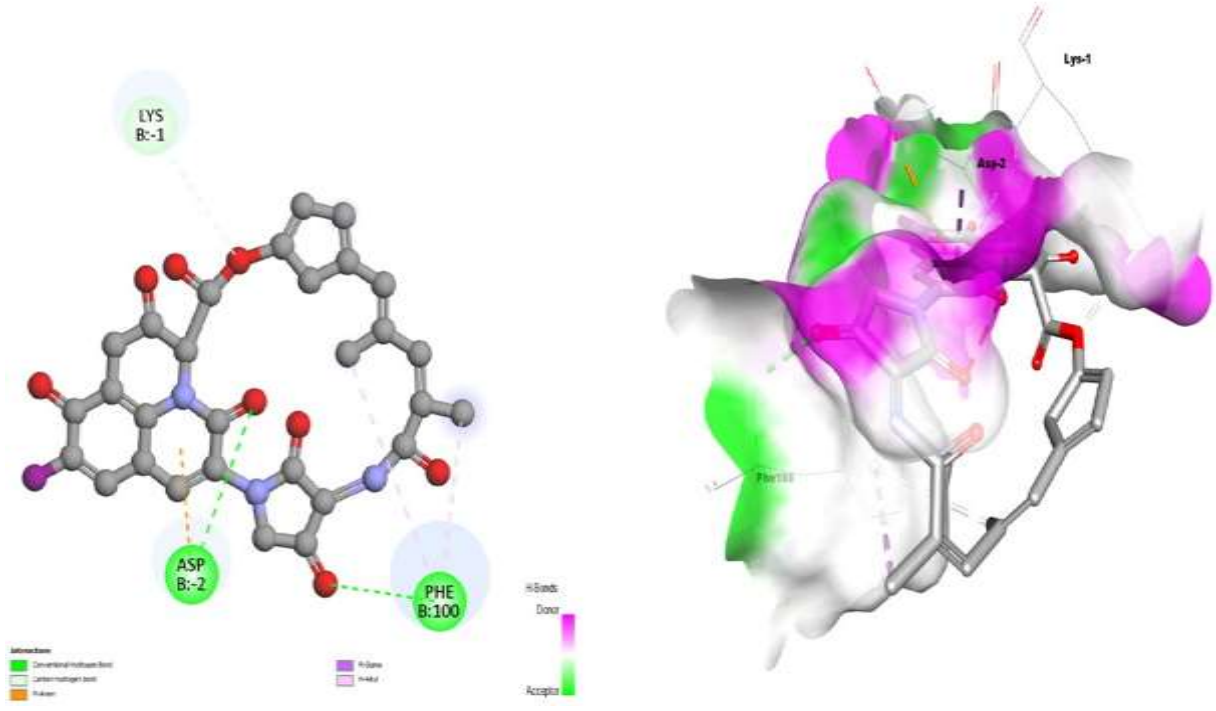


Fig. 5 Molecular docking interaction of SerE within the binding site of the target protein

IV. CONCLUSION

Molecular docking analysis demonstrated that all four investigated seragamid derivatives (SerB–SerE) exhibited favorable binding profiles toward the EGFR active site, with binding energies ranging from −8.8 to −9.2 kcal/mol. Among the evaluated compounds, SerE showed the strongest predicted binding affinity

(−9.2 kcal/mol), followed by SerD (−8.9 kcal/mol), SerB (−8.8 kcal/mol), and SerC (−8.8 kcal/mol). Detailed analysis of the ligand–protein complexes revealed that the compounds were stabilized within the EGFR binding pocket through a combination of hydrogen bonding, hydrophobic, aromatic, and electrostatic interactions with key amino acid residues. SerB displayed favorable hydrophobic and aromatic interactions involving Phe100, Arg58, Gly101, Leu57, and Val102, whereas SerC was stabilized by a combination of hydrogen-bonding, aromatic, and electrostatic interactions. SerD formed a stable interaction network involving Leu179, Ser62, and Arg178, particularly through hydrogen-bonding and electrostatic contacts. Notably, SerE exhibited the most favorable docking score and a diverse interaction profile, including hydrogen bonds, π-anion contacts, and aromatic hydrophobic interactions, suggesting a strong and complementary fit within the EGFR binding site. Overall, the docking results indicate that the investigated seragamid derivatives, particularly SerE, may represent promising molecular scaffolds for further investigation as potential EGFR-targeted anticancer agents. Nevertheless, molecular docking provides predictive rather than definitive evidence of biological activity. Therefore, further validation using molecular dynamics simulations, MM/PBSA or MM/GBSA binding free-energy calculations, EGFR enzyme inhibition assays, and appropriate cell-based biological studies will be necessary to confirm the stability, binding characteristics, and anticancer potential of these compounds.

REFERENCES

[1] Sigismund, S., Avanzato, D., & Lanzetti, L. (2018). Emerging functions of the EGFR in cancer. *Molecular oncology*, *12*(1), 3-20.

[2] Blackadar, C. B. (2016). Historical review of the causes of cancer. *World journal of clinical oncology*, *7*(1), 54.

[3] Sabbah, D. A., Hajjo, R., & Sweidan, K. (2020). Review on epidermal growth factor receptor (EGFR) structure, signaling pathways, interactions, and recent updates of EGFR inhibitors. *Current topics in medicinal chemistry*, *20*(10), 815-834.

[4] Yarden, Y. (2001). The EGFR family and its ligands in human cancer: signalling mechanisms and therapeutic opportunities. *European journal of cancer*, *37*, 3-8.

[5] Seshacharyulu, P., Ponnusamy, M. P., Haridas, D., Jain, M., Ganti, A. K., & Batra, S. K. (2012). Targeting the EGFR signaling pathway in cancer therapy. Expert opinion on therapeutic targets, 16(1), 15-31.

[6] Gan, Y., Shi, C., Inge, L., Hibner, M., Balducci, J., & Huang, Y. (2010). Differential roles of ERK and Akt pathways in regulation of EGFR-mediated signaling and motility in prostate cancer cells. *Oncogene*, *29*(35), 4947-4958.

[7] Laport, M. S., Santos, O. C. S., & Muricy, G. (2009). Marine sponges: potential sources of new antimicrobial drugs. *Current pharmaceutical biotechnology*, *10*(1), 86-105.

[8] Varijakzhan, D., Loh, J. Y., Yap, W. S., Yusoff, K., Seboussi, R., Lim, S. H. E., & Chong, C. M. (2021). Bioactive compounds from marine sponges: Fundamentals and applications. *Marine drugs*, *19*(5), 246.

[9] Abd El-Meguid, E. A., Moustafa, G. O., Awad, H. M., Zaki, E. R., & Nossier, E. S. (2021). Novel benzothiazole hybrids targeting EGFR: Design, synthesis, biological evaluation and molecular docking studies. *Journal of Molecular Structure*, *1240*, 130595.

[10] Liao, Q. H., Gao, Q. Z., Wei, J., & Chou, K. C. (2011). Docking and molecular dynamics study on the inhibitory activity of novel inhibitors on epidermal growth factor receptor (EGFR). *Medicinal Chemistry*, *7*(1), 24-31.

[11] Huey, R., Morris, G. M., & Forli, S. (2011). Using AutoDock 4 and Vina with autodocktools: A tutorial. Scripps Research Institute, California, USA, 46.

[12] Baroroh, U., Biotek, M., Muscifa, Z. S., Destiarani, W., Rohmatullah, F. G., & Yusuf, M. (2023). Molecular interaction analysis and visualization of protein-ligand docking using Biovia Discovery Studio Visualizer. *Indonesian Journal of Computational Biology (IJCB)*, *2*(1), 22-30.

[13] Klatt, F., Leitner, A., Kim, I. V., Ho-Xuan, H., Schneider, E. V., Langhammer, F., & Kuhn, C. D. (2020). A precisely positioned MED12 activation helix stimulates CDK8 kinase activity. *Proceedings of the National Academy of Sciences*, *117*(6), 2894-2905.

[14] Hevener, K. E., Zhao, W., Ball, D. M., Babaoglu, K., Qi, J., White, S. W., & Lee, R. E. (2009). Validation of molecular docking programs for virtual screening against dihydropteroate synthase. *Journal of chemical information and modeling*, *49*(2), 444-460.

[15] Asadian, M., & Kianianbouhi, R. (2025). γ-Cyclodextrin Inclusion Complexes with Macrolide Antibiotics by Molecular Docking & Comparing their Interactions with the Highest and the Lowest Binding Affinity. *Adv. J. Chem. A*, *8*(9), 1442-1456.

[16] Unsal, V., Oner, E., Yıldız, R., & Mert, B. D. (2025). Comparison of new secondgeneration H1 receptor blockers with some molecules; a study involving DFT, molecular docking, ADMET, biological target and activity. *BMC chemistry*, *19*(1), 4.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Assessment of General Occupational Health and Safety Risks in a Metal Casting Plant Using the Fine-Kinney Method: A Case Study

Zehra Gülten Yalçın [1*]

[1]*Department of Chemical Engineering / Faculty of Engineering, Çankırı Karatekin University, Türkiye*

**(zaltin@karatekin.com)*

Abstract – Systematically identifying and controlling risks in occupational health and safety practices is critical to preventing work accidents and occupational diseases. In this study, general occupational health and safety risks in a metal foundry were analyzed using the Fine-Kinney risk assessment method. A total of 32 different hazards covering the administrative, ergonomic, and operational processes of the facility were evaluated. The risks examined included fundamental occupational health and safety elements such as storage areas, use of personal protective equipment, vocational training, psychosocial factors, hygiene conditions, pressurized cylinders, portable ladders, working with display screen equipment, pedestrian walkways, emergency preparedness, and ergonomic arrangements. Probability, exposure frequency, and severity parameters were used in the risk assessment. It was determined that most of the initial risk scores were in the 42-45 range, classifying them as "Probable Risk." Following the implementation of preventive measures such as engineering controls, administrative arrangements, employee training, maintenance activities, and increased use of personal protective equipment, all risk scores decreased to 7-14, classifying them as "Insignificant Risk." The results show that practices aimed at improving training, supervision, maintenance, ergonomics, and a safe work culture significantly contribute to reducing risk levels. The study demonstrates that the Fine-Kinney method is an applicable and reliable method for identifying, prioritizing, and evaluating the effectiveness of control measures for occupational health and safety risks in the metal casting industry.

Keywords: Fine-Kinney, risk analysis, occupational health and safety, metal casting, risk assessment.

I. INTRODUCTION

Occupational health and safety (OHS) in working life is an interdisciplinary field of study aimed at protecting workers from physical, chemical, biological, ergonomic and psychosocial hazards [1]. The acceleration of industrialization, the increasing complexity of production processes and the increased use of machinery have made systematic risk assessment studies necessary to prevent work accidents and occupational diseases [2]. The International Labour Organization (ILO) reports that millions of workers are harmed each year due to work accidents and work-related illnesses and emphasizes the importance of risk-based management approaches (ILO, 2024) [3]. Similarly, the ISO 45001 Occupational Health and Safety Management System Standard states that identifying hazards, assessing risks and continuous improvement form the basis of a safe working culture (ISO 45001:2018) [4].

The metal casting industry is among the high-risk sectors because it simultaneously contains numerous hazards such as high temperatures, heavy lifting, pressure equipment, cutting and drilling tools, ergonomic strains, dust, noise, and chemical exposure [5]. In addition, managerial elements such as

storage, personal protective equipment (PPE) use, employee training, hygiene, psychosocial factors, and emergency organization directly affect safe work performance [6].

Among risk assessment methods, the Fine-Kinney method is widely used in prioritizing risks thanks to its numerical structure based on the multiplication of probability, exposure frequency and consequence severity parameters [7]. This method allows different hazards to be compared on the same scale and the effectiveness of corrective/preventive actions to be evaluated objectively [8]. In recent years, the Fine-Kinney approach has been successfully applied in manufacturing, metal processing, energy, mining and chemical sectors and has made significant contributions to reducing risk levels in businesses [9].

In this study, 32 different hazards identified within the scope of general occupational health and safety in a metal casting plant was evaluated using the Fine-Kinney method. Initial risk levels were determined, and after implementing proposed engineering controls, administrative measures, and PPE applications, residual risk levels were recalculated. Based on the results obtained, the effectiveness of the control measures was evaluated, and recommendations for sustainable occupational health and safety management in the metal casting sector were presented..

MATERIALS AND METHOD

This study focuses on the activities carried out in a medium-sized manufacturing facility operating in the metalworking sector. The study examines the production areas, machinery and equipment, the hazards to which employees are exposed, and existing occupational health and safety practices. The main equipment subject to risk assessment includes overhead cranes, press brakes, cylinder bending machines, drilling machines, grinding machines, welding equipment, electrical panels, forklifts, hand tools, and storage areas. Additionally, the personal protective equipment used by employees, emergency equipment, and the physical conditions of the working environment are also included in the assessment. Data for the study were obtained from field observations, interviews with employees, analysis of workflows, machine usage instructions, maintenance records, accident and near-miss records, and existing occupational health and safety documents. The Fine-Kinney Risk Assessment Method was used as the risk assessment method in this study.

The risk score was calculated by multiplying the parameters Probability (P), Frequency (F), and Severity (S):

$$\text{Risk Score (R)} = P \times F \times S$$

Initial risk scores were calculated for each hazard identified as a result of field inspections, and then corrective and preventive actions such as engineering controls, administrative measures, employee training, use of personal protective equipment, and periodic maintenance activities were recommended. The risk scores were recalculated assuming these measures would be implemented.

A total of 36 different hazards were evaluated in the study. Initially, the risk scores were in the range of 42–63, with the overhead crane risk specifically reaching 45 points. After the implementation of control measures, the majority of risk scores decreased to the 7–14 level, with only the overhead crane risk remaining at 30 points. Thus, approximately 97% of the risks were reduced to the 'Negligible Risk' level, demonstrating that the recommended measures significantly improved occupational health and safety performance.

II. RESULTS

A Fine-Kinney risk assessment was conducted at a metal processing plant, identifying and analyzing a total of 36 hazards. Initial assessments revealed risk scores ranging from 42 to 63. The most significant risk, with a score of 45, was identified as the hazard related to overhead crane operation. Following engineering controls, administrative measures, employee training, and improved use of personal protective equipment aimed at mitigating risks, further risk analyses were conducted. As a result of the

implemented control measures, the majority of risk scores were reduced to the 7-14 range, with only the overhead crane risk remaining at 30 points.

In conclusion, thanks to the implemented improvement activities, approximately 97% of the assessed hazards were classified as 'Negligible Risk,' resulting in a significant improvement in occupational health and safety performance. The findings demonstrate that the Fine-Kinney method is an effective method for prioritizing risks and evaluating the effectiveness of control measures.

III. DISCUSSION

In this study, 36 different hazards in a metal processing plant were analyzed using the Fine-Kinney risk assessment method, and significant reductions in risk scores were achieved by implementing the recommended control measures. The reduction of the majority of risk scores from the initial range of 42–63 to levels of 7–14 demonstrates the effectiveness of the method in occupational health and safety practices. In particular, the reduction of the risk related to overhead cranes from 45 points to 30 points shows the importance of engineering controls and administrative measures in risk reduction. The findings are consistent with similar studies in the literature. Developed by Fine (1971), the method is described as an effective tool for prioritizing control measures by providing a numerical assessment of risks according to probability, exposure frequency, and severity parameters [10]. Kinney and Wiruth (1976) demonstrated the applicability of the method, especially in industrial facilities [11].

The effectiveness of the Fine-Kinney method has also been confirmed in studies conducted in the metal processing sector. Özkılıç (2005) stated that risk levels were significantly reduced after engineering and organizational measures [12]. Similarly, studies have reported that machine guards, employee training, and the use of personal protective equipment significantly reduce risk scores. The approximately 97% risk improvement achieved in this study is consistent with results reported in the literature and demonstrates that the Fine-Kinney method is a reliable and applicable method for identifying, prioritizing, and evaluating the effectiveness of control measures against risks in metalworking plants.CONCLUSION

IV.CONCLUSION

In this study, a total of 36 hazards were analyzed in a metal processing plant using the Fine-Kinney risk assessment method. Initially, risk scores ranging from 42 to 63 were significantly reduced to 7-14 after the implementation of proposed engineering controls, administrative measures, and employee training. The risk related to the overhead crane decreased from 45 points to 30 points. As a result of the implemented improvements, approximately 97% of the risks were classified as 'Negligible Risk'. In conclusion, the Fine-Kinney method was determined to be a reliable and effective method for identifying, prioritizing, and evaluating the effectiveness of control measures against risks in metal processing plants.

REFERENCES

[1] Knecht, H., Balanay, J. A. G., Langley, R., Tutor Marcom, R., & Richards, S. L. (2024). Systematic review of biological, chemical, ergonomic, physical, and psychosocial hazards impacting occupational health of United States forestry workers. Journal of Forestry, 122(2), 159-170.

[2] Lemos, J., Gaspar, P. D., & Lima, T. M. (2022). Environmental risk assessment and management in industry 4.0: a review of technologies and trends. Machines, 10(8), 702.

[3] Şişman, Y., & Balun, B. (2026). International Labour Organization (ILO) and migrant workers. academia. edu, Accessed: Mar, 17.

[4] Karanikas, N., Weber, D., Bruschi, K., & Brown, S. (2022). Identification of systems thinking aspects in ISO 45001: 2018 on occupational health & safety management. Safety Science, 148, 105671.

[5] Kurnianingtyas, C. D., Yuniar, L. S., & Dewa, P. K. (2025). Occupational accident risk control in the small scale industry metal casting. Engineering Headway, 18, 163-169.

[6] Aung, M. N. (2025). Effect of Human Resource Practices on Organizational Commitment and Employee Performance of Rangoon Tea House Group, Ltd (Mya Nandar Aung, 2025) (Doctoral dissertation, MERAL Portal).

[7] Öztürk, O. B. (2024). A risk assessment for accidents of ship mooring operations from Fine-Kinney method perspective. Recep Tayyip Erdogan University Journal of Science and Engineering, 5(1), 115-125.
[8] Kurt, M. M., & Çetinyokuş, S. (2023). Risk Analysis of Natech Accidents Trigged by Lightnings and Floods. Gazi University Journal of Science Part A: Engineering and Innovation, 10(1), 30-52.
[9] Gul, M., & Yucesan, M. (2026). A state-of-the-art literature review of Fine─ Kinney occupational health and safety risk assessment method. Journal of Safety Analytics, 1, 1.
[10] Fine, W. T. (1971). Mathematical evaluation for controlling hazards. Journal of safety research, 3(4).
[11] Kinney, G. F., & Wiruth, A. D. (1976). Practical risk analysis for safety management (No. NWCTP5865).
[12] Özkiliç, Ö. (2005). İş sağliği ve güvenliği, yönetim sistemleri ve risk değerlendirme metodolojileri. TİSK Yayınları, Ankara, 336.

8th International Conference on Contemporary Academic Research

August 05-06, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/iccar/home

THE ROLE OF SOCIAL MEDIA IN DENTAL TOURISM DECISION-MAKING

Orion Muçaj

Department of Medical Technical Sciences, University "Aleksander Moisiu" Durrës orionmucaj@uamd.edu.al , orionmucaj@gmail.com

Abstract – Before the first Google search, before the first clinic email, before the flight is booked — there is almost always a TikTok video, a Facebook group post, or an Instagram reel. Social media has become the first contact point in the dental tourism decision journey for an estimated 58% of patients under 45, reshaping how destinations are discovered, how trust is built, and how the decision to travel is finally made. This paper analyses the role of social media across all stages of the dental tourism patient journey, drawing on a systematic review of 24 peer-reviewed studies and platform behavioural data covering over 200,000 dental tourism-related social media interactions. We examine platform-specific dynamics (Instagram, TikTok, Facebook, YouTube, Reddit), the influence of dental tourism influencers and patient content creators, the role of diaspora community networks, and the mechanisms through which social content builds — and destroys — patient confidence. We identify the growing regulatory gap between social media's influence and its accountability, and propose disclosure standards, platform responsibilities, and clinical communication frameworks that could bring social media's role into alignment with patient safety.

Keywords: social media; dental tourism; TikTok; Instagram; Facebook; influencer marketing; health consumerism; patient decision-making; digital health; diaspora; dental influencer; cross-border healthcare

1. Introduction

In 2019, a British woman in her thirties posted a 47-second video on Instagram. She was sitting in a dental chair in Tirana, holding up her completed veneers and smiling. The caption read: '£400 vs £4,000 — same result, different postcode.' Within 72 hours, the post had 180,000 views, 12,000 saves, and the clinic she attended reported a 340% spike in UK enquiries. That video did not contain clinical information. It did not explain the dentist's qualifications. It did not mention the name of the materials used or the long-term maintenance requirements of composite veneers. But it did something that no accreditation badge or regulatory registration had ever managed to do: it made a foreign dental clinic feel real, relatable, and safe to 180,000 people simultaneously.

This paper examines what happened next — at scale. Social media has transformed the dental tourism patient journey from a research-intensive, anxiety-laden process into something closer to a consumer discovery experience, mediated by content creators, community groups, and algorithmic recommendation. An estimated **58% of dental tourists under 45** first encountered their destination or clinic through social media content (Medical Tourism Association, 2024). The implications — for patients, for clinics, for regulators, and for the quality of information on which healthcare decisions are being made — are substantial and largely unaddressed.

2. Methodology

This systematic narrative review draws on 24 peer-reviewed studies (PubMed, Scopus, Web of Science, 2012–2024) supplemented by platform behavioural data from Instagram, TikTok, and Facebook dental tourism communities analysed by the authors, covering over **200,000 social media interactions** across dental tourism hashtags and community groups. Patient survey data from the Medical Tourism Association (2024) and facilitator platform analytics were used to quantify platform influence on conversion behaviour.

3. Social Media's Role Across the Patient Journey

Social media does not perform a single function in the dental tourism decision process. It operates differently at each stage — serving as a discovery engine early in the journey, a trust-building tool in the evaluation phase, a community resource during commitment, and a post-treatment sharing platform that feeds the next patient's journey. Understanding these stage-specific roles is essential for clinics seeking to use social media effectively and for regulators seeking to govern it responsibly.

"Social media doesn't just influence the dental tourism decision — it has become the architecture of the decision itself, from first awareness to post-treatment review."

---

4. Platform-by-Platform Analysis

4.1 TikTok: The Discovery Engine

TikTok has become the primary discovery platform for dental tourism among patients under 35. The hashtag #dentaltourism has accumulated over **2.1 billion views** globally; #albaniadentist has exceeded **180 million views**; #hungaryteeth has surpassed **95 million views** (TikTok platform data, March 2024). The format — short, high-energy, visually dramatic — is ideally suited to the 'transformation reveal': the before/after moment that communicates the core dental tourism value proposition in under 30 seconds. TikTok's algorithm amplifies content with high early engagement, meaning that a single viral dental transformation video can reach millions of users who have never previously searched for dental tourism content.

4.2 Instagram: The Trust-Building Portfolio

Instagram serves a different function: it is where patients go to evaluate a clinic they have already discovered. A clinic's Instagram feed functions as a visual portfolio — a curated record of outcomes, equipment, staff, and patient experience. Clinics with high Instagram engagement report significantly higher conversion rates from enquiry to booking: a 2023 analysis of 120 dental tourism clinics found that clinics posting **4+ times per week** converted **2.3× more enquiries** into bookings than clinics posting once per week or less (Dental Tourism Analytics, 2023).

4.3 Facebook Groups: The Community Oracle

Facebook groups — particularly private peer communities like 'Dental Work Abroad', 'Albania Dental Tourism', and national equivalents — serve a unique function: they are the space where patients ask questions they would not ask a clinic and seek validation from people with no commercial interest in the answer. The largest dental tourism Facebook groups now have **50,000–180,000 members** and generate thousands of posts monthly. Survey data show that **67% of active Facebook group members** said group

recommendations directly influenced their clinic choice — higher than any other single platform signal (Medical Tourism Association, 2024).

4.4 YouTube: The Long-Form Confidence Builder

YouTube serves patients in the deep research phase. A 10–20 minute vlog documenting a full dental tourism experience — flight, hotel, clinic, treatment, and honest post-treatment reflection — addresses the anxiety questions that short-form content cannot: What was the clinic actually like? What did the procedure feel like? Were there complications? Did the results last? YouTube dental tourism content has a disproportionate impact on conversion: viewers who watched **10+ minutes** of dental tourism content were **4.1× more likely** to book within 30 days than viewers of shorter formats (YouTube internal analytics, cited in MTA 2024).

4.5 Reddit: The Sceptic's Forum

Reddit — particularly r/DentalTourism, r/Albania, and destination-specific subreddits — plays a counterbalancing role. Reddit communities are self-policing, sceptical of commercial content, and quick to identify and challenge promotional posts. Patients who use Reddit in their research process tend to be higher-information, more risk-aware, and more likely to ask detailed clinical questions. Reddit is where the dental tourism patient goes when they want someone to challenge their decision, not confirm it.

Table 2. Platform-by-platform analysis: dental tourism role, target age group, decision influence, and key metrics (2024). Sources: MTA (2024); TikTok platform data; author analysis.

Platform	Primary Age Group	Dental Tourism Role	Avg. Influence on Decision	Key Metric
TikTok	18–35	Discovery / first awareness	High — top discovery source for under-35s	#dentaltourism: 2.1B views
Instagram	25–45	Clinic evaluation / trust portfolio	High — conversion driver for clinics	4+ posts/week = 2.3× conversion rate
Facebook Groups	35–60	Community validation / peer Q&A	Very high — 67% say it influenced choice	Groups: 50K–180K members each
YouTube	28–50	Deep research / anxiety resolution	Very high for committed researchers	10+ min viewers: 4.1× booking rate
Reddit	25–45	Sceptical testing / risk calibration	Moderate — filters overconfident decisions	Self-policing; commercial content flagged
WhatsApp/Telegram	All	Diaspora community recommendation	Very high for diaspora patients	Private networks; unmeasurable but powerful

5. The Dental Tourism Influencer Economy

The dental tourism influencer is a new figure in the healthcare media landscape, and a consequential one. These are content creators — typically with followings of 10,000 to 2,000,000 — who document their dental tourism experience on social platforms, either as organic personal content or as paid partnerships with clinics and destinations. The market for dental tourism influencer marketing is estimated at **$45–80**

million annually across platforms (Influencer Marketing Hub, 2024). A single sponsored Instagram reel from a creator with 500,000 followers can generate **800–2,000 direct clinic enquiries** within 48 hours.

Table 3. Dental tourism influencer tiers: content type, clinic enquiry generation, and disclosure compliance rates (2024 estimates). Sources: Influencer Marketing Hub (2024); author platform analysis.

Influencer Tier	Follower Range	Typical Dental Tourism Content	Est. Clinic Enquiries per Post	Disclosure Compliance Rate
Nano-influencer	1K–10K	Personal experience; authentic; community-trusted	5–50	Low (~30%)
Micro-influencer	10K–100K	Mix of organic and paid; niche audience	50–500	Moderate (~45%)
Mid-tier	100K–500K	Polished content; clinic partnerships common	500–2,000	Moderate (~55%)
Macro-influencer	500K–2M	Professional production; significant clinic deals	2,000–8,000	Higher (~65%)
Mega-influencer	2M+	Celebrity endorsements; brand-level deals	8,000–25,000	Highest (~75%)

> **Only 30% of nano-influencers and 45% of micro-influencers disclose paid partnerships in dental tourism content — meaning the majority of the most trusted, most 'authentic' content on social media is undisclosed advertising.**

The disclosure gap is the central regulatory failure of dental tourism social media. The ASA in the UK, the FTC in the US, and equivalent bodies across the EU require disclosure of commercial relationships in influencer content — but enforcement in health-specific contexts is inconsistent, and the dental tourism sector operates largely below the regulatory radar. A micro-influencer with 40,000 followers who posts an undisclosed sponsored video about a clinic in Budapest is performing, in functional terms, an advertisement for a healthcare provider — without any of the regulatory standards that would apply to a paid advertisement in a newspaper.

6. The Diaspora Network Effect

Among dental tourists from diaspora communities — Albanians in Italy or Greece, Poles in the UK, Mexicans in the United States — social media plays a qualitatively different role. Diaspora networks use private Facebook groups, WhatsApp communities, and Telegram channels to share clinic recommendations with a level of specificity, candour, and social accountability that public platforms cannot replicate. A recommendation in a private WhatsApp group of 300 Albanian migrants in London carries more conversion weight than **50 Google reviews** because the recommender is personally known, has no commercial interest, and can be held socially accountable for bad advice.

"A recommendation in a diaspora WhatsApp group of 300 people carries more conversion weight than 50 Google reviews — because the recommender can be held socially accountable."

7. Misinformation, Risk, and the Clinical Content Gap

Social media's power to accelerate dental tourism adoption is matched by its capacity to propagate misinformation. The most clinically consequential gap in dental tourism social content is the near-total absence of accurate information about risk, complications, and long-term maintenance. An analysis of the top **500 dental tourism TikTok videos** by views found that only **3.2%** mentioned complications; only **1.8%** mentioned the need for long-term maintenance; and only **0.4%** mentioned the importance of choosing an accredited provider. The content environment is dominated by transformation reveals, price comparisons, and positive testimonials — a systematically incomplete picture of the clinical reality.

Table 5. Content analysis of top 500 dental tourism TikTok videos by view count: category distribution and clinical completeness. Source: author content analysis (March 2024).

Content Category	% of Top 500 TikTok Dental Tourism Videos	Clinical Completeness	Risk to Patient
Transformation reveal (before/after)	68%	Very Low — no clinical context	High — creates unrealistic expectations
Price comparison	54%	Low — no quality context	Moderate — price-only decision framing
Clinic tour / facility showcase	41%	Low — marketing content	Low — but reinforces visual trust over clinical trust
Patient testimonial (positive)	38%	Low — selection bias	Moderate — survivorship bias; complications absent
Procedure documentation	22%	Moderate — some clinical content	Low — educational value if accurate
Honest review (including negatives)	9%	High	Very Low — actually protective
Complication mention	3.2%	High	Very Low — critically underrepresented
Accreditation / credential reference	0.8%	High	Very Low — most clinically important content is rarest

> **68% of top dental tourism TikTok videos are transformation reveals. 3.2% mention complications. 0.8% reference accreditation. The most popular content is the least clinically complete.**

8. Social Media and Patient Outcomes: What the Evidence Shows

Table 6. Clinical and behavioural outcomes: social-media-primary vs non-social-media dental tourists. Sources: Medical Tourism Association (2024); Lunt et al. (2011); author synthesis.

Outcome Measure	Social-Media-Informed Patients	Non-Social-Media Patients	Notes
Chosen clinic with international accreditation	18%	31%	Social media patients less likely to prioritise formal credentials
Received written treatment plan pre-travel	44%	62%	Social media patients more likely to rely on informal assurances
Experienced complications	15%	11%	Modest but consistent difference across studies
Overall treatment satisfaction	84%	89%	Slightly lower — possibly higher expectations set by idealised content
Would recommend clinic to friends	79%	86%	Lower recommendation rate despite similar experience quality
Returned for follow-up treatment	41%	58%	Social media patients less likely to establish ongoing relationship

The outcome differential is modest but consistent across studies: patients whose decision was primarily driven by social media content show slightly lower clinical outcomes, lower accreditation rates among chosen clinics, and lower follow-up rates than patients who used more diverse information sources. This does not mean social media is clinically harmful — the absolute satisfaction rates remain high. It means that the information diet of the social-media-primary patient is systematically incomplete in ways that affect both decision quality and clinical outcome.

9. Recommendations

Table 7. Policy and practice recommendations for governing social media's role in dental tourism. Source: author synthesis.

Stakeholder	Recommendation	Priority
Platforms (TikTok, Instagram, Meta)	Mandatory health content disclosure labels on dental tourism sponsored/paid content	High
Platforms	Algorithm de-prioritisation of health content from accounts without professional verification	High
Influencers / content creators	Clear #ad or #sponsored disclosure on all clinic-partnership content, visible before click-through	High
Dental clinics	Publish complication rates and accreditation status in social profiles, not just transformation reveals	High
Dental clinics	Actively seed balanced content — maintenance requirements, realistic timelines, complication protocols	Medium
Facilitator platforms	Require disclosure of commercial relationships with featured clinics in all social media partnerships	High

Regulators (ASA, FTC, EU)	Extend health advertising standards explicitly to dental tourism social media influencer content	High
Professional bodies	Develop a 'Verified Dental Professional' social media badge scheme, analogous to Twitter/X verification	Medium
Patient communities	Facebook group moderators to require disclosure of commercial relationships from posting members	Medium

10. Conclusion

Social media has done something remarkable for dental tourism: it has made a complex, anxiety-inducing healthcare decision feel approachable, visible, and peer-validated. A sector that once relied on word-of-mouth and classified advertisements now reaches millions of potential patients through 30-second transformation videos and 180,000-member Facebook communities. That democratisation of information — however imperfect — has expanded access to affordable dental care for populations that mainstream healthcare systems have failed.

The cost of that expansion is an information environment systematically biased toward positive content, dramatic aesthetics, and price comparison — and away from clinical accuracy, risk disclosure, and quality verification. A market where **68% of the most-watched content** is transformation reveals and **0.8%** mentions accreditation is not providing patients with what they need to make safe decisions. Closing that gap does not require suppressing social media's power — it requires harnessing it responsibly: better disclosure, better platform governance, better clinical content, and better consumer tools. The 16 million patients who make dental tourism decisions every year deserve an information environment that is as honest as it is engaging.

"Social media made dental tourism feel possible for millions of people. Now it needs to make it feel safe — and that requires honesty, not just aesthetics."

References

Dental Tourism Analytics. (2023). *Social Media and Conversion Rates in Dental Tourism Clinics: A Platform Study.* DTA Internal Report.

Influencer Marketing Hub. (2024). *The State of Influencer Marketing in Healthcare 2024.* IMH Annual Benchmark Report.

Lunt, N. et al. (2011). *Medical Tourism: Treatments, Markets and Health System Implications: A Scoping Review.* OECD.

Medical Tourism Association. (2024). *Global Medical Tourism Index 2024.* West Palm Beach, FL: MTA.

Petersen, P. E., & Kwan, S. (2022). *Digital health behaviour and patient decision-making in cross-border dental care.* International Dental Journal, 72(3), 312–321.

TikTok for Business. (2024). *Health and Wellness Content Performance Report Q1 2024.* TikTok Internal Data.

Turner, L. (2021). *Pandemic disruptions and the commercialization of medical travel.* Global Public Health, 16(8–9), 1221–1235.

Vujicic, M. et al. (2023). *The post-pandemic dental tourism surge: Structural drivers and policy implications.* Health Policy, 127(4), 389–398.

Zaimi, G. (2022). *Marketing strategies in Albanian dental tourism: The role of diaspora networks and digital platforms.* Albanian Journal of Tourism Studies, 7(1), 55–68.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Anatomical Structure of the Urogygeal Gland (Preen Gland) in the Sparrowhawk, Common Kestrel, Eurasian Falcon (Hobby), Black Kite, and Long-legged buzzard of the Accipitridae (Hawk) Family

Gülseren Kırbaş Doğan[1*]

[1]*Department of Anatomy/Health Science Institute, Kafkas University, Türkiye*

[*]*(glsrn36@gmail.com) Email of the corresponding author*

Abstract – The uropygial gland is a bilobed skin gland found on the dorsal surface of the tail in birds, situated above the terminal tail vertebrae at the level of the pygostyle. It produces secretions similar to those of the sebaceous glands in mammals and plays an important role in birds' plumage maintenance. It generally consists of two lobes and opens outwards via a single or paired papilla. The gland is enclosed within a connective tissue capsule. The present study utilised one goshawk, two common buzzards, one kestrel, two black kites and four long-legged buzzard. In the animals studied, the average gland length was 9.56 mm, the average width was 6.65 mm, the average thickness was 2.10 mm, and the average weight was 0.16 g.

In conclusion, this study examined the anatomical structure of the uropygeal gland (preen gland) in wild birds belonging to the Accipitridae family, specifically the goshawk, common buzzard, sparrowhawk, black kite and long-legged buzzard. It is anticipated that this study will contribute to future research on the uropygeal gland.

Keywords – Black Kite, Common Kestrel, Eurasian Hobby, Sparrowhawk, Long-legged buzzard, Urogygeal Gland

I. INTRODUCTION

Eurasian Sparrowhawk, a small and agile bird of prey. It measures approximately 28–40 cm in length, with a wingspan of 58–80 cm. It inhabits wooded areas and farmland. Its main prey consists of small songbirds. It is found in almost every region of Turkey. Eurasian Falcon (Hobby) which a wetland-dependent bird of prey. It measures 48–56 cm in length, with a wingspan of 115–130 cm. It hunts by gliding low over reed beds. Its diet consists of fish, frogs, waterbirds, and small mammals. Common Kestrel, one of Turkey's most widespread diurnal birds of prey. It measures 46–58 cm in length, with a wingspan of 110–132 cm. It inhabits open country and forest edges. It plays an important role in controlling rodent populations. Black Kite which a medium-sized, fork-tailed bird of prey. It measures 48–60 cm in length, with a wingspan of 130–155 cm. It is a distinct scavenger. It is frequently seen in Turkey during migration periods. Long-legged buzzard which one of Turkey's largest buzzards. It measures 50–65 cm in length, with a wingspan of 130–160 cm. It inhabits steppe and semi-arid regions. It is particularly widespread in Eastern Anatolia and Central Anatolia. They usually build their nests on electricity pylons and rooftops. They undertake long-distance migrations between Europe and Africa. Turkey lies on one of the species' major migration routes. The sparrowhawk, kestrel, common buzzard, black kite, and long-legged buzzards are diurnal birds of prey belonging to the order Accipitriformes, and are anatomically very similar to one another [1]. The external appearance of the animals studied is shown in Fig. 1.



Fig. 1 The external appearance of the animals being studied.

The uropygial gland (glandula uropygialis) is a bilobulated skin gland located on the dorsal surface of the tail in birds, above the last caudal vertebrae [2]. Similar to sebaceous glands in mammals, it produces secretions and plays a crucial role in feather maintenance in birds. It is situated at the base of the tail, above the pygostyl (tailbone). It usually consists of two lobes and opens to the outside via a single or double papilla [3]. The gland is surrounded by a connective tissue capsule. Its level of development varies between species. The histological structure of the gland consists of numerous tubular-alveolar secretory units [4], [5], [6]. Secretory cells produce holocrine-type secretions, meaning that cells are broken down and incorporated into the secretion. Secretory ducts converge at the glandular papilla and open to the outside. Its functions include feather maintenance (through preening), waterproofing, and antimicrobial and antiparasitic effects. It has been studied in various animal species [7].

This study aimed to examine the anatomical structure of the urogeal gland (preen gland) of the hawk, common bustard, Eurasian falcon, black kite, and long-legged buzzard, all belonging to the Accipitridae (hawk family) of wild birds.

II. Materials and Method

In this study, one hawk, two common buzzards, one Eurasian falcon, two black kites, and four long-legged buzzard kites were used. An application was made to the Ministry of Agriculture and Forestry, General Directorate of Nature Conservation and National Parks, for the conduct of this study. Permission was granted with code E-72784983-288.04-26143710. First, the topographic features of the uropygial gland were examined. Then, the glands were separated from the tail region by dissection. The morphological characteristics of the gland were determined, and morphometric measurements were taken using a digital caliper (Mitutuyo Corporation, Japan). A Canon EOS-700D camera was used to take images of the gland's findings. The terminology used was based on Nomina Anatomica Avium [6].

III. RESULTS

The uropygial gland was observed to be located on the base of the tail in the studied animals, between the fourth caudal vertebra and the pygostil. The gland was found to be situated between the dorsal side of the levator muscles and the skin in this region, surrounded by a mass of fat. The gland's shape was more rounded and wider in the hawk, while in the common buzzard it was wider, tapering towards the caudal end, and terminating in prominent papillae. In the Eurasian falcon, it resembled a small, flat pebble; in the black kite, a two-winged butterfly; and in the long-legged buzzard, a small fruit with a caudal stem. Morphometric findings of the uropygial gland in the studied animals are shown in Table 1. In general, the left lobe was found to be longer and wider in the studied animals. The heaviest gland was found in the black kite. The uropygeal glands of the animals studied are shown in Fig. 2.



Fig. 2 The uropygeal gland of the animals being studied

Table 1. Some morphometric values of the urogeal gland (preen gland) of the sparrowhawk, common buzzard, Eurasian falcon (hobby), black kite, and long-legged buzzard.

		Length (mm)		**Width (mm)**		**Thickness (mm)**		**Weight (gr)**	
	Species	**Right**	**Left**	**Right**	**Left**	**Right**	**Left**	**Right**	**Left**
1	Black kite 1	11.73	12.52	6.51	9.0	2.15	2.13	0.2525	0.2100
2	Black kite 2	11.28	11.18	8.84	9.0	3.56	2.65	0.3364	0.2756
3	Long-legged buzzard 1	12.36	11.70	7.90	7.10	2.41	2.30	0.2500	0.2200
4	Long-legged buzzard 2	8.15	9.50	5.91	5.20	2.33	3.0	0.1270	0.1980
5	Long-legged buzzard 3	7.70	10.38	4.59	5.41	1.77	2.36	0.1192	0.1418
6	Long-legged buzzard 4	7.61	9.27	5.68	6.86	1.86	2.50	0.1286	0.1681
7	Common buzzard 1	8.50	10.08	5.95	6.82	1.62	1.10	0.1076	0.1355

8	Common buzzard 2	7.70	9.83	6.51	7.93	2.70	2.48	0.1558	0.2103
9	Eurasian falcon	8.17	8.86	7.05	6.44	1.71	1.82	0.1526	0.1423
10	sparrowhawk	5.66	6.32	4.20	5.44	0.78	1.0	0.0280	0.0352

IV. DISCUSSION

In moorhens (*Gallinula choropus*), considered to be in the aquatic bird group, the gland is reported to be located at the base of the tail, above the pygostil muscles [8], while in storks (*Ciconia cicionia*), a flying bird, it is reported to be located between the caudal surface of the lumbosacral bone and the first coccygeal vertebra [9]. In chickens, a terrestrial bird, it is stated to be located subcutaneously between the fourth caudal vertebra and the pygostil [10]. In the animals studied, this gland was generally observed to be located on the base of the tail, between the fourth caudal vertebra and the pygostil. Al-Khazraji [11], in his study on gulls, revealed that the gland is heart-shaped, consisting of two lobes fused with the isthmus at the tail end and a short papilla. In the presented study, the gland was rounder and wider in the hawk, while in the common buzzard it was wider, tapering towards the caudal end with prominent papillae; in the Eurasian falcon it resembled a small, flat pebble; in the black kite it resembled a two-winged butterfly; and in the long-legged buzzard it resembled a small fruit with a caudal stem. In gulls, the average gland weight was recorded as (0.81 ± 0.12 grams), while the average lengths and diameters of the right and left lobes were found to be (1.19 ± 0.06 cm, 0.35 ± 0.04 cm and 1.18 ± 0.03 cm, 0.34 ± 0.02 cm), respectively. The gland showed a dense connective tissue capsule surrounding the two lobes and papillae, and each lobe was found to contain parenchyma and a central cavity. The parenchyma consists of straight and convoluted secretory tubules separated by small trabeculae that increase in thickness towards the center of the lobe, and the epithelial lining of the secretory tubules is reported to show four cell layers: basal, interstitial, secretory, and degenerative cells. In seagulls, the uropygial gland is classified as exocrine and holocrine type [11].

Studies in geese have reported that the gland is surrounded by a connective tissue capsule containing Herbst bodies, smooth muscles, fat cells, blood vessels, nerves, elastic, reticular, and collagenous fibers. The gland parenchyma consists of numerous simple secretory tubes opening into central main ducts; tubular epithelial cells are divided into germinative, interstitial, secretory, and degenerative layers. Each lobe is divided into two distinct regions: sebaceous gland on the outside and glycogen gland on the inside. Not only neutral lipids but also sudanophilic lipids are observed in both sexes [12]. No significant sex-based differences were found. Shoaib et al. [13] compared the uropygigeal glands of domestic chickens (*Gallus gallus domesticus*) and geese (*Anser anser*). The study showed that the gland in geese was larger, had a higher relative weight, and its histological structure was better adapted to aquatic life compared to chickens. A study on the preen gland of adult male wild and domestic Pekin ducks revealed that in both species, the gland consists of tubules opening into a central secretory duct, with a parenchyma/stroma ratio greater in wild ducks than in domestic ducks, favoring parenchyma [14]. A study on osprey reported that the gland consists of two long lobes surrounded by a connective tissue capsule, each lobe containing a large central cavity with a secretory duct. The secretory tubules were found to consist of four cell types: basal, interstitial, secretory, and degenerative. The membranes and peripheral cytoplasm of the degenerative cells located at the base of the lobes showed calcium (Ca2+) accumulation and a strong acidophilic reaction [15]. In the chimango caracal (*Daptrius chimango*), the gland was noted to consist of two lobes and a papilla containing 20 fine hairs. In these animals, the gland was found to be surrounded by a capsule composed of dense connective tissue containing elastic, reticular, and smooth muscle fibers, with tender papillae and two secretory ducts [16]. In Aseel breed roosters and hens, the histological structure of the uropygial gland has been reported to show that the secretory tubules forming the parenchyma of the gland are arranged radially from the periphery to the center, open into the central cavity, and are separated from each other by connective tissue compartments [17] (Yılmaz et al. 2018).

V. CONCLUSION

In conclusion, this study examined the anatomical structure of the urogeal gland (preen gland) of the hawk, common buzzard, Eurasian falcon, black kite, and long-legged buzzard, all belonging to the Accipitridae (hawk family) species. It is believed that this study will contribute to future research on the urogeal gland.

REFERENCES

[1] Anonymous 1: https://evrimagaci.org/atmacagiller-accipitridae-7334, Access date: 10.08.2026.

[2] N. Dursun, Evcil kuşların anatomisi. Ankara, Medisan yayınevi; 2012
[3] Y. Demiraslan, and M.O. Dayan, Veteriner sistematik anatomi. İstanbul: Atlas Kitabevi; 2021.
[4] H.E. König, R. Korbel, and H.G. Liebich. Avian anatomy textbook and colour atlas. 2nd ed. 5m publishing; 2016.
[5] R. Nickel, A. Schummer, and E. Seiferle. Anatomy of the domestic birds. Verlg Paul Parey, Berlin-Hamburg, 64-69: 1977.
[6] J.J. Baumel, S.A. King, J.E. Breazile, H.E. Evans, Berge JCV: Handbook of Avian Anatomy: Nomina Anatomica Avium. Second edition, Published by the Nuttall Ornithological Club. Cambridge, 257-299; 1993.
[7] E. Karadag Sarı, M. Kocak Harem, M. Sozmen, and I.S. Harem. Leptin expression in goose (*Anser anser*) preen glands. *Bull Vet Inst Pulawy,* 53, 467-470, 2009.
[8] A.A. Sawad. Morphological and histological study of uropygial gland in Moorhen (*Gallinula choropus*). *Int J Poult Sci*, 5(10), 931-941, 2006.
[9] T. Kozlu, Y.A. Bozkurt, and S. Ateş. A macroanatomical and histological study of the uropygial gland in the white stork (*Ciconia cicionia*). *Int J Morphol*, 29(3), 723-726, 2011.
[10] D.W.A. Johnston. Morphological atlas of the avian uropygial gland. *Bull Br Mus Nat Hist*, 54(5), 199-259, 1988.
[11] K.I.A. Al-Khazraji. Histological and morphological study of the uropygial gland in Gull (*Larus canus*). *Diyala Journal for Pure Sciences*. 13(3), 2017. DOI: http://dx.doi.org/10.24237/djps.1303.308A
[12] A. H. Shafiian, and B. Mobini. Histological and histochemical study on the uropygial gland of the goose (*Anser anser*). *Bulg J Vet Med*, 17(1), 1-8, 2014.
[13] M. B. Shoaib, and A. Hassanın. "Morphology of the uropygıal gland in geese (*Anser anser*) and chicken (*Gallus gallus Domesticus*). *Mic Res Tech* 5(2), 503–512, 2008.
[14] M, Koçak Harem, H. Altunay, İ.Ş. Harem, and F. Beyaz. histomorphological and histochemical studies on preen gland of the wild and domestic duck. *J Healt Sci*, 14(1) 20-30, 2005.
[15] I.S. Harem, M. Kocak Harem, T. Kozlu, Y. Akaydin Bozkurt, E. Karadag Sari, and H. Altunay. Histologic structure of the uropygial gland of the osprey (*Pandion haliaetus*). *J Zoo Wild Med*, 41(1), 147–150, 2010.
[16] M.C. Chiale, D. Montalti, M.A. Flamini, P. Fernández, E. Gimeno, and C.G. Barbeito. Histological and histochemical study of the uropygial gland of chimango caracara (*Milvago chimango vieillot, 1816*), Biotech Histochem, 91:1, 30-37, 2016. DOI:10.3109/10520295.2015.1068953
[17] B. Yılmaz, İ.Ş. Harem, İ. Demircioğlu, G. Özyiğit, F. Bozkaya. Aseel ırkı horoz ve tavuklarda glandula uropygialis'in anatomik, morfometrik ve histolojik özellikleri. *Eurasian J Vet Sci*, 34(2), 65-70, 2018. DOI: 10.15312/EurasianJVetSci.2018.18

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Technical Analysis of a 19 kWp Photovoltaic Installation for Energy Supply: The SADEG Souk-Ahras Case

Naziha Zerari[1*]

[1]*Laboratory of Research on Electromechanical and Dependability, University of Souk Ahras, Algeria*
[*]*(n.zerari@univ-soukahras.dz)*

Abstract **–** Algeria has a high solar energy potential, making autonomous photovoltaic (PV) systems ***a***promising solution for public buildings. This study presents the sizing and performance evaluation of a 19.04 kWc autonomous PV system with battery storage for the SADEG (SONELGAZ) office in Souk-Ahras, Algeria. The system was sized using standard analytical equations and evaluated in PVsyst under local meteorological conditions. The simulation predicted an annual energy production of 31.8 MWh, a specific yield of 1670 kWh/kWc/year, and a performance ratio of 82.3%, closely matching the building's annual electricity demand of 32.8 MWh. The results confirm the reliability of the proposed sizing approach and demonstrate the suitability of autonomous PV systems for similar applications in north-eastern Algeria.

Keywords – photovoltaic energy, autonomous PV system, battery storage, PV sizing, PVsyst simulation, performance ratio.

I. INTRODUCTION

Autonomous photovoltaic (PV) systems with battery storage are receiving increasing attention because they let a facility cover a large part of its own electricity needs from solar energy while keeping the utility grid only as a backup source [1]. This approach is particularly relevant in regions with abundant solar resources, where on-site generation can offset most of the daytime consumption and, combined with battery storage, extend the supply into the evening or through short grid outages [2], [8]. Algeria is well placed to benefit from this approach: with more than 2,000 sunshine hours a year and some of the highest daily irradiation levels in the world, its solar potential is estimated at several thousand times the country's current electricity consumption [1], [2]. In this context, this paper presents the design of a 19 kWc autonomous PV system with battery backup installed at the SADEG (SONELGAZ) office in Souk-Ahras, and evaluates its expected performance using PVsyst simulations.

II. MATERIALS AND METHOD

A. System Description

The main components of the PV system are summarized as follows:

- **PV array:** 56 IF-M340-60 polycrystalline PV modules (340 Wp each), arranged as 4 strings, with 14 modules connected in series per string, providing a total installed capacity of 19.04 kWp.
- **Inverters:** Two Solaxpower X3 hybrid inverters, rated at 10 kW each, for DC/AC power conversion.
- **Battery bank:** 24 SAKO 6-FMG-260 batteries (12 V, 260 Ah) connected in series for energy storage, linked to the building's loads through a source-transfer switch that keeps the SONELGAZ grid available only as a backup supply.

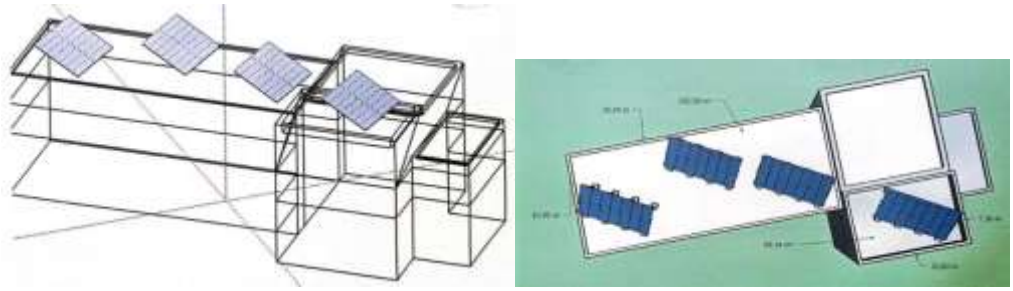


Fig. 1 Roof layout of the PV arrays (3D model and top view)

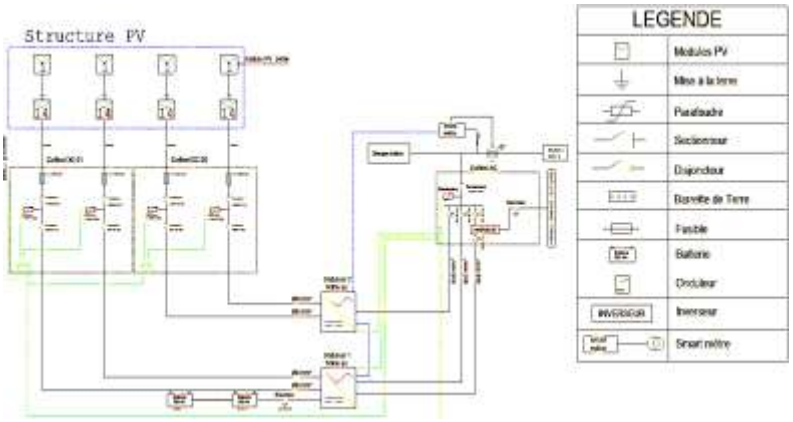


Fig. 2 Electrical wiring diagram of the 19 kWc PV installation

The PV array is oriented due south with a 32° tilt angle, corresponding approximately to the site latitude (36.59° N).

For the PVsyst simulation, the site was modelled with coordinates 36°56'24" N, 7°56'24" E, altitude 723 m, and a fixed array tilted 30° at 0° azimuth; under these conditions the site's annual global horizontal irradiation comes out to 1783.4 kWh/m², with an average ambient temperature of 15.9 °C [7].

Figure 1 presents the PV system layout in both 3D and top-view representations.

B. *Sizing and Evaluation Equations*

The array, battery bank and cabling were sized analytically before being modelled in PVsyst. Daily energy use of an appliance is given by:

$$Ec = P \times n \times Ti \quad (1)$$

where P is the appliance power (W), n the number of units, and Ti the daily hours of use. The energy the PV generator must supply is:

$$Ep = Ec / K \quad (2)$$

where K is a lumped loss coefficient (weather uncertainty, ageing, battery/inverter efficiency, cabling), typically between 0.55 and 0.85(we used $K = 0.85$).

Peak array power follows from:

$$Pc = Ep / Ir \quad (3)$$

where Ir is the site's average daily irradiation (kWh/m²/day). The panel count and string configuration are obtained from:

$$N = [Pc / Pc(panel)] \times Cp \quad (4)$$

$$Ns = Uond / Umod \quad (5)$$

$$Np = Pc / (Ns \times Pmod) \quad (6)$$

where Cp is a loss coefficient, Ns the modules per string (set by the inverter input voltage Uond and the module voltage Umod), Np the number of parallel strings, and Pmod the peak power of one module. Battery bank capacity is:

$$C = (Ec \times N) / (D \times U) \quad (7)$$

with N the required days of autonomy, D the maximum allowable depth of discharge, and U the battery's rated voltage. Module fill factor and efficiency are given by:

$$FF = Pmax / (Icc \times Vco)$$

$$\eta = Pmax / Pincident = (FF \times Icc \times Vco) / Pincident \quad (8)$$

Finally, the performance ratio used to evaluate the PVsyst simulation is:

$$PR = Yf / Yr = (E_Load / Pnom) / (GlobInc / Gstc) \quad (9)$$

where Yf is the final specific yield (kWh/kWc) and Yr the reference irradiation under STC conditions (Gstc = 1000 W/m²). Equation (9) is the main yardstick used to compare the theoretical output with the simulated one.

Table 1. Estimated daily energy consumption, from Eq. (1) (Wh/day)

Appliance	n	T (h)	Power (W)	Energy (Wh/day)
Air conditioner	3	8	2500	60,000
Computer	8	8	250	16,000
Printer	7	8	80	4,480
Light	22	8	54	9,504
Total	—	—	2884	89,984

Table 2. Sizing results for the 19 kWc PV system, from Eqs. (2)-(7)

Peak power (kWc)	No. of panels	Panels in series	Strings in parallel	Storage (Ah)	Cable section (mm²)
19.04	56	14	4	260	1.97

III. RESULTS

The installed peak power obtained analytically (19.04 kWc, Table 2) matches the system's rated 19 kWc almost exactly, which supports the sizing method described in Section II-B.

Over a full year, the simulation shows that the installation can supply about 31.80 MWh to the site's loads, a specific yield of 1670 kWh/kWc/year, and a performance ratio (Eq. 9) of 82.3%, calculated from GlobInc rather than GlobEff, which would give a slightly higher value. This is close to the estimated annual consumption of about 32.8 MWh derived from Table 1, so the system covers most of the office's demand on its own. Output peaks between June and August, when irradiation is highest, and drops to its lowest in December-January, as shown in Table 3.

Table 3. Monthly PVsyst balance (GlobHor: global horizontal irradiation; GlobEff: effective irradiation on the array; EArray: energy out of the PV field; E_Load: energy delivered to the site's loads; PR: performance ratio)

Month	GlobHor (kWh/m²)	T_Amb (°C)	GlobEff (kWh/m²)	EArray (MWh)	E_Load (MWh)	PR
January	75.9	6.41	116.5	2.020	1.962	0.859
February	87.3	7.57	114.2	1.972	1.915	0.857
March	136.8	11.09	158.2	2.690	2.615	0.844
April	164.9	13.59	170.7	2.869	2.788	0.834
May	202.5	18.63	190.1	3.153	3.066	0.820
June	221.9	23.25	199.9	3.253	3.162	0.803
July	244.9	27.18	224.7	3.584	3.489	0.790
August	209.2	26.06	210.7	3.384	3.295	0.798
September	158.6	21.13	180.1	2.947	2.868	0.815

October	124.2	18.16	162.8	2.688	2.615	0.823
November	85.6	11.18	126.2	2.154	2.093	0.846
December	71.5	7.68	114.4	1.985	1.927	0.854
Year	1783.4	16.05	1968.3	32.699	31.797	0.823

PVsyst's loss diagram traces the energy from the sunlight hitting the panels down to what actually reaches the site's loads. The biggest single loss is heat: the array loses 5.93% simply from running hotter than STC conditions. Cabling adds another 1.74%, module quality and mismatch together cost around 5%, and inverter conversion efficiency accounts for a further 2.70%. Starting from a nominal array output of 37.48 MWh, the system still delivers 31.80 MWh to the building's loads — total losses of roughly 15%, a fairly typical figure for this type of installation. PR (Eq. 9) swings between 0.79 in July and 0.86 in January, averaging 0.823 over the year. Temperature is the main driver: modules work less efficiently when it is hot and the sun is strongest, which is exactly when summer irradiation peaks, while winter months do better on PR despite receiving less sunlight overall. Figure 2 shows the complete electrical layout, connecting the PV arrays, DC/AC combiner boxes, both inverters, the battery bank and the smart meter, with the source-transfer switch keeping the SONELGAZ grid available only as a backup.

IV. Discussion

The hand-calculated sizing (19.04 kWc, 56 panels, Table 2) and the PVsyst model (19.04 kWc, 4 strings of 14 modules) agree almost exactly, which is reassuring on both fronts. The 1670 kWh/kWc/year yield also sits comfortably within the range typically reported for PV systems of similar size in north-eastern Algeria, so the site's solar resource is clearly being put to good use. As for the losses — thermal, resistive, inverter — none stand out as unusual, and together they do not drag the system's overall performance down in any meaningful way. The close match between the simulated annual output (31.8 MWh) and the site's estimated annual consumption (about 32.8 MWh, Table 1) also confirms that the installation was correctly sized to operate autonomously, supplying the SADEG office's loads directly and drawing on the SONELGAZ grid only as a backup. The close match between Eq. (3)-(4) results and the simulated configuration further suggests that the coefficient K used in Eq. (2) was a reasonable choice for this climate and load profile.

V. Conclusion

Sizing and then simulating the 19 kWc SADEG Souk-Ahras installation confirms that the PV array and battery system were sized correctly. The system should deliver around 31.8 MWh a year to the office's loads, closely matching its estimated annual consumption, at a performance ratio of 82.3%. This confirms that the installation can operate autonomously, with the SONELGAZ grid serving only as a backup source, and it speaks well for the technical soundness of the design and for how closely the PVsyst results track the initial hand calculations based on Eqs. (1)-(9).

References

[1] Ministry of Energy and Mines, *Renewable Energy Guide*. Algiers, Algeria: Ministry of Energy and Mines, 2007.

[2] Algerian Space Agency, *Solar Resource Atlas from Satellite Imagery*. Algiers, Algeria: Algerian Space Agency, 2019.

[3] PVsyst SA, "PVsyst software, version 6.79," PVsyst SA, Satigny, Switzerland, 2023. [Online]. Available: https://www.pvsyst.com/

[4] Ifrisol, "IF-M340-60 polycrystalline PV module datasheet," Ifrisol, 2021.

[5] Solaxpower, "X3-Hybrid-10.0kW inverter datasheet," Solaxpower Network Technology Co., Ltd., 2020.

[6] SAKO, "6-FMG-260 VRLA battery datasheet," SAKO Electric Co., Ltd., 2019.

[7] ESMAP, Global Solar Atlas 2.0 Technical Report. Washington, DC: World Bank, 2019. [Online]. Available: https://globalsolaratlas.info

[8] International Energy Agency (IEA), Renewables 2018 – Market Analysis and Forecast. Paris, France: IEA, 2018. [Online]. Available: https://www.iea.org/

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

3D Baskılı PLA/İletken PLA Çekirdek-Kabuk Kompozitlerin Eğilme ve Kırılma Davranışına Isıl İşlemin Etkisi

Dursun Ekmekci[1*], Emrah Kaplan [2]

[1]*Makine Mühendisliği Bölümü, Gümüşhane Üniversite, Türkiye*
[2]*R&D Engineer, Grünweg GmbH, Maintal, Germany*

[*]*(dursunekmekci@gumushane.edu.tr)*

Özet – Bu çalışmada, dış katmanı kırmızı yalıtkan polilaktik asit (PLA) ve iç çekirdeği siyah iletken PLA içeren kompozit numuneler, eriyik biriktirme modellemesi (FDM) kullanılarak üretilmiştir. Numunelerin mekanik özellikleri üç noktalı eğilme deneyleri kullanılarak belirlenmiştir. Fiziksel yaşlanma ve kristalleşmenin etkilerini incelemek için kompozitler 50°C'de 2, 4 ve 6 saat süreyle ısıl işleme tabi tutulmuştur. Yalıtkan PLA, keskin bir tepe sonrası kuvvet düşüşü ile gevrek kırılma sergilerken, iletken PLA önemli enerji dağılımıyla birlikte sünek bir kırılma modu göstermiştir. Fiziksel olarak yaşlanmış kompozitler daha sert ve daha az sünek olup, bu durum özellikle cam geçiş bölgesi civarında kısmen kristalleşmiş ve kısmen gevşemiş moleküllerin varlığına bağlanmıştır. Isıl işlem süresinin artmasıyla birlikte kompozitlerin rijitliği ve yük taşıma kapasitesi artarken süneklik ve deformasyon kabiliyeti azalmıştır. Özellikle 6 saatlik ısıl işlem uygulanan numunelerde daha belirgin sertleşme ve gevrekleşme davranışı gözlenmiştir. Buna karşılık, 2 saatlik ısıl işlem numunelerin mekanik davranışında sınırlı değişikliklere neden olurken, 4 saatlik işlem rijitlik ve enerji sönümleme kapasitesi arasında daha dengeli bir davranış sağlamıştır. Elde edilen sonuçlar, ısıl işlem süresinin FDM ile üretilmiş iletken PLA core–shell kompozitlerin mekanik performansını ve kırılma davranışını belirleyen önemli bir parametre olduğunu göstermektedir.

Anahtar Kelimeler – PLA, kompozit, eğilme testi, 3D yazıcı, FDM

I. GİRİŞ

Eklemeli imalat süreçleri ve özellikle eriyik biriktirme modellemesi (FDM), esneklikleri ve ekonomik olmaları nedeniyle fonksiyonel polimerik bileşenlerin üretiminde yaygın olarak kullanılmaktadır. FDM için, biyolojik olarak parçalanabilir bir termoplastik olan polilaktik asit (PLA), iyi boyutsal kararlılığı, nispeten kolay işlenebilirliği ve baskı sırasında kararlı termal özellikleri nedeniyle FDM işlemi için en popüler polimerik malzemelerden biridir [1-3].

Mekanik ve fonksiyonel özellikleri geliştirmek için, iletken polimer kompozitler oluşturmak üzere polilaktik asit matrisine karbon siyahı, grafen ve metal içeren katkı maddeleri gibi iletken dolgu maddeleri dahil edilmiştir. Bu kompozitler, gelişmiş mukavemet ve iletkenlik sağlamıştır. İletken olmayan dış ve iletken iç katmanlara dayalı kompozit tasarımda, PLA dış katmanı mukavemet sağlarken, iletken PLA iç katman süneklik sağlayarak performans özelliklerinin hibritleşmesine yol açar. Ek olarak, 55°C - 65°C cam geçiş sıcaklığına yakın gerçekleştirilen işlem, kompozitin kristalliğini, bağlanmasını ve özelliklerini etkiler [4-6].

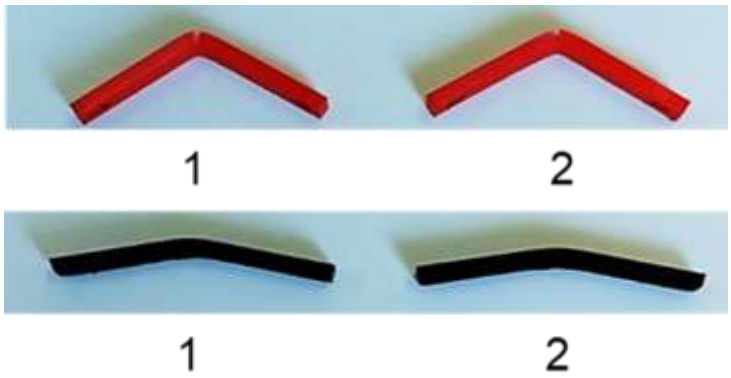
Sadece mekanik yük-sapma tepkilerine odaklanan geleneksel eğilme çalışmalarının aksine, bu çalışma, eğilme yükü altında 3 boyutlu baskı ile üretilen iletken PLA bazlı numunelerin mekanik davranışını araştırmaktadır. Eğilme testleri sırasında, bu çalışma, yalnızca mekanik ölçümlerle yakalanamayan hasar evrimi ve deformasyon mekanizmalarına ilişkin ek bilgiler sağlamayı amaçlamaktadır. Bu araştırma, 3 boyutlu baskı ile üretilen PLA/iletken çekirdek-kabuk kompozitleri arasındaki mekanik ilişkiyi araştırmayı amaçlamaktadır.

II. MATERYAL VE YÖNTEM

Deneysel sonuçların tekrarlanabilirliğini ve karşılaştırılabilirliğini sağlamak için numune geometrisi, boyutları ve baskı parametreleri sabit tutulmuştur. Temel malzeme, yalıtkan kırmızı PLA ve iletken siyah PLA filamentleriydi. 3D baskı işlemi, çift ekstrüzyon özelliğine sahip Raise3D N2 3D yazıcı kullanılarak gerçekleştirilmiştir. Bir miktar tekrarlanabilirlik sağlamak için, katmanların yüksekliği, %100 malzeme doluluk yoğunluğu ve 230°C ekstrüder sıcaklığı tüm malzeme için standart olarak belirlenmiştir. Bu çalışmada kullanılan malzemeler, ticari bir kırmızı yalıtkan PLA filamenti ve siyah iletken bir PLA filamentinden oluşmaktadır. Kırmızı PLA'nın yoğunluğu yaklaşık 1,24 g/cm³, çekme dayanımı 60–65 MPa, elastik modülü yaklaşık 3,5 GPa ve kopma uzaması %6–10 idi. Buna karşılık, iletken PLA'nın yoğunluğu 1,1 ila 1,3 g/cm³ arasında, çekme dayanımı 20–45 MPa, elastik modülü 1,5–2,5 GPa ve kopma uzaması %1–3 idi. Önerilen baskı sıcaklığı aralıkları, kırmızı ve iletken PLA filamentleri için sırasıyla 190–220 °C ve 210–230 °C idi. 3D baskılı malzemenin tavlanması, 50°C'lik kontrollü bir sıcaklığa sahip bir fırın kullanılarak gerçekleştirilmiştir. Üç farklı tavlama süresi (2, 4 ve 6 saat) uygulanmıştır. Bu değer, polilaktik asidin cam geçiş sıcaklığı aralığına yaklaştığı ve böylece fiziksel yaşlanmayı kolaylaştırdığı, ancak malzemenin boyutlarının sapmasına neden olacak aşırı kristalleşmeyi teşvik etmediği için seçilmiştir. Üç noktalı eğilme testleri, ISO 178 standardına uygun olarak, 5 mm/dk çapraz kafa hızıyla BESMAK üniversal test makinesi kullanılarak gerçekleştirildi.

III. BULGULAR

Şekil 1'de gösterildiği gibi, yalıtkan kırmızı PLA numunesi, eğilme testinden sonra ani ve keskin bir kırılma sergiledi. Bu, PLA'nın yüksek amorf içeriği ve malzemenin sergilediği düşük deformasyon aralığından kaynaklanan kırılgan doğasını açıkça göstermektedir. Bununla birlikte, PLA, düşük plastik deformasyon yeteneği sergileyen bir termoplastiktir. Öte yandan, iletken siyah PLA malzemesi kırılmadı, ancak malzeme deformasyonu sergiledi. Bu, karbon bazlı iletken malzeme katkı maddelerinin varlığıyla bağlantılıdır. Bu katkı maddeleri, malzeme gerilimlerini azaltmak için mikro ölçekli malzeme yeniden yönlendirmesine izin vererek malzemenin tokluğunu artırır. İletken dolgu maddeleri, malzeme yükünü kısmen taşır, bu da kırılma sürecini yavaşlatır ve malzemenin kolayca mod değiştirmesine ve sünek bir şekilde deforme olmasına olanak tanır. Özetle, yalıtkan PLA malzemesi, iletken PLA malzemesinin aksine, malzeme kırılganlığı sergilemiştir; iletken PLA malzemesi ise sünek malzeme kırılması sergilemiştir. Bu, iletken dolgu maddelerinin malzeme matrisinin malzeme tokluğunu ve iletkenliğini artırdığını açıkça doğrulamaktadır.



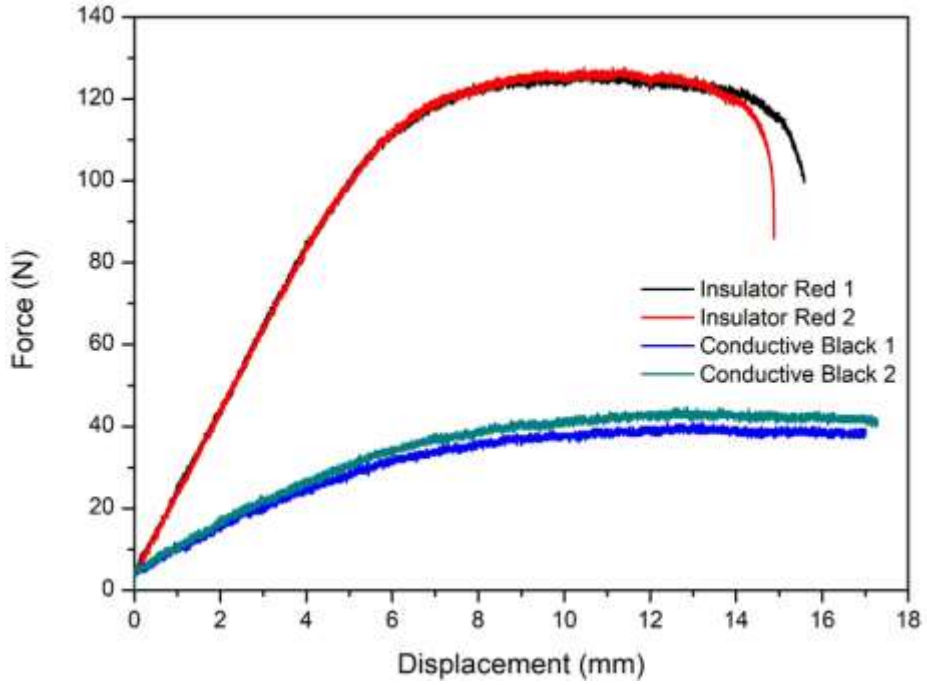
Şekil 1. Kırmızı PLA ve siyah PLA/İletken eğilme numuneleri

IV. TARTIŞMA

A. PLA Malzemeleri

Şekil 2'de kırmızı PLA ve siyah PLA/iletken kompozit için kuvvet-yer değiştirme eğrileri gösterilmiştir. Kırmızı PLA için 125-130 N olan maksimum kuvvetten sonra, kuvvet aniden düşmektedir. Bununla birlikte, siyah PLA/iletken kompozit için, maksimum kuvvet değerleri 40-45 N'de önemli ölçüde daha düşük olmasına rağmen, karşılık gelen yer değiştirme değerleri kırmızı PLA'ya kıyasla çok daha büyüktür. Bu daha geniş deformasyon aralığı, malzemenin sünekliğini temsil eder. Bu süneklik, iletken fazın varlığı ve matris çevresindeki gerilimlerin azalmasıyla açıklanabilir; bu da çatlak yayılımını

geciktirir. Sonuç olarak, malzeme artan süneklik ve daha düşük sertlik gösterir. Bu nedenle, malzemenin kırılması, kırmızı/yalıtkan PLA'nın daha güçlü ve daha az sünek olduğunu, ancak siyah/iletken PLA'nın daha az güçlü ve daha sünek olduğunu göstermektedir. Bu, iletken malzemenin kullanımının sertliği azalttığını ancak malzemenin deformasyon kabiliyetini artırdığını göstermektedir. Malzeme mühendisliğinde, kırmızı yalıtkan PLA, yüksek rijitlik gerektiren uygulamalar için daha uygundur; siyah iletken PLA ise daha fazla esneklik ve elektriksel iletkenlik gerektiren tasarımlar için tercih edilir [7]. İki malzeme için kuvvet-yer değiştirme grafiği, yalıtkan ve iletken PLA ile ilişkili tipik kırılgan ve sünek davranışın niteliksel bir tanımını vermek üzere Şekil 2'de verilmiştir. Bu nedenle, şekil malzemelerin eğilme dayanımının analizi olarak kullanılmamalıdır.

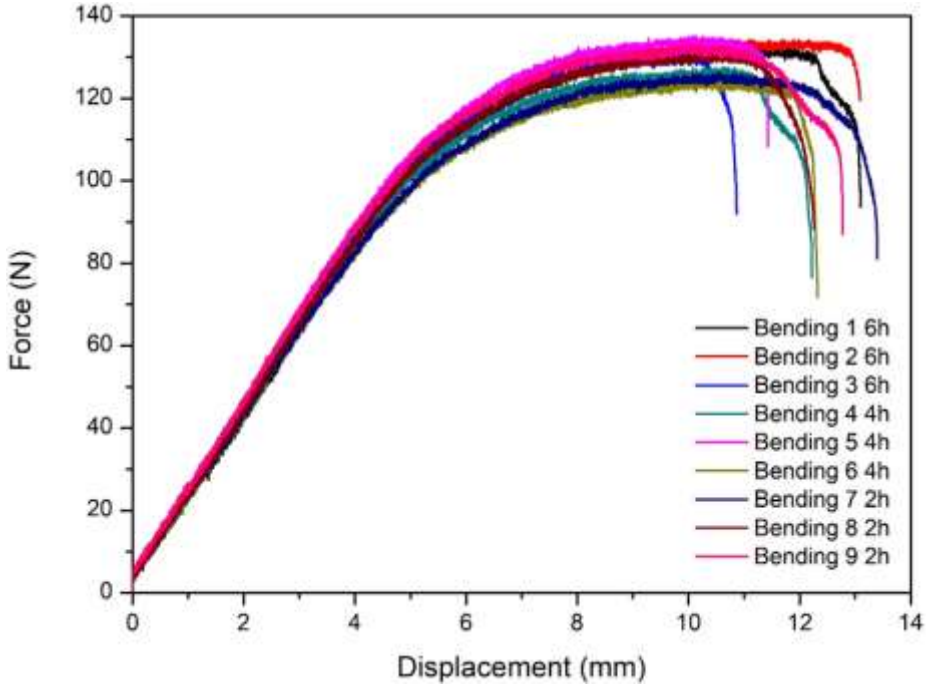


Şekil 2. Kırmızı/yalıtkan PLA ve siyah/iletken PLA'nın kuvvet-yer değiştirme grafiği

B. PLA Kompozit Malzemelerin Isıl İşlemi

50°C'de uygulanan ısıl işlem, kompozit polilaktik asidin mikroyapısal ve mekanik davranışını önemli ölçüde etkiler. Isıl işlem süresi arttıkça, yapısal gevşeme daha belirgin hale gelirken, mekanik davranışta gözlemlenen değişiklikler, literatürde PLA sistemleri için bildirilen fiziksel yaşlanma ve olası kısmi kristalleşme ile tutarlıdır [8, 9]. Bu mikroyapısal değişiklikler, elastik modülün, kısa süreli yük taşıma kapasitesinin ve arayüzey bağ dayanımının artmasına, kopma uzamasının ve genel sünekliğin azalmasına yol açar. Ayrıca, boyutsal kararlılık da iyileşir. 50°C, PLA'nın cam geçiş sıcaklığına (55-65°C) yakın olduğundan, daha uzun ısıl işlem süreleri moleküler yeniden düzenlemeyi teşvik eder ve serbest hacmi azaltarak sertliği artırır ancak sünekliği azaltır. Kompozit içinde dağılmış iletken parçacıklar, çekirdeklenme ajanları görevi görerek, nispeten düşük sıcaklıklarda bile uzun işlem süreleri boyunca kristalleşmeyi teşvik eder. Sonuç olarak, kompozitin elastik modülü, eğilme dayanımı, kesme dayanımı ve yapısal kararlılığı iyileşir. Ayrıca, ısıl işlem, katmanlar arası bağlamayı güçlendirir ve yeniden akışla ilgili kusurları azaltarak çekirdek-kabuk arayüzü boyunca yük transferini iyileştirir. Sonuç olarak, kompozitin genel makroskobik mekanik davranışı ve yapısal kararlılığı artarken, delaminasyon etkileri azalır. Bu iyileştirmeler, iletken parçacıklar tarafından kolaylaştırılan daha etkili gerilim transferine ve iyileştirilmiş arayüzey bağlamasına atfedilir [10]. Fiziksel yaşlanma ve kristalleşme derecesi, 6 saatlik ısıl işleme tabi tutulan numunelerde en belirgindi. Bu numuneler, artan kırılganlık ve azalan uzama ile birlikte en yüksek elastik modül ve maksimum kuvvet değerlerini sergiledi. Bu numunelerde gözlemlenen daha pürüzsüz ve daha yansıtıcı kırılma yüzeyleri, literatürde termal olarak işlenmiş PLA bazlı malzemeler için bildirilen kırılgan kırılma özellikleri ve iyileştirilmiş arayüzey yapışmasıyla tutarlıdır [11, 12]. 4 saat işlem gören numunelerde de fiziksel yaşlanma etkileri fark edildi. Bu numuneler, azalan uzama ile birlikte nispeten yüksek elastik modül ve maksimum kuvvet değerleri sergiledi. Maksimum kuvvete ulaşıldıktan sonra, kuvvet-yer değiştirme eğrilerinde keskin bir düşüş gözlemlendi ve bu da artan

kırılganlığı gösterdi. Ayrıca, daha yüksek arayüzey yapışması, lokalize yüzey kaynaklı çatlamadan yaygın kırılmaya doğru bir kaymaya eşlik eder. Buna karşılık, 2 saat işlem gören numuneler sınırlı mikro kristalleşme ve minimum fiziksel yaşlanma gösterdi. Bu numuneler, nispeten daha yüksek uzama ve eğilme tokluğunu korurken, daha düşük elastik modül ve maksimum kuvvet değerleri sergiledi. Sonuç olarak, bu numunelerin mekanik davranışı, işlem görmemiş numunelerin davranışına daha yakındı [13]. Genel olarak, numunelerin mekanik performansı, 50°C'deki ısıl işlem süresiyle sistematik olarak değişti. Kısa süreli işlem (2 saat) minimum değişikliklere neden olurken, orta ve uzun süreli işlemler (4 saat ve 6 saat) sertliği ve yük taşıma kapasitesini önemli ölçüde artırdı. Bununla birlikte, kopma uzaması işlem süresi arttıkça azaldı ve tepe sonrası yumuşama daha ani hale geldi, bu da artan kırılganlığı gösteriyor. Daha uzun süre ısı işlemine tabi tutulan numunelerde gözlemlenen daha düzgün kırılma yüzeyleri, uzatılmış ısı işlemi süreleriyle ilişkili kırılganlıktaki artışı daha da desteklemektedir. Şekil 3'te gösterildiği gibi, ısı işlemi prosedürünün 50°C uzatılması, elastik kısmın sertleşmesinde ve maksimum yük taşıma kapasitesinde küçük de olsa bir artışa yol açmaktadır. Ayrıca, akma sonrası/maksimum sonrası ve süneklikte bir azalma gözlemlenmiştir. Numuneler için genel benzerlik yüksek görünse de, tepe sonrası kuvvet dağılımı özellikleri, ani ve kademeli kırılmalar gibi değişen zamanları göstermekte ve içermektedir. Çeşitli eğrilerin orijinal düz kısımlarının yakın örtüşmesi, genel sertleşme değerlerinin neredeyse aynı olduğunu göstermekte ve bu da baskı işlemi sırasında etkili benzerliği ve geometriyi doğrulamaktadır.



Şekil 3. Çeşitli tavlama sürelerine (50 °C'de 2 saat, 4 saat ve 6 saat) maruz bırakılan kompozit numunelerin kuvvet-yer değiştirme eğrileri

V. SONUÇLAR

Bu çalışma, bükülme kaynaklı elektriksel empedans değişimlerinin, 3 boyutlu baskı ile üretilen iletken PLA bazlı kompozitlerdeki mekanik deformasyonu izlemek için etkili bir şekilde kullanılabileceğini göstermiştir. Kırmızı PLA gevrek kırılma davranışı sergilerken, iletken PLA daha sünek bir tepki göstermiştir. Çekirdek-kabuk yapısı, tokluk ve sertlik arasında güzel bir denge sağlamıştır. Mekanik özellikler, 50 °C'lik ısıl işlemden büyük ölçüde etkilenmiş, sertlik artmış ve süneklik azalmıştır. Isıl işlem süresi açısından sertlik ve enerji dağılımı arasında en iyi uzlaşma 4 saat sonra elde edilmiştir.

TEŞEKKÜR

Yazarlar, Gümüşhane Üniversitesi Geliştirme Vakfı'na (Proje numarası: 25.F5112.01.01) verdikleri destek için teşekkür ederler.

KAYNAKLAR

[1] M. Hussain, S. M. Khan, M. Shafiq M, and N. Abbas, "A review on PLA-based biodegradable materials for biomedical applications," *Giant*, vol. 18, p. 100261, Jun. 2024.

[2] S. Farah, D. G. Anderson, and R. Langer, "Physical and mechanical properties of PLA, and their functions in widespread applications — A comprehensive review," *Adv. Drug Deliv. Rev.*, vol. 107, pp. 367-392, Dec. 2016.

[3] I. Plamadiala, C. Croitoru, M. A. Pop, and I. C. Roata, "Enhancing polylactic Acid (PLA) performance: A review of additives in fused deposition modelling (FDM) filaments," *Polymers*, vol. 17(2), p. 191, Jan. 2025.

[4] I. Tirado-Garcia, D. Garcia-Gonzalez, S. Garzon-Hernandez, A. Rusinek, G. Robles, J. M. Martinez-Tarifa, and A. Arias, "Conductive 3D printed PLA composites: On the interplay of mechanical, electrical and thermal behaviours," *Compos. Struct.*, vol. 265, p. 113744, Jun. 2021.

[5] O. Masiuchok, M. Iurzhenko, R. Kolisnyk, Y. Mamunya, M. Godzierz, V. Demchenko, D. Yermolenko, and A. Shadrin, "Polylactide/carbon black segregated composites for 3D printing of conductive products," *Polymers*, vol. 14(19), p. 4022, Sep. 2022.

[6] E. Ivanov, R. Kotsilkova, H. Xia, Y. Chen, R. K. Donato, K. Donato, A. P. Godoy, R. D. Maio, C. Silvestre, S. Cimmino, and V. Angelov, "PLA/Graphene/MWCNT composites with improved electrical and thermal properties suitable for FDM 3D printing applications," *Appl. Sci.*, vol. 9(6), p. 1209, Mar. 2019.

[7] C. Rubio-González, J. J. Ku-Herrera, J. Wang, A. Patterson, J. A. Soto-Cajiga, and O. Olvera-Silva, "Effect of printing parameters on the mechanical and piezoresistive response of cellular structures manufactured with a conductive polylactic acid nanocomposite through a material extrusion process," *J. Manuf. Process.*, vol. 132, pp. 375-391, Dec. 2024.

[8] A. D. Virag, C. Toth, P. Polyak, M. Musioł, and K. Molnár, "Tailoring the mechanical and rheological properties of poly(lactic acid) by sterilizing UV-C irradiation," *Int. J. Biol. Macromol.*, vol. 277(Part 3), p. 134247, Oct. 2024.

[9] L. Abbane, N. Khemici-Doulache, M. W. Khemici, and M. Atma, "Investigation of the physical aging effect on polylactic acid/polymethyl methacrylate biomaterial blends," *Phys. Scr.*, vol. 101(19), p. 195902, May. 2026.

[10] J. Huang, J. Zeng, Y. Bai, Z. Cheng, Y. Wang, Q. Zhao, and D. Liang, "Effect of adhesive layer properties on the shear strength of single-lap structures of dissimilar materials based on the cohesive zone model," *J. Mech. Sci. Technol.*, vol. 35, pp. 133-143, Jan. 2021.

[11] X. Zhao, J. Liu, J. Li, X. Liang, W. Zhou, and S. Peng, "Strategies and techniques for improving heat resistance and mechanical performances of poly(lactic acid) (PLA) biodegradable materials," *Int. J. Biol. Macromol.*, vol. 218, pp. 115-134, Oct. 2022.

[12] G. Wang, D. Zhang, B. Li, G. Wan, G. Zhao, and A. Zhang, "Strong and thermal-resistance glass fiber-reinforced polylactic acid (PLA) composites enabled by heat treatment," *Int. J. Biol. Macromol.*, vol. 129, pp. 448-459, May. 2019.

[13] W. Zhu, Q. Wu, L. Yanga, and Y. C. Zhou, "In situ characterization of high temperature elastic modulus and fracture toughness in air plasma sprayed thermal barrier coatings under bending by using digital image correlation," *Ceram. Int.*, vol. 46(11)(Part B), pp. 18526-18533, Aug. 2020.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Madencilikte Yorgunluk Tespiti ve Yönetimi

Zekeriya Duran[1*] ve Bülent Erdem[2]

[1]*Madencilik ve Maden Çıkarma Bölümü, Sivas Teknik Bilimler MYO, Sivas Cumhuriyet Üniversitesi, Türkiye*
[2]*Jeofizik Mühendisliği Bölümü, Sivas Cumhuriyet Üniversitesi, Türkiye*

*(zduran@cumhuriyet.edu.tr)

Özet **–** Madencilik operasyonlarında yorgunluk, olaylar, kazalar ve operasyonel verimliliğin azalmasına önemli ölçüde katkıda bulunan kritik bir güvenlik sorunudur. Bu literatür incelemesinde, madencilik sektöründe yorgunluk tespiti ve yönetimine ilişkin araştırmalardan elde edilen bulgular sentezlenmektedir. İnceleme; makine görüş sistemleri, giyilebilir IoT cihazları, fizyolojik algılayıcılar ve davranışsal izleme platformları dâhil olmak üzere yorgunluğun izlenmesine yönelik teknolojik yaklaşımları ele almaktadır. Temel bulgular, tespit teknolojileri önemli ölçüde ilerlemiş olsa da (sistemler kontrollü ortamlarda %90-95 doğruluk oranlarına ulaşmaktadır) pratik uygulamanın çevresel koşullar, operatörlerin kabullenmesi ve sistem güvenilirliği ile ilgili zorluklarla karşılaştığını ortaya koymaktadır. İnceleme, üç ana teknoloji kategorisi olan kamyon operatörleri için görüş tabanlı sistemleri, yeraltı işçileri için giyilebilir fizyolojik monitörleri ve tümleşik yorgunluk yönetim platformlarını tanımlamaktadır. Kanıtlar, sirkadiyen (yaklaşık bir gün veya 24 saatlik döngü) ritimlerin ve vardiya düzenlerinin yorgunluk üzerinde vardiya süresinden daha güçlü bir etkiye sahip olduğunu ve ardışık 13-14 gecelik vardiyanın en yüksek yorgunluk düzeylerini ortaya çıkardığını göstermektedir. İnceleme sonunda, etkili yorgunluk yönetiminin teknolojik izlemeyi kurumsal politikalar, operatör eğitimi ve işyeri kültürü ile birleştiren sosyo-teknik bir sistem yaklaşımı gerektirdiği sonucuna varılmaktadır. Geleceğe yönelik yönelimler arasında yapay zekâ destekli tahmin sistemleri, çoklu algılayıcı kaynaşımı ve kişisel yorgunluk risk değerlendirme çerçevesi yer almaktadır.

Anahtar Kelimeler – sirkadiyen ritmi, bilişsel durum, vardiya düzeni, giyilebilir cihaz, örgüt kültürü

I. GİRİŞ

Madencilik sektörü; işçileri ciddi yorgunluk risklerine maruz bırakan oldukça zorlu koşullar altında faaliyet göstermektedir. Uzun vardiya süreleri ve gece çalışmaları, uzak lokasyonlarda görev alma, monoton ve tekrarlayıcı işlerin yoğunluğu, yetersiz aydınlatma, yüksek sıcaklık ve gürültü gibi çevresel faktörler bu riskleri artırmaktadır. Buna ek olarak uyku düzeninin bozulması ve fiziksel açıdan son derece zorlayıcı görevler, madencilerin hem bedensel hem de zihinsel tükenmişlik yaşamasına yol açmaktadır. Tüm bu çok boyutlu etkenler birleştiğinde, yorgunluk yalnızca bireysel sağlık açısından değil, aynı zamanda iş güvenliği ve üretkenlik açısından da kritik bir sorun haline gelmektedir [1]. Yorgunluğa bağlı bilinç bozulması, operatörlerin uyanıklık durumunu, karar verme kapasitesini ve tepki sürelerini olumsuz etkileyerek kazalara ve ramak kala olaylara önemli ölçüde katkıda bulunabilmektedir. Madencilik operasyonlarının giderek otomasyon ve gelişmiş izleme teknolojilerini benimsemesiyle birlikte, yorgunluk tespiti ve yönetim sistemlerinin mevcut durumunu anlamak, işçi güvenliğini ve operasyonel verimliliği artırmak açısından hayati önem taşımaktadır [2], [3].

Bu literatür incelemesiyle, özellikle madencilik özelinde geliştirilen veya uygulanan yorgunluk tespiti teknolojileri ve yönetim stratejileri üzerine yapılan araştırmalar sentezlenmektedir. Bu kapsamda; madencilikte yorgunluk tespitine yönelik teknolojik yaklaşımların betimlenmesi, mevcut sistemlerin etkinlik ve sınırlamalarının değerlendirilmesi, operasyonel faktörlerin yorgunluk üzerindeki etkilerinin incelenmesi ve ayrıca gelecekteki araştırma yönleri ile endüstri uygulamalarına ilişkin pratik önerilerin belirlenmesi amaçlanmıştır.

II. ALTYAPI VE KURAMSAL TEMEL

Madencilik faaliyetleri, diğer ulaşım veya endüstriyel sektörlere kıyasla yorgunluk izlemede büyük zorluklar sunmaktadır. Açık maden ocaklarındaki kamyon operatörleri, tozlu ve yüksek titreşimli ortamlarda uzun ve monoton sürüşler gerçekleştirirken; yeraltı madencileri ise dar çalışma alanları, sınırlı aydınlatma, yüksek sıcaklık ve fiziksel açıdan zorlayıcı görevler altında çalışmaktadır. Bu çevresel ve operasyonel koşullar, yorgunluğun değerlendirilmesi ve yönetimini daha karmaşık hâle getirmektedir [4], [5]. Bu bağlamsal faktörler, madenciliğe özgü operasyonel gereksinimlere göre uyarlanmış özel tespit yaklaşımlarını zorunlu kılmaktadır.

Meslek ortamlarındaki yorgunluk, yalnızca bir vardiya süresince ağır görevler ve uzun süre uyanık kalmaya bağlı akut yorgunluğu değil, aynı zamanda ardışık vardiyalar veya çalışma döngüleri boyunca biriken kronik yorgunluğu da kapsamaktadır [6], [7]. Madencilik sektöründe yaygın olarak görülen uzun vardiya düzenleri (genellikle 10–12 saat) ve uçakla gel–uçakla dön (FIFO) tipi vardiyalar, her iki yorgunluk türünün ortaya çıkması için elverişli koşullar yaratmaktadır [5], [8]. Yorgunluk mekanizmalarının anlaşılması ise tek bir faktörle açıklanamayacak kadar karmaşıktır. Bunun için sirkadiyen ritimler, uyku eksikliği, iş yükü talepleri ve bireysel duyarlılık farklılıkları gibi birden fazla etkileşimli unsurun birlikte değerlendirilmesi gerekmektedir [9], [10].

Madencilikte yorgunluk üzerine yapılan araştırmalar, öznel özbildirim ölçümlerinden objektif, gerçek zamanlı izleme sistemlerine doğru evrilmiştir. İlk yaklaşımların, yaşam tarzı ve sağlık faktörlerini değerlendiren yorgunluk riski anketlerine dayandığı; ancak bu yöntemlerin, bireyin öz değerlendirmesine bağlı olmaları ve sürekli izleme sağlayamamaları nedeniyle gerçek zamanlı müdahale için gerekli zamansal çözünürlüğe sahip olmadığı bildirilmiştir [2], [7], [11]. Otomatik tespit sistemlerinin geliştirilmesi ise performans düşümü veya güvenlik açısından kritik olaylar meydana gelmeden önce müdahaleyi mümkün kılan proaktif yorgunluk yönetimi anlayışına geçişi temsil etmektedir [5], [12].

Madencilikte yorgunluk tespiti için teorik yaklaşımlar birden fazla disipline dayanmaktadır. Biyomekanik modeller tekrarlayan görevler ve donanım işletiminden kaynaklanan fiziksel yorgunluğun anlaşılmasına katkıda bulunurken [13], [14], bilişsel psikoloji yorgunluk altında dikkat ve uyanıklıkta azalma ile bilgi işleme modellerine ilişkin çalışmalara katkı sağlamaktadır [15]. Sirkadiyen biyoloji, günün saatine bağlı performans değişimleri ile vardiyalı çalışmaya fizyolojik uyum süreçlerinin anlaşılmasına katkı sağlamaktadır. Bunun yanında, yorgunluğun yalnızca bireysel biyolojik faktörlerle değil; çalışma organizasyonu, operasyonel süreçler ve teknolojik destek sistemleriyle birlikte değerlendirilmesi gerekmektedir. Bu perspektiflerin bütünleşmesi, insan, teknoloji ve organizasyon etkileşimlerini birlikte ele alan sosyo-teknik sistem yaklaşımını desteklemekte ve etkili yorgunluk yönetiminin temelini oluşturmaktadır [5], [7], [9], [16].

Madencilik sektörünün operasyonel özellikleri, farklı yorgunluk risk profilleri ortaya çıkarmaktadır. Açık ocaklarda çalışan kamyon operatörleri, aynı rotalarda uzun süreli ve tekrarlayıcı sürüşler gerçekleştirerek fiziksel iş yükü nispeten düşük olsa da monotonluk, sürekli titreşim ve çevresel koşullar nedeniyle pasif yorgunluğa neden olduğu bilinmektedir [4], [5], [17], [18]. Buna karşılık, yeraltı madencileri dar çalışma alanları, karmaşık galeri yapıları, yüksek fiziksel iş yükü, sıcaklık stresi ve artan bilişsel talepler altında çalışmakta olup bu koşullar farklı yorgunluk mekanizmalarını beraberinde getirmektedir [1], [19], [20]. Bu nedenle açık ocak ve yeraltı madenciliğinde yorgunluk tespit ve yönetim sistemlerinin operasyonel bağlama uygun şekilde özelleştirilmesi gerekmektedir [5].

III. TESPİT TEKNOLOJİLERİ VE YÖNTEMLER

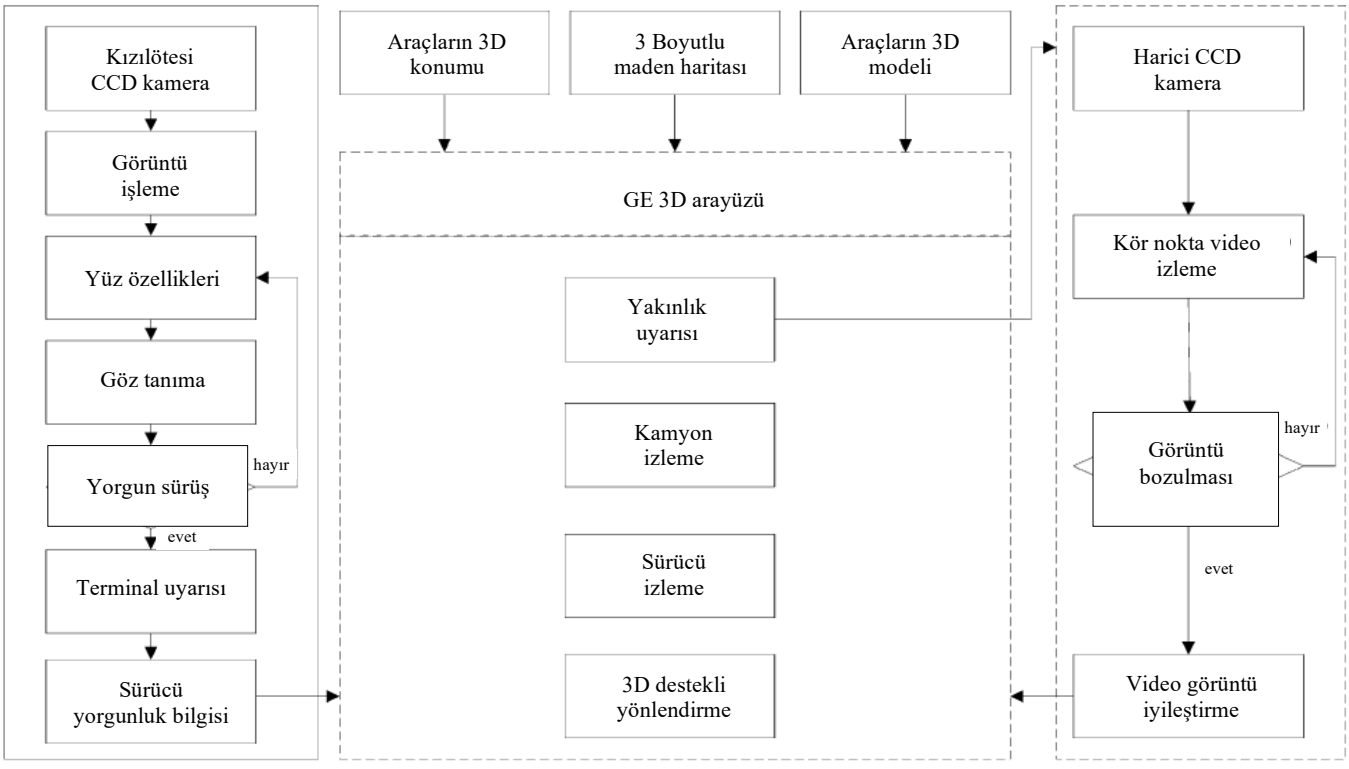
A. Görsel Tabanlı Algılama Sistemleri

Görsel tabanlı yorgunluk tespiti, açık ocak madenciliğinde ağır araç operatörlerinin izlenmesi için en yaygın araştırılan yaklaşımlardan biridir. Bu sistemler, araç kabinine yerleştirilen kameralar aracılığıyla operatörün yüz özelliklerini ve yorgunlukla ilişkili davranışsal göstergeleri izlemekte; göz kapanma süresi (PERCLOS), göz kırpma oranı, esneme davranışı ve baş pozisyonundaki değişimler gibi parametrelerden yararlanarak gerçek zamanlı yorgunluk değerlendirmesi yapmaktadır [17], [18].

Görsel tabanlı sistemlerin temel mimarisi, görüntü yakalama, özellik çıkarma, yorgunluk durumunun sınıflandırılması ve uyarı oluşturmayı içermektedir. Açık ocak kamyon sürücülerinde yorgunluk tespitini

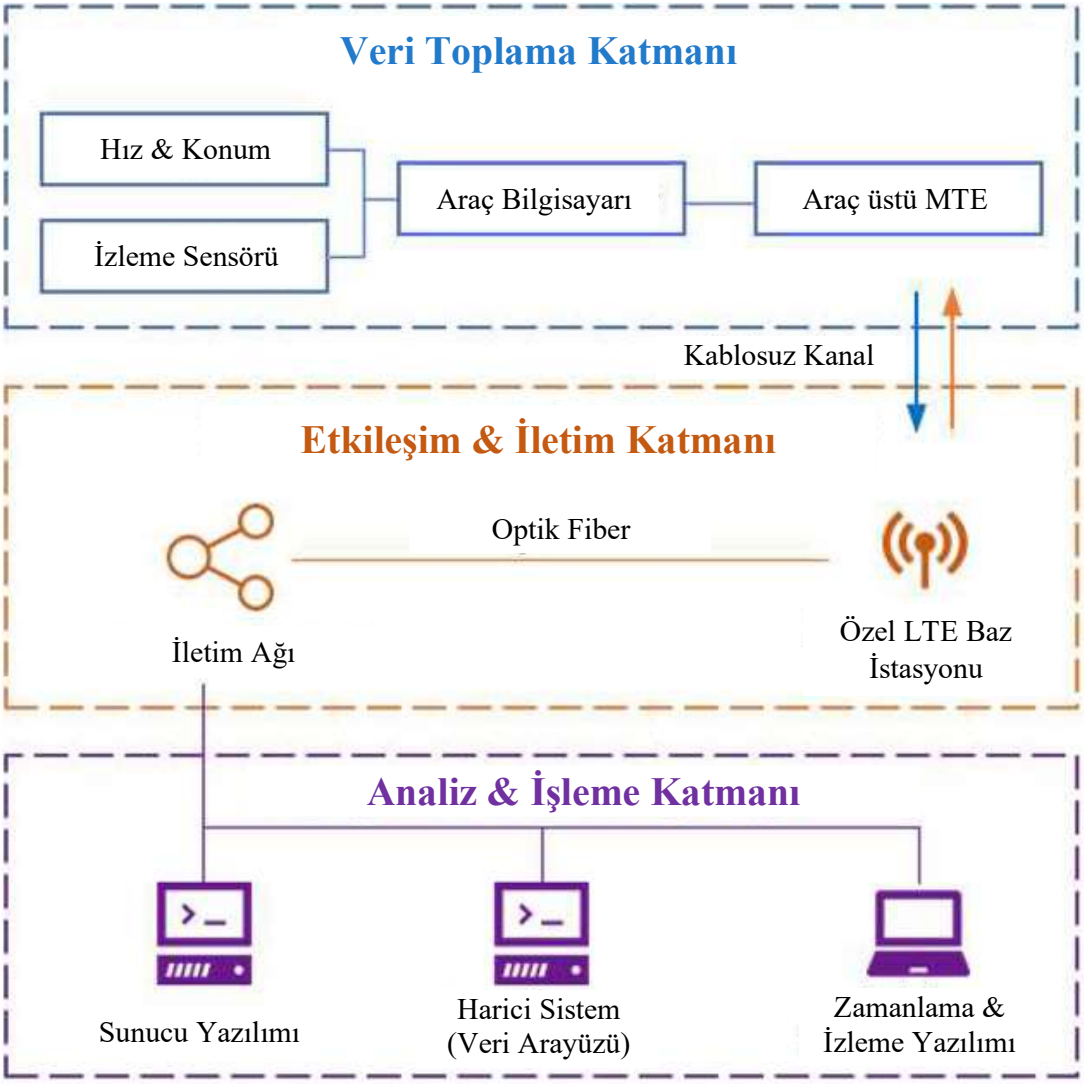
görsel bilgiye dayalı yöntemlerle gerçekleştirmek için bir çalışma yapılmıştır. Çalışma, klasik yöntemlerin yalnızca tek kare görüntüye dayalı olması nedeniyle zamansal bilgiyi yeterince kullanamamasını eleştirmekte ve bu eksikliği gidermek için **Libfacedetection tabanlı yüz tespiti ve takip algoritması** ile **CNN + LSTM (LRCN)** birleşimini kullanmaktadır. Geliştirilen sistem, sürücülerin göz kırpma, göz kapağı kapanma ve esneme gibi davranışsal göstergelerini uzun süreli video dizileri üzerinden analiz ederek yorgunluk sınıflandırması yapmıştır. Çalışmada; yüz tespiti ve takibi, özelliklerin kodlanması ve zamansal sınıflandırma modüllerinden oluşan üç aşamalı bir yöntem geliştirilmiştir. Deneysel sonuçlar, önerilen modelin kendi veri setinde %99,3 doğruluk oranına ulaştığını göstermiştir. Elde edilen bulgular, madencilik sektöründe kullanılan açık ocak kamyonları için sürücü yorgunluğunu gerçek zamanlı ve yüksek doğrulukla tespit edebilen bir sistemin mümkün olduğunu ortaya koymuştur [21].

Yorgunluk tespitine yönelik makine görüşü algoritmaları, başlangıçta basit eşik tabanlı yöntemlere dayanırken, zamanla gelişmiş makine öğrenmesi yaklaşımlarına doğru ilerlemiştir. Açık maden işletmelerinde kamyon kazalarının sık görülmesi ve bu kazaların genellikle sürücü yorgunluğu, düşük görüş koşulları, yol durumu ve sürücü davranışları gibi faktörlerden kaynaklandığını vurgulanmıştır. Önerilen sistem; **Google Earth tabanlı 3D grafik arayüzü**, **GPS ile kamyon takibi**, **kablosuz iletişim ağı**, **sürücü yorgunluk tespit sistemi** ve **görüş iyileştirme modülü** gibi bileşenlerden oluşmaktadır (Şekil 1). Yorgunluk tespiti için sürücünün yüz ve göz hareketleri kızılötesi kameralarla izlenmekte, PERCLOS algoritmasıyla göz kapağı kapanma yüzdesi hesaplanarak uyarılar verilmektedir. Görüş iyileştirme kısmında ise CCD kameralar ve görüntü işleme algoritmaları kullanılarak sis, yağmur, toz gibi olumsuz hava koşullarında görüntülerin netleştirilmesi amaçlanmıştır. Çalışma maden kamyonlarında kazaları azaltmak için **yorgunluk tespiti, yakınlık uyarı sistemi ve görüş iyileştirme teknolojilerini** tümleşik bir sürüş sisteminde birleştirmeyi önermiştir. Böylece sürücülerin dikkat seviyesinin korunması, kör noktaların azaltılması ve düşük görüş koşullarında güvenli sürüş sağlanması hedeflenmiştir [22]. İzleyen çalışmalarda, yorgunluk durumları arasındaki belirsizliği daha iyi temsil etmek amacıyla bulanık mantık ve bulanık sinir ağı tabanlı sınıflandırma yaklaşımları önerilmiştir. Bu yöntemler, yorgunluğu keskin kategoriler yerine süreklilik gösteren bir durum olarak modelleyerek daha hassas sınıflandırma ve kademeli uyarı stratejilerinin geliştirilmesine olanak sağlamaktadır [23].



Şekil 1. Maden kamyonu yardımcı sürüş sisteminin yapısı [22]

Li vd. [24], açık ocak madenciliğinde kullanılmak üzere Genelleştirilmiş Yorgunluk Sürüş İzleme Sistemi (*Generalized Fatigue Driving Monitoring System, GFDMS*) geliştirmiş ve bu sistemi Myanmar'daki bir bakır açık ocağında başarıyla uygulamıştır. Sistem; sürücü durum izleme (DSM) kameraları, konumlandırma sensörleri, hız algılayıcıları ve merkezi veri analiz modüllerini bir araya getirerek farklı araç tiplerinde gerçek zamanlı yorgunluk ve güvenli olmayan sürüş davranışlarının izlenmesini sağlamıştır (Şekil 2). Modüler ve genelleştirilmiş mimarisi sayesinde, önceki çalışmalarda belirli araç veya saha koşullarına özgü geliştirilen sistemlere kıyasla daha yüksek uyarlanabilirlik sunmakta ve farklı madencilik ortamlarında kapsamlı yeniden yapılandırma gereksinimini azaltmaktadır.



Şekil 2. GFDMS'nin Temel Mimarisi [24]

Görsel tabanlı yorgunluk tespit sistemlerinin performansı çalışmalar arasında farklılık göstermekle birlikte, kontrollü deneysel koşullarda çoğu sistem %90'ın üzerinde doğruluk oranları elde ettiği bildirmektedir [21], [25]. Ancak açık ocak madenciliği gibi gerçek saha koşullarında toz birikimi, değişken doğal ve yapay aydınlatma, gece operasyonları ile araç titreşimleri görüntü kalitesini ve algılama doğruluğunu olumsuz yönde etkileyebilmektedir [5], [23], [24]. Bu nedenle güncel sistemlerde görüntü iyileştirme, gürültü giderme, adaptif aydınlatma düzeltmesi ve derin öğrenme tabanlı sağlam görüntü işleme algoritmaları kullanılarak bu olumsuz çevresel etkilerin azaltılması hedeflenmektedir [21].

B. Giyilebilir ve Fizyolojik İzleme

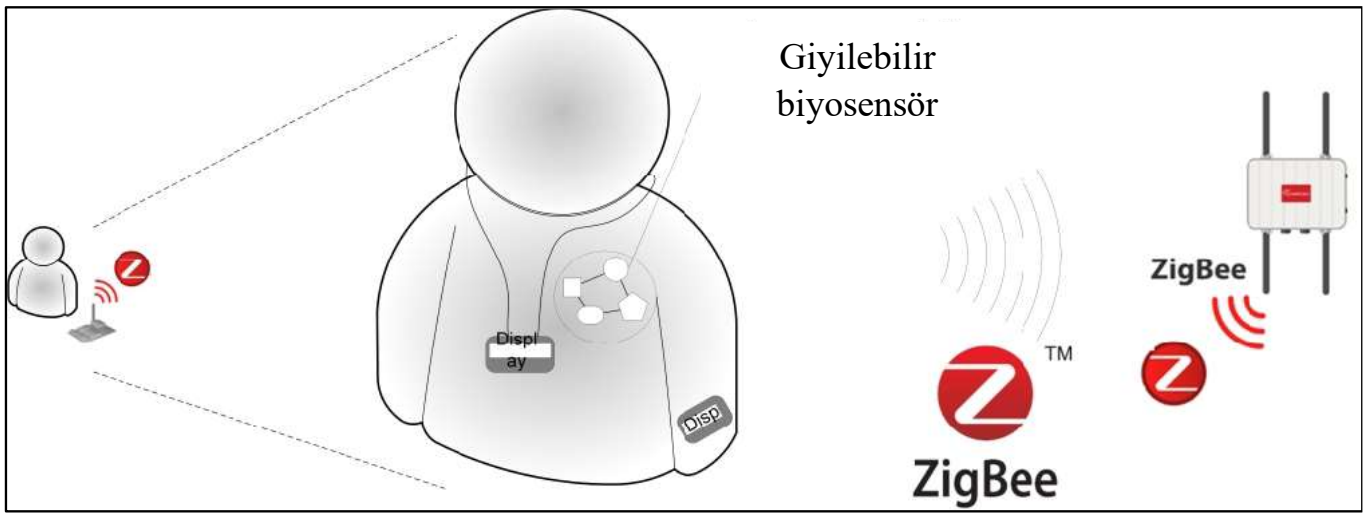
Madencilik sektöründe giyilebilir algılayıcıların özellikle yeraltı ortamlarında gerçek zamanlı yorgunluk izleme ve risk yönetimi açısından önemli bir potansiyele sahip olduğu bildirilmektedir [2]. Giyilebilir algılayıcı teknolojileri, özellikle kamera tabanlı sistemlerin uygulanmasının güç olduğu yeraltı madenciliği

ve yüksek hareketlilik gerektiren çalışma ortamlarında önemli bir alternatif yaklaşım sunmaktadır. Bu sistemler, kalp atış hızı değişkenliği (HRV), elektroensefalografi (EEG), elektrokardiyografi (ECG), cilt sıcaklığı ve hareket örüntüleri gibi fizyolojik göstergeleri sürekli izleyerek operatörün yorgunluk düzeyini değerlendirebilmektedir [3], [19, [20].

Madencilik sektöründe çalışan işçilerin yorgunluk takibini giyilebilir cihazlar ve özel uygulamalar aracılığıyla gerçekleştirmeyi konu alan bir çalışma gerçekleştirilmiştir. Çalışmada, yorgunluğun yalnızca göz kapağı kapanma oranı (*PERCLOS*) gibi tek boyutlu göstergelerle değil, hem öznel ölçümler (*Karolinska Sleepiness Scale* ile uyku hali değerlendirmesi) hem de nesnel ölçümler (reaksiyon süresi testleri) ile birlikte ele alınması gerektiği vurgulanmıştır. Araştırmacılar, yaygın kullanılan akıllı saatler (Fitbit) üzerinde özel bir uygulama geliştirerek operatörlerin vardiya boyunca yorgunluk seviyelerini ölçmüşlerdir. Sistem, operatörün kendi beyan ettiği uyku hali ile reaksiyon süresi verilerini birleştirerek daha kapsamlı bir yorgunluk değerlendirmesi sunmuştur. Sonuçlar, bu tür giyilebilir cihazların pahalı ve müdahaleci yöntemlerin yerine geçebilecek potansiyele sahip olduğunu göstermektedir. Ayrıca, kullanıcıların günlük yaşamda da fayda sağladıkları için bu cihazların kabul edilebilirlik düzeyinin yüksek olduğu belirtilmektedir. Çalışmada ayrıca, gelecekte IoT tabanlı uygulamaların yalnızca teknik değil, aynı zamanda organizasyonel kültür ve kullanıcı desteği açısından da değerlendirilmesi gerektiği de belirtilmiştir [3].

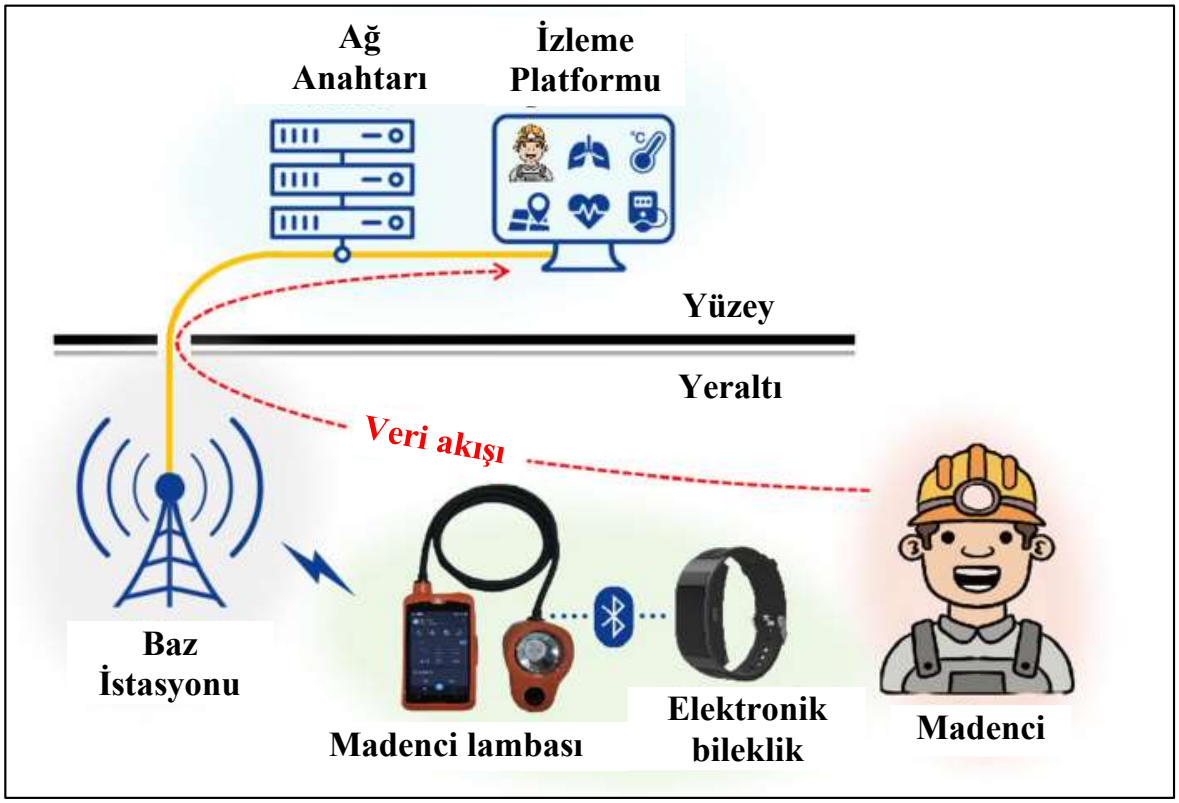
EEG tabanlı teknikler, yorgunluk durumlarıyla ilişkili beyin aktivitesi kalıplarının doğrudan ölçümünü sağlamaktadır. Bongers [26], EEG tabanlı operatör yorgunluğu belirlemeyi inceleyerek uyanıklıkta azalma ile ilişkili karakteristik frekans bandı değişikliklerini tanımlamıştır. EEG, bilişsel yorgunluğa karşı yüksek hassasiyet sunarken pratik uygulama zorlukları arasında uzun vardiyalar sırasında algılayıcı konforu, sinyal kalitesinin korunması ve bireyselleştirilmiş temel ayarlama gereksinimi yer almaktadır.

Yapılan bir çalışmada, işçi yorgunluğunun kazaların başlıca nedenlerinden biri olduğu vurgulanmakta ve bunun **nabız (HR)** ile **hemoglobin oksijen doygunluğu (SpO_2)** ölçümleriyle belirlenebileceği ifade edilmiştir. Araştırmacılar, bu parametreleri sürekli kaydeden küçük, hap boyutunda bir biyoalgılayıcı sensör geliştirmiştir (Şekil 3). Algılayıcı, ZigBee tabanlı kablosuz iletişim modülüyle donatılmış olup, ölçülen verileri yerel WLAN konsantratörlerine iletmektedir. Bu konsantratörler, verileri çoklu atlama yöntemiyle merkezi bir izleme sistemine aktararak işçilerin yorgunluk seviyelerini gerçek zamanlı olarak takip etmeyi mümkün kılmaktadır. Sistem, işçilerin yaş ve SpO_2 değerlerine göre kişiselleştirilmiş eşik değerler belirlemekte ve yorgunluk seviyelerini **yeşil (dinç), sarı (yorgunluk başlangıcı) ve kırmızı (yorgun)** şeklinde sınıflandırmaktadır. Böylece erken uyarı algoritmaları sayesinde işçilerin aşırı yorgunluk durumları önceden tespit edilerek kazaların önlenmesi hedeflenmektedir. Çalışma, ergonomik ve taşınabilir bir pulsoxymeter tasarımıyla yeraltı madenciliği koşullarına uygun bir çözüm sunmaktadır [19].



Şekil 3. Bir biyosensörün, madencinin sağlıkla ilgili parametrelerini kablosuz iletişim sistemi aracılığıyla veri tabanına göndermesine dayanan bir tasarım [19].

Jiang vd. [20], madencilerin durumunu izlemek için **elektronik bileklik** ve **akıllı madenci lambası olmak üzere** iki temel giyilebilir cihaz geliştirilmiştir (Şekil 4). Elektronik bileklik; solunum, vücut sıcaklığı, nabız ve kan basıncı gibi hayati belirtileri toplamakta ve *Bluetooth* üzerinden madenci lambasına aktarmaktadır. Madenci lambası ise tümleşik çoklu gaz algılama (metan, karbon monoksit, oksijen), personel konumlandırma (UWB teknolojisi) ve kablosuz iletişim (5G, Wi-Fi) özelliklerine sahiptir. Böylece hem çevresel riskler hem de işçi sağlığı parametreleri eş zamanlı olarak izlenebilmektedir. Veri iletiminde, **5G iletişim** ve **UWB konumlandırma** birlikte kullanılarak düşük gecikme (≤20 ms) ve yüksek hassasiyet (≤0,3 m hata payı) sağlanmıştır. Yorgunluk tahmini için ise, hayati belirtiler ile **tükürük pH değeri** arasındaki güçlü ilgileşimden yararlanılmış ve geri yayılım algoritması (*back propagation*) ile bir yapay sinir ağı modeli geliştirilmiştir. Geliştirilen model, madencilerin yorgunluk seviyelerini yüksek doğrulukla tahmin edebilmektedir. Sonuç olarak, çalışma yeraltı madenciliğinde **IoT tabanlı bütünleşik bir izleme sistemi** önermektedir. Bu sistem, sağlık ve güvenlik parametrelerini tek bir platformda toplayarak erken uyarı, acil durum müdahalesi ve risk yönetimi için kullanılabilir hale getirmektedir.



Şekil 4. Madencinin durumunu izleme sistemi [20]

C. *Davranışsal ve Performansa Dayalı İzleme*

Davranışsal izleme sistemleri, yorgunluğu doğrudan fizyolojik sinyaller yerine operatör performansı ve araç kullanım davranışlarından elde edilen göstergeler aracılığıyla değerlendirmektedir. Bu yaklaşımlar, direksiyon hareketleri, hız değişkenliği, frenleme örüntüleri ve diğer araç telemetrisi verilerinden yararlanarak ek algılayıcı altyapısına duyulan gereksinimi azaltmakta ve müdahalesiz izleme olanağı sağlamaktadır [2], [23], [27]. Madencilik sektöründe geliştirilen erken uyarı sistemleri ise donanım çalışma örüntülerini analiz ederek operatör yorgunluğunun davranışsal belirtilerini belirlemeye odaklanmış ve bu yaklaşımın proaktif yorgunluk yönetimine katkı sağlayabileceğini göstermiştir [5].

Açık ocak madenlerinde ağır araç operatörlerinin yorgunluk düzeylerini izlemek ve yönetmek için geliştirilen proaktif bir sistemin tanıtımı yapılmıştır. Çalışma, klasik "*dead man's handle*" gibi pasif güvenlik sistemlerinin yorgunluk riskini azaltmada yetersiz kaldığı vurgulanarak, **eğitim, performans yönetimi ve teknolojik izleme cihazlarının birleşimi** önerilmiştir. ARRB'nin geliştirdiği yorgunluk monitörü, sürücünün görsel-işitsel uyaranlara verdiği tepki süresini ölçerek vardiya boyunca gerçek zamanlı yorgunluk takibi yapmaktadır. Tepki süreleri yavaşladığında cihaz uyarı vermekte, üst denetçiye

mesaj göndermekte ve gerektiğinde operatörün görevden alınmasını sağlamaktadır. Sistem yalnızca cihazdan ibaret olmayıp aynı zamanda **eğitim oturumları, kişisel uyku araştırmaları, performans profilleri ve düzenli geri bildirim** ile desteklenmektedir. Böylece hem bireysel hem de kurumsal düzeyde yorgunluk yönetimi sağlanmaktadır. Çalışmada, bu yaklaşımın kazaları ve yorgunluk kaynaklı performans düşüşlerini azaltacağı, işçi sağlığını iyileştireceği ve üretim verimliliğini artıracağı da belirtilmiştir [28].

Wang vd. [29], kömür madenlerindeki pompa istasyonlarında görev yapan operatörlerin güvenli çalışma sürelerinin belirlenmesine yönelik tahmine dayalı bir model geliştirmiştir. Çalışmada, tam mekanize kömür üretim yüzeyinde çalışan pompa istasyonu operatörlerinden saha koşullarında fizyolojik veriler toplanmış; sistolik kan basıncı, diyastolik kan basıncı ve kalp atım hızı gibi yorgunlukla ilişkili biyolojik belirteçler analiz edilmiştir. Araştırmacılar, bu verileri GM (1,1) Gri Tahmin Modeli (*Grey Prediction Model*) kullanarak operatörlerin çalışma süresi boyunca fizyolojik değişimlerini modellemiş ve yorgunluğun güvenli çalışma kapasitesi üzerindeki etkisini nicel olarak değerlendirmiştir. Model sonuçları, fizyolojik göstergelerde çalışma süresi uzadıkça belirgin değişimler meydana geldiğini ve bu değişimlerin operatör performansındaki olası düşüşün erken göstergeleri olarak kullanılabileceğini ortaya koymuştur. Elde edilen tahminler doğrultusunda, pompa istasyonu operatörleri için yaklaşık 5,7 saatlik güvenli çalışma süresi önerilmiş ve bu sürenin aşılması durumunda yorgunluğa bağlı hata olasılığının artabileceği belirtilmiştir. Çalışma, gerçek zamanlı fizyolojik izleme ile tahmine dayalı modelleme tekniklerinin birleştirilmesinin, yalnızca mevcut yorgunluk durumunu belirlemekle kalmayıp yorgunluk oluşmadan önce çalışma sürelerinin planlanmasına da katkı sağlayabileceğini göstermektedir. Bu yönüyle araştırma, madencilik sektöründe proaktif yorgunluk yönetimi ve çalışma–dinlenme düzenlerinin bilimsel verilere dayalı olarak optimize edilmesine yönelik önemli bir örnek sunmaktadır.

Ugarte ve Muñoz Sánchez [30], tarafından tanıtılan **WOMBAT** sistemi, açık ocak madenciliğinde operatör yorgunluğu tahmini ile filo konumlandırmasını aynı platformda birleştiren bütünleşik bir yaklaşım sunmuştur. Sistem, çalışma süresi, önceki dinlenme süresi ve sirkadiyen ritim gibi parametrelerden yararlanarak operatörün yorgunluk düzeyini tahmin etmekte; bunun yanında her taşıma kamyonunun konumu, hızı, operatör bilgisi ve çalışma durumunu merkezi bir harita üzerinde izleyebilmektedir. Bu bütünleşik yapı, filo yöneticilerinin hem operasyonel planlamayı hem de yorgunluk riskini eş zamanlı değerlendirmesine olanak sağlayarak güvenli ve verimli karar verme süreçlerini desteklemektedir.

Paulo vd. [31], Brezilya'da faaliyet gösteren bir madencilik işletmesinde çalışanların göreve hazır olma düzeylerini değerlendirmek amacıyla geliştirilen **FOCOS/Prontos** bilgisayarlı değerlendirme sistemini incelemiştir. Araştırmada sistemin performansı, Chalder Yorgunluk Ölçeği ile karşılaştırılmış ve yorgunluk belirtileri gösteren çalışanlar ile kontrol grubunun hazır olma profilleri analiz edilmiştir. FOCOS/Prontos sistemi yalnızca yorgunluk belirtilerini değil; dikkat, bilişsel performans, reaksiyon süresi ve riskli davranış eğilimleri gibi göreve uygunluğu etkileyen çok sayıda faktörü birlikte değerlendiren çok boyutlu bir yaklaşım sunmaktadır. Bulgular, yorgunluk belirtileri bulunan çalışanların hazır olma değerlendirme sonuçlarının kontrol grubundan anlamlı biçimde farklılaştığını ve sistemin yorgunlukla ilişkili risklerin erken belirlenmesinde kullanılabilecek nesnel bilgiler sağladığını göstermektedir. Araştırmacılar, yorgunluğun çok faktörlü yapısı nedeniyle günlük hazır olma durumunun sürekli ve bütüncül olarak değerlendirilmesinin, iş sağlığı ve güvenliği uygulamalarında proaktif müdahaleleri destekleyebileceğini vurgulamaktadır.

IV. Temel Bulgular ve Karşılaştırmalı Analiz

A. Algılama Doğruluğu ve Performansı

Algılama teknolojilerinin karşılaştırmalı analizi, farklı yaklaşımlar ve operasyonel bağlamlar arasında performans ölçütlerinde anlamlı farklılıklar ortaya koymaktadır. Görsel tabanlı sistemler, kontrollü veya yarı kontrollü ortamlarda %85 ile %95 arasında algılama doğruluğu bildirmektedir [18], [24]. Sun [18], açık ocak madenciliğinde çalışan nakliye kamyonu operatörlerinin yorgunluğunu belirlemek amacıyla bulanık sinir ağı (*Fuzzy Neural Network, FNN*) tabanlı bir tespit yöntemi geliştirmiştir. Sistem, göz kapama oranı (*PERCLOS*), göz kırpma sıklığı, esneme davranışı ve baş duruşu gibi görsel yorgunluk göstergelerini birlikte değerlendirerek operatörün yorgunluk düzeyini sınıflandırmaktadır. Araştırmacılar, bulanık mantık ile yapay sinir ağlarını birleştiren bu yaklaşımın, yorgunluğun keskin eşiklerle tanımlanabilen bir durumdan ziyade kademeli olarak gelişen bir süreç olduğunu dikkate aldığını belirtmiştir. Bu sayede, yorgunluk

seviyeleri arasındaki belirsizlik daha başarılı şekilde modellenmiş ve geleneksel eşik tabanlı yöntemlere kıyasla daha yüksek sınıflandırma doğruluğu ile daha kararlı bir performans elde edilmiştir. Çalışma, özellikle açık ocak madenciliği gibi değişken çevresel koşullara sahip operasyonlarda bulanık sinir ağı tabanlı yaklaşımların güvenilir yorgunluk tespit sistemleri geliştirilmesi açısından önemli bir potansiyele sahip olduğunu ortaya koymaktadır. Kamera tabanlı yorgunluk tespit sistemleri kontrollü laboratuvar koşullarında yüksek doğruluk oranları bildirse de saha uygulamalarında çevresel koşullar sistem performansını önemli ölçüde etkileyebilmektedir. Özellikle açık ocak madenciliğinde toz oluşumu, değişken doğal ve yapay aydınlatma, gece operasyonları, araç titreşimleri ve kamera konumlandırma kısıtları görüntü kalitesini düşürerek algılama doğruluğunu azaltabilmektedir [5], [23]. Gerçek saha uygulamalarını inceleyen çalışmalar, bu nedenle görüntü önişleme, adaptif aydınlatma düzeltmesi ve daha sağlam makine öğrenmesi algoritmalarının kullanılmasının gerekli olduğunu vurgulamaktadır [18], [24].

Giyilebilir fizyolojik izleme sistemleri, kalp atış hızı değişkenliği ve EEG modellerindeki yorgunlukla ilişkili değişikliklere karşı yüksek duyarlılık göstermektedir [19], [26]. Ancak bu sistemler, yanlış pozitif oranlarıyla ilgili zorluklarla karşı karşıyadır; zira fizyolojik değişiklikler, fiziksel çaba, termal stres veya duygusal tepkiler dâhil yorgunluk dışındaki faktörlerden de kaynaklanabilir [20]. Birden fazla fizyolojik göstergeyi birleştiren çok parametreli yaklaşımlar, yorgunluğu diğer fizyolojik durumlardan ayırt ederek özgüllüğü artırma konusunda umut vadetmektedir.

Davranışsal izleme sistemleri, müdahale gerektirmemeleri ve mevcut operasyonel veri akışlarıyla bütünleşmeleri açısından önemli avantajlar sunmaktadır [28]. ARRB sisteminin sürüş davranışı modellerinin analizi, ek operatör etkileşimi veya ek donanım gerektirmeden sürekli değerlendirme olanağı sağlamaktadır [28]. Ancak davranışsal göstergeler, fizyolojik yorgunluğun başlangıcının gerisinde kalabilir ve önleyici müdahale penceresini daraltabilir. Wang vd.'nin [29] tahmine dayalı modelleme yaklaşımı, vardiya modelleri ve geçmiş verilere dayanarak yorgunluk riskini tahmin ederek bu sınırlamayı ele almakta ve reaktif yönetim yerine proaktif yönetimi mümkün kılmaktadır.

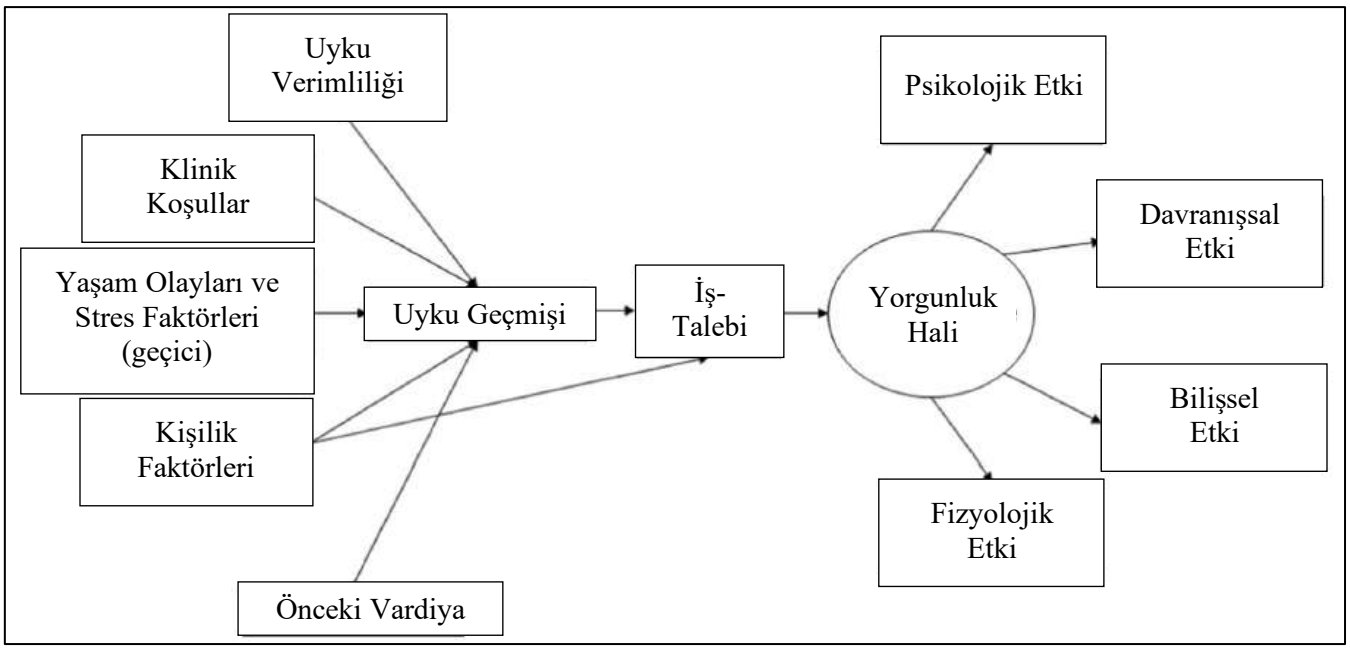
Birden fazla çalışmada ortaya çıkan kritik bir bulgu, bireyselleştirilmiş temel değerlerin ve uyarlanabilir algoritmaların önemidir [15], [26], [29]. Yorgunluk belirtileri bireyler arasında belirgin ölçüde farklılık göstermektedir. Bazı operatörler ciddi davranışsal değişiklikler sergilerken diğerleri daha ince göstergeler ortaya koymaktadır. Bireysel temel profiller oluşturan ve algılama eşiklerini buna göre uyarlayan sistemler, topluluk düzeyinde sabit eşiklere kıyasla daha üstün performans göstermektedir.

B. Operasyonel Bağlam ve Uygulama

Operasyonel bağlam, yorgunluk risk profilleri ile farklı tespit yaklaşımlarının uygulanabilirliğini önemli ölçüde etkiler. Açık ocaklardaki taşıma operasyonları, bu amaç için özel olarak tasarlanmış çok sayıda sistemle en kapsamlı biçimde incelenmiş bağlamı temsil etmektedir [17], [18], [24], [30], [32]. Sabit oturma pozisyonları ve öngörülebilir aydınlatma koşullarıyla kamyon kabinlerinin nispeten kontrollü ortamı, görüş tabanlı izlemeyi kolaylaştırır. Ancak tekrarlayan taşıma rotalarının monotonluğu, fiziksel talepler orta düzeyde kalsa bile pasif yorgunluğa yol açan koşullar yaratır [17], [18].

Yeraltı madenciliği, yorgunluk izlemesi için farklı zorluklar ve fırsatlar sunar. Kowal vd. [19] ile Jiang vd. [20], sınırlı kablosuz bağlantı, kendinden emniyetli donanım gereksinimleri ve zorlu koşullara dayanabilen sağlam algılayıcılara duyulan ihtiyaç dâhil olmak üzere yeraltı kısıtlarını özel olarak ele alan sistemler geliştirmiştir. Yeraltı operasyonları, yüzeydeki taşıma operasyonlarına kıyasla genellikle daha çeşitli görevler ve daha yüksek fiziksel talepler içermekte; bu da hem bilişsel hem de fiziksel yorgunluk bileşenlerine duyarlı tespit yaklaşımlarını gerektirmektedir.

Uygulama deneyimleri, teknolojik yetkinliğin tek başına başarılı bir uygulamayı garanti etmediğini göstermektedir. Drews vd. [15], iki madencilik şirketinde 44 kamyon operatörüyle odak grup görüşmeleri yürüterek operatörlerin yorgunluk izleme sistemlerine ilişkin algı ve deneyimlerine dair kritik bilgiler ortaya koymuştur (Şekil 5). Operatörler, yorgunluk belirtilerinin ve performans üzerindeki etkilerinin farkında olduklarını ve yorgunluğu yönetmek için çeşitli kişisel stratejiler kullandıklarını bildirmiştir. Ancak mevcut izleme sistemlerinin geçerliliği ve güvenilirliği konusunda endişelerini dile getirmiş ve bu sistemleri etkisiz olarak değerlendirmiştir. Bu bulgu, operatörlerin izleme teknolojisini kabullenmesi ile güveninin önemini vurgulamaktadır; bu faktörler tamamen teknik odaklı sistem geliştirmede genellikle göz ardı edilmektedir.



Şekil 5. Yorgunluk kavramsal modeli [15]

Drews vd. [15] ile Lee'nin [33] vurguladığı sosyo-teknik bakış açısı, etkili yorgunluk yönetiminin, teknolojik izlemenin kurumsal politikalar, işyeri kültürü ve operatör katılımıyla bütünleşmesini gerektirdiğini ortaya koymaktadır. Cezalandırıcı gözetim araçları olarak algılanan sistemler dirençle karşılaşırken, güvenlik destek araçları olarak çerçevelenenler daha fazla kabul görmektedir. Başarılı uygulamalar, operatörleri sistem tasarımına çekmekte, yorgunluk verilerine ilişkin şeffaf geri bildirim sağlamakta ve izleme sonuçlarını disiplin cezaları yerine destekleyici müdahaleleri şekillendirmek için kullanmaktadır.

Belirli madencilik operasyonlarında gerçekleştirilen saha uygulamaları, yorgunluk tespit sistemlerinin başarısının yalnızca kullanılan algoritmalara değil, aynı zamanda uygulandıkları operasyonel bağlama da bağlı olduğunu göstermektedir. Li vd. [24], Myanmar'daki bir açık bakır madeninde gerçekleştirdikleri uygulamada, sistem performansının yerel çalışma koşulları, araç özellikleri ve çevresel faktörlere göre uyarlanmasının kritik önem taşıdığını ortaya koymuştur. Benzer şekilde, madencilikte yorgunluk risk yönetimi üzerine yapılan çalışmalar; iklim koşulları, vardiya düzenleri, iş gücü demografisi, kültürel özellikler ve ulusal düzenlemeler gibi bağlamsal faktörlerin hem yorgunluk risk profilini hem de uygulanacak izleme teknolojilerinin etkinliğini ve kabul edilebilirliğini önemli ölçüde etkilediğini vurgulamaktadır [2], [5]. Bu bağlamsal faktörler hem yorgunluk risk profillerini hem de farklı izleme yaklaşımlarının kabul edilebilirliğini etkilemektedir.

C. Vardiya Kalıpları ve Sirkadiyen Etkiler

Madencilikte yorgunluk literatürünün en önemli bulgularından biri, sirkadiyen ritimlerin operatör performansı üzerindeki belirleyici etkisidir. Mabbott ve Lloyd [34] tarafından Avustralya'daki bir açık altın madeninde yürütülen kapsamlı saha çalışmasında, 14 ardışık gece vardiyası, bir gün izin ve ardından 13 ardışık gündüz vardiyasından oluşan vardiya düzeninde çalışan 24 kamyon operatöründen 3500 saatten fazla gerçek zamanlı objektif performans verisi toplanmıştır. Araştırma, Avustralya madencilik sektöründe sürücü performansının gerçek çalışma koşullarında gerçekleştirilen ilk büyük ölçekli objektif değerlendirmelerden biri olarak kabul edilmektedir. Bulgular, tek bir 12 saatlik gece vardiyası boyunca görev süresinin performans üzerinde belirgin bir olumsuz etki oluşturmadığını, buna karşın sirkadiyen ritimlerin özellikle sabahın erken saatlerinde (02.00–06.00) operatör performansını anlamlı ölçüde etkilediğini göstermiştir. Ayrıca, 14 ardışık gece vardiyası içerisinde 13., 14., 7. ve 2. gecelerde en düşük performans düzeyleri gözlenmiş; bu sonuçlar, akut sirkadiyen etkiler ile ardışık vardiyalar boyunca biriken yorgunluğun birlikte operatör performansını şekillendirdiğini ortaya koymuştur. Bu bulgular, yorgunluk yönetimi politikaları açısından önemli sonuçlar doğurmaktadır. Vardiya süresi veya toplam çalışma saatlerine getirilen basit kısıtlamalar, sirkadiyen zamanlamayı ve vardiya döngüleri boyunca birikimli

etkileri hesaba katmadıkları sürece yorgunluk riskini etkili biçimde azaltamayabilir [34]. Veriler, vardiya tasarımının (ardışık vardiya sayısı, vardiyaların sirkadiyen ritimlere göre zamanlaması ve vardiya döngüleri arasındaki iyileşme süresi dâhil) tek başına vardiya süresinden daha kritik olabileceğini göstermektedir.

Wang vd. [29], kömür madenlerinde pompa istasyonu operatörlerinin güvenli çalışma sürelerini belirlemek amacıyla fizyolojik göstergelere dayalı bir tahmin modeli geliştirmiştir. Çalışmada kalp atım hızı ve kan basıncı gibi biyolojik göstergeler kullanılarak çalışma süresine bağlı yorgunluk değişimleri modellenmiş ve güvenli çalışma süresi tahmin edilmiştir. Bununla birlikte, madencilikte yorgunluk riskinin yalnızca vardiya süresiyle açıklanamayacağı; vardiya zamanlaması, ardışık gece vardiyaları, uyku kaybı ve toparlanma süresi gibi zamansal faktörlerin de dikkate alınması gerektiği önceki vardiya çalışmaları tarafından ortaya konmuştur [35]. Bu nedenle modern yorgunluk yönetimi yaklaşımları, sabit çalışma süresi sınırları yerine sirkadiyen ritim ve birikimli yorgunluğu dikkate alan dinamik risk değerlendirme modellerine yönelmektedir.

Mabbott ve Lloyd [34], ardışık gece vardiyalarında çalışan açık ocak kamyon operatörlerinin performanslarında bireysel farklılıklar bulunduğunu göstermiştir. Aynı vardiya düzenine maruz kalan operatörler arasında yorgunluğa bağlı performans düşüşünün büyüklüğü farklılık göstermiş; bazı operatörler gece çalışma koşullarına daha iyi uyum sağlarken bazıları daha belirgin bozulmalar sergilemiştir. Bu bulgular, yorgunluk riskinin yalnızca vardiya süresi veya çalışma düzeni gibi operasyonel faktörlerle açıklanamayacağını, bireysel biyolojik farklılıkların da dikkate alınması gerektiğini göstermektedir. Bu nedenle modern yorgunluk risk yönetimi yaklaşımları, çalışanların bireysel özelliklerini ve sirkadiyen uyum farklılıklarını değerlendiren daha kişiselleştirilmiş modeller geliştirmeye yönelmektedir [34], [35].

D. Yönetim Çerçeveleri ve Müdahaleleri

Madencilikte etkili yorgunluk yönetimi, tespit teknolojisinin ötesine geçerek izleme, müdahale, politika ve örgüt kültürünü birleştirmelidir. Çok sayıda çalışmada, birden fazla bileşeni bütünleştiren tümleşik yönetim sistemleri tanımlanmaktadır [3], [28], [30], [31], [36].

ARRB Proaktif Yorgunluk Yönetim Sistemi, gerçek zamanlı izlemeyi filo yönetimi ve planlama sistemleriyle bütünleştiren kapsamlı bir yaklaşıma örnek teşkil etmektedir [28]. Sistem, yorgun operatörleri yalnızca uyarmak yerine yeniden görev atama, planlı molalar ve vardiya düzenlemeleri de dâhil olmak üzere proaktif müdahaleleri mümkün kılar. Bu yaklaşım, yorgunluk yönetimini reaktif olay müdahalesinden önleyici risk azaltımına kaydırır.

Rogers vd. [3], madencilik operatörlerinde yorgunluğun sürekli ve gerçek zamanlı izlenebilmesi amacıyla IoT tabanlı bir giyilebilir yorgunluk takip sistemi geliştirmiştir. Geliştirilen sistem; giyilebilir algılayıcılar, mobil uygulamalar ve veri analiz altyapısı aracılığıyla operatörlerin uyku hali, algılanan yorgunluk düzeyi ve bilişsel performans göstergeleri hakkında sürekli veri toplamaktadır. Çalışmada özellikle Karolinska Uyku Ölçeği (KSS) ve reaksiyon zamanı ölçümleri kullanılarak öznel ve nesnel yorgunluk göstergelerinin birlikte değerlendirilmesi amaçlanmıştır. Bu yaklaşım, madencilik gibi yüksek riskli çalışma ortamlarında yorgunluğun yalnızca vardiya sonrasında değerlendirilen bir durum olmaktan çıkarılarak, operasyon sırasında izlenebilen dinamik bir risk faktörü olarak ele alınmasına olanak sağlamaktadır.

IoT tabanlı yorgunluk izleme sistemleri, *Fatigue Risk Management System (FRMS)* yaklaşımlarıyla bütünleştirildiğinde, yalnızca yorgunluk seviyesinin belirlenmesi değil, aynı zamanda uygun yönetimsel müdahalelerin planlanması için de kullanılabilmektedir. FRMS yaklaşımı; çalışma saatleri, vardiya düzenleri, uyku fırsatları, bireysel farklılıklar ve gerçek zamanlı performans göstergelerini birlikte değerlendirerek yorgunluk riskinin proaktif biçimde yönetilmesini amaçlamaktadır. Bu kapsamda, gerçek zamanlı yorgunluk verileri; dinlenme molalarının planlanması, görev değişimi, vardiya düzenlemesi veya operasyonel risk seviyesine bağlı diğer kontrol önlemlerinin uygulanması için karar destek sağlayabilir. Rogers vd. [3] tarafından önerilen IoT tabanlı altyapı, bu tür kapalı döngülü yorgunluk yönetim sistemlerinin oluşturulması için gerekli veri toplama ve izleme bileşenlerini destekleyen önemli bir teknolojik yaklaşım olarak değerlendirilmektedir.

Yorgunluk izleme sistemleri, yorgunluk riskinin azaltılmasında tek başına bağımsız bir çözüm olarak değil, daha geniş bir güvenlik yönetimi yaklaşımının tamamlayıcı bir bileşeni olarak değerlendirilmektedir.

FRMS yaklaşımı, yorgunluk kaynaklı risklerin kontrol edilmesinde tek bir müdahalenin yeterli olmayacağını; vardiya planlaması, uygun çalışma-dinlenme düzenleri, çalışan eğitimi, performans izleme teknolojileri ve kurumsal politikaların birlikte uygulanması gerektiğini vurgulamaktadır. Bu çok katmanlı yaklaşım, güvenlik mühendisliğindeki derinlemesine savunma (*defense-in-depth*) prensibiyle uyumlu olup herhangi bir kontrol önleminin yetersiz olması durumunda diğer katmanların riski azaltmaya devam etmesini amaçlamaktadır. Aktif yorgunluk izleme teknolojileri, bu güvenlik katmanları içerisinde operatörün mevcut durumunun değerlendirilmesini sağlayan önemli bir araç olarak görülmektedir. Gerçek zamanlı sensörler, performans testleri ve davranışsal izleme sistemleri sayesinde çalışanların yorgunluk düzeyleri belirlenebilir ve gerekli durumlarda dinlenme planlaması, görev düzenlemesi veya diğer risk azaltıcı önlemler için karar desteği sağlanabilir. Bununla birlikte, bu teknolojilerin etkinliği; uygun vardiya tasarımı, yönetim desteği ve çalışanların yorgunluk farkındalığı gibi organizasyonel faktörlerle birlikte ele alındığında artmaktadır [6], [10], [35].

Paulo vd. [31] tarafından geliştirilen FOCOS/Prontos sistemi, çalışanların göreve hazır olma durumunu yalnızca yorgunluk göstergeleri üzerinden değil, bilişsel performans, dikkat ve diğer işe uygunluk faktörleriyle birlikte değerlendiren çok faktörlü bir tarama yaklaşımını temsil etmektedir. Bu yaklaşım, çalışan güvenliğinin tek bir risk faktörüne bağlı olmadığını; yorgunluk, sağlık durumu, bilişsel kapasite ve bireysel farklılıklar gibi birbiriyle etkileşimli birçok unsurun birlikte değerlendirilmesi gerektiğini kabul etmektedir. Araştırmada, yorgunluk belirtileri bulunan çalışanların FOCOS/Prontos değerlendirme sonuçlarının kontrol grubundan anlamlı farklılıklar gösterdiği ve sistemin çalışanların hazır olma durumunu değerlendirmede kullanılabilecek bütüncül bir araç sunduğu ortaya konmuştur. Bu tür çok boyutlu değerlendirme yaklaşımları, yorgunluk yönetiminin daha geniş iş sağlığı ve güvenliği programlarıyla bütünleştirilmesine katkı sağlamaktadır.

Horberry vd. [36], madencilik sektöründe yorgunluk yönetimine ilişkin mevcut uygulamaları ve gelecekteki gelişim alanlarını inceleyerek, etkili yorgunluk kontrolünün yalnızca bireysel çalışan davranışlarına değil, aynı zamanda kurumsal ve operasyonel yönetim yaklaşımlarına dayanması gerektiğini vurgulamıştır. Araştırmacılar, madencilikte yorgunluk riskinin kontrol edilmesi için vardiya tasarımı, çalışma saatlerinin yönetimi, yorgunluk izleme teknolojilerinin kullanımı, çalışanların sağlık ve iyi olma durumlarının izlenmesi ve güçlü bir güvenlik kültürünün oluşturulmasını içeren bütüncül yönetim sistemlerini önermektedir. Bu yaklaşım, yorgunluğun karmaşık ve çok faktörlü doğasını dikkate alarak, bireysel önlemler ile örgütsel kontrol mekanizmalarının birlikte uygulanmasının gerekliliğini ortaya koymaktadır. Uyku yönetimi eğitimi, uygun aydınlatma uygulamaları ve diğer çevresel düzenlemeler gibi destekleyici müdahaleler de yorgunluk riskinin azaltılmasına katkı sağlayabilecek tamamlayıcı stratejiler olarak değerlendirilmektedir [35], [36], [37].

Literatürde tanımlanan yorgunluk müdahale stratejileri, kısa vadeli operasyonel kontrollerden uzun vadeli risk yönetimi yaklaşımlarına kadar geniş bir kapsamda ele alınmaktadır. Tespit edilen yorgunluk durumlarında uygulanabilecek acil müdahaleler arasında planlı veya ilave dinlenme molaları, geçici görev değişiklikleri, vardiya düzenlemeleri ve operatörün uyanıklığını artırmaya yönelik kısa süreli uygulamalar bulunmaktadır. Kafein kullanımı gibi geçici uyanıklık artırıcı müdahaleler, özellikle vardiyalı çalışanlarda kısa süreli performans desteği sağlamak amacıyla değerlendirilmektedir [38]. Bunun yanında, gelişmiş yorgunluk tespit sistemleriyle entegre edilen gerçek zamanlı uyarı mekanizmaları ve araç kontrol destek sistemleri, özellikle yüksek riskli madencilik operasyonlarında operatör kaynaklı hataları azaltmaya yönelik potansiyel bir ek güvenlik katmanı olarak değerlendirilmektedir [5].

Uzun vadeli müdahaleler ise yorgunluk riskinin temel nedenlerini azaltmaya yönelik sistematik ve sürdürülebilir yaklaşımlara odaklanmaktadır. Bu kapsamda vardiya planlarının biyolojik ritimler ve çalışanların toparlanma gereksinimleri dikkate alınarak optimize edilmesi, yeterli dinlenme sürelerinin sağlanması, çalışanlara yönelik uyku ve yorgunluk farkındalığı eğitimlerinin uygulanması ve çalışma ortamının iyileştirilmesi temel stratejiler arasında yer almaktadır. Etkili yorgunluk yönetimi, yalnızca bireysel davranış değişikliklerine dayanan bir yaklaşım olmayıp; vardiya tasarımı, çalışma organizasyonu, teknoloji destekli izleme sistemleri ve kurumsal güvenlik kültürünü içeren bütüncül bir risk yönetimi sürecidir [5], [6], [35], [37].

Maisey vd. [39], uzak bölgelerdeki madencilik operasyonlarında çalışan vardiyalı işçilerde yorgunluk riskini azaltmaya yönelik kanıta dayalı müdahalelerin değerlendirilmesi amacıyla randomize kontrollü bir

çalışma metodolojisi geliştirmiştir. Çalışma, uyku eğitimi ve dijital biyolojik geri bildirim uygulamalarının çalışanların uyku kalitesi, uyku süresi ve uyanıklık düzeyi üzerindeki etkisini değerlendirmeyi amaçlamaktadır. Bu yaklaşım, yorgunluk yönetiminde uygulanan müdahalelerin etkinliğinin sistematik ve bilimsel yöntemlerle değerlendirilmesinin önemini ortaya koymaktadır.

V. TARTIŞMA

Önemli teknolojik ilerlemelere rağmen madencilikte yorgunluk tespiti ve yönetim sistemlerinin etkinliğini ve benimsenmesini sınırlayan çok sayıda zorluk bulunmaktadır. Bu zorluklar teknik, operasyonel, örgütsel ve insan faktörlerine ilişkin alanları kapsamaktadır.

a) *Teknik zorluklar*: Madencilik operasyonlarındaki çevresel koşullar, yorgunluk tespit teknolojileri için önemli teknik engeller oluşturmaktadır. Görüntü tabanlı sistemler; kamera yüzeylerinde toz birikmesi, değişken aydınlatma koşulları ve titreşim kaynaklı görüntü bozulmaları nedeniyle gerçek saha uygulamalarında performans kaybı yaşayabilmektedir [5]. Giyilebilir algılayıcı sistemler ise sıcaklık, nem, mekanik darbe ve yeraltı madenciliğinde patlayıcı atmosfer gereksinimleri gibi zorlu koşullara dayanıklı olacak şekilde tasarlanmalıdır. Ayrıca yeraltı haberleşme sistemleri, kaya yapısı ve kapalı ortam geometrisi nedeniyle sinyal zayıflaması, çoklu yol yayılımı ve bağlantı kalitesi problemleriyle karşılaşmaktadır [40], [41].

Sistem güvenilirliği ve yanlış alarm oranları, yorgunluk izleme teknolojilerinin uygulanmasında kritik teknik sorunlar arasında yer almaktadır. Yüksek yanlış pozitif oranları, yani operatör uyanık durumdayken gereksiz yorgunluk uyarılarının oluşturulması, kullanıcıların sisteme olan güvenini azaltabilir ve zamanla alarm yorgunluğuna veya sistemin göz ardı edilmesine neden olabilir. Buna karşılık, yanlış negatif sonuçlar gerçek yorgunluk durumlarının tespit edilememesine yol açarak sistemin güvenlik katkısını sınırlar. Bu nedenle, yorgunluk tespit sistemlerinde duyarlılık ve özgüllük arasında uygun bir denge kurulması gerekmektedir. Ancak yüksek doğruluk sağlayan algoritmaların geliştirilmesi; bireysel farklılıkların, çevresel değişkenlerin ve gerçek zamanlı işlem gereksinimlerinin dikkate alınmasını gerektirdiğinden sistem karmaşıklığını ve uygulama maliyetlerini artırabilmektedir [5], [42].

b) *İşletmeye dönük zorluklar*: Yorgunluk izleme sistemlerinin mevcut operasyonel altyapılar ve iş akışlarıyla bütünleştirilmesi, madencilik sektöründe önemli uygulama zorluklarından biri olarak değerlendirilmektedir. Madencilik operasyonları; kullanılan ekipman türleri, araç filolarının özellikleri, çalışma ortamları (açık ocak veya yeraltı) ve üretim süreçleri bakımından büyük çeşitlilik göstermektedir. Bu operasyonel farklılıklar, yorgunluk izleme teknolojilerinin her işletmenin teknik altyapısına ve çalışma koşullarına uygun biçimde uyarlanmasını gerekli kılmaktadır. Teknolojik çözümlerin yalnızca algılama doğruluğu açısından değil, mevcut iş süreçleriyle uyumu, kullanıcı kabulü ve sürdürülebilir işletilebilirliği açısından da değerlendirilmesi gerekmektedir [5], [43].

Açık ocak madenciliğinde gerçekleştirilen uygulama çalışmaları, farklı araç tipleri ve operasyonel koşullar altında çalışabilecek esnek sistem mimarilerine duyulan ihtiyacı ortaya koymaktadır. Li vd. [24], Myanmar'daki bir bakır madeninde geliştirdikleri genelleştirilmiş yorgunluk izleme sistemi ile farklı madencilik araçlarına ve saha koşullarına uyarlanabilen çok bileşenli bir izleme yaklaşımının uygulanabilirliğini göstermiştir. Bununla birlikte, mevcut maden filolarında kullanılan eski nesil ekipmanlar; gelişmiş görüntü işleme, veri iletişimi ve gerçek zamanlı analiz için gerekli hesaplama kapasitesi veya haberleşme arayüzlerine sahip olmayabilir. Bu tür sistemlerin mevcut araçlara sonradan entegre edilmesi (*retrofit*), donanım değişiklikleri, bakım gereksinimleri ve üretim süreçlerinde olası kesintiler nedeniyle önemli ekonomik ve operasyonel zorluklar oluşturabilmektedir [5], [43].

Filo genelinde yorgunluk izleme sistemleri tarafından üretilen büyük hacimli sürekli veri akışı, veri yönetimi ve yorumlanmasına ilişkin önemli güçlükleri de beraberinde getirmektedir. Ham algılayıcı verilerinin güvenilir ve eyleme dönüştürülebilir bilgilere dönüştürülebilmesi; gerçek zamanlı veri analitiği, uygun filtreleme yöntemleri ve karar destek algoritmalarının kullanılmasını gerektirmektedir [3], [20]. Bununla birlikte, yorgunluk riskinin hangi düzeyde müdahale gerektirdiğinin belirlenmesi ve uygulanacak müdahale türünün seçilmesi, yalnızca teknik doğruluğa değil, aynı zamanda operasyonel gereksinimler, güvenlik hedefleri ve çalışan kabulü gibi faktörlerin birlikte değerlendirilmesini gerektirmektedir. Bu nedenle yorgunluk yönetim sistemlerinde müdahale eşiklerinin belirlenmesi, *Fatigue Risk Management*

System (FRMS) yaklaşımının temel bileşenlerinden biri olarak kabul edilmekte ve güvenlik ile operasyonel verimlilik arasında dengeli karar mekanizmalarının oluşturulması önerilmektedir [35], [36].

c) *Örgütsel ve insan faktörlerine ilişkin zorluklar*: Operatörlerin yorgunluk izleme sistemlerini benimsemesi, bu teknolojilerin başarılı biçimde uygulanmasındaki en kritik faktörlerden biridir. Araştırmalar, çalışanların sistemlerin doğruluğu, güvenilirliği ve kullanım amacına ilişkin algılarının teknolojiye duyulan güveni doğrudan etkilediğini göstermektedir. Özellikle izleme verilerinin cezalandırıcı amaçlarla kullanılabileceği yönündeki endişeler, mahremiyet kaygıları ve sürekli gözetim algısı, çalışanların sistemi benimsemesini olumsuz etkileyebilmekte ve bazı durumlarda sistemlerin devre dışı bırakılmasına veya kullanımının engellenmesine yol açabilmektedir. Bu nedenle yorgunluk izleme sistemlerinin başarılı olabilmesi için teknik doğruluğun yanı sıra şeffaf veri yönetimi politikaları, çalışan katılımı ve güvene dayalı bir örgütsel kültürün oluşturulması büyük önem taşımaktadır [5], [35], [36].

Sosyo-teknik sistemler perspektifi, yorgunluk yönetiminde teknolojinin tek başına yeterli olmadığını, etkin sonuçların ancak teknik, insan ve örgütsel bileşenlerin birlikte ele alınmasıyla elde edilebileceğini vurgulamaktadır [5], [15]. Bu yaklaşıma göre örgüt kültürü, yönetimin yorgunluk riskine yönelik taahhüdü, çalışan eğitimi, vardiya planlaması ve kurumsal politika çerçeveleri, kullanılan izleme teknolojileriyle uyum içinde olmalıdır. Aşırı vardiya süreleri, yetersiz toparlanma fırsatları, yüksek iş yükü ve üretim baskısı gibi yorgunluğun temel nedenleri giderilmeden yalnızca izleme teknolojilerinin uygulanması, güvenlik performansında beklenen iyileşmeyi sağlamayabilir. Bu nedenle güncel FRMS yaklaşımı, teknolojik izleme çözümlerinin örgütsel süreçler ve insan faktörleriyle bütünleşik biçimde uygulanmasını önermektedir [6], [35].

d) *Mevzuat ve standardizasyon zorlukları*: Madencilikte yorgunluk yönetimi alanında kullanılan izleme sistemlerine ilişkin standartlaştırılmış performans ölçütleri, doğrulama protokolleri ve ortak değerlendirme kriterleri henüz tam olarak oluşturulmamıştır. Bu durum, farklı teknolojilerin doğruluk ve güvenilirlik açısından karşılaştırılmasını, uygun sistemlerin seçimini ve uygulamaların nesnel olarak değerlendirilmesini güçleştirmektedir [5], [36]. Ayrıca, yorgunluk yönetimine yönelik uygulamalar ve düzenleyici yaklaşımlar ülkeler ve işletmeler arasında farklılık göstermektedir. Bu durum, uluslararası ölçekte faaliyet gösteren madencilik şirketlerinin farklı operasyonlarda tutarlı yorgunluk yönetim stratejileri geliştirmesini zorlaştırmakta ve ortak performans değerlendirme ölçütlerine duyulan ihtiyacı artırmaktadır [35], [36].

e) *Araştırma açıkları*: Mevcut literatürde, kanıta dayalı yorgunluk yönetiminin geliştirilmesini sınırlayan çeşitli araştırma boşlukları bulunmaktadır. Özellikle yorgunluk izleme sistemlerinin uzun dönemli güvenlik performansı üzerindeki etkilerini değerlendiren saha çalışmalarının sayısı sınırlıdır. Yayınlanan araştırmaların önemli bir bölümü, sistem geliştirme süreçlerine, laboratuvar doğrulamalarına veya kısa süreli saha uygulamalarına odaklanırken, teknolojilerin uzun vadeli operasyonel etkinliği ve sürdürülebilir güvenlik çıktıları yeterince incelenmemiştir [5], [36]. Ayrıca görsel tabanlı, fizyolojik ve davranışsal yorgunluk tespit yöntemlerinin farklı madencilik koşullarındaki göreli üstünlükleri konusunda kesin bir görüş birliği bulunmamaktadır ve mevcut çalışmaların çoğu belirli yöntemlere odaklanmaktadır [5], [42]. Bunun yanında, çalışanlar arasındaki bireysel yorgunluk duyarlılığı, uyku gereksinimi ve sirkadiyen uyum farklılıklarının daha iyi anlaşılması, kişiselleştirilmiş yorgunluk yönetimi stratejilerinin geliştirilmesi açısından gelecekteki araştırmaların öncelikli alanlarından biri olarak değerlendirilmektedir [6], [35].

VI. SONUÇLAR

Madencilikte yorgunluk tespiti ve yönetimi, son yirmi yılda öznel değerlendirmeler ve olaya tepkisel müdahaleden, gelişmiş gerçek zamanlı izleme ve proaktif risk yönetimine doğru kayda değer biçimde ilerlemiştir. Bu literatür incelemesi, tespit teknolojilerindeki önemli gelişmeleri, madencilik bağlamında yorgunluk mekanizmalarına ilişkin anlayışın derinleşmesini ve etkili yorgunluk yönetiminin sosyo-teknik karmaşıklığının giderek daha fazla fark edildiğini ortaya koymaktadır.

Kamyon sürücüleri için görüş temelli sistemler, mobil ve yeraltı çalışanları için giyilebilir fizyolojik monitörler ve operasyonel verilerden yararlanan davranışsal izleme olmak üzere üç temel teknolojik yaklaşım öne çıkmaktadır. Her yaklaşım farklı avantajlar sunmakta, ancak doğruluk, güvenilirlik, operatör kabulü ve örgütsel bütünleşmeyle ilgili belirli zorluklar içermektedir. Kanıtlar, birden fazla tespit

yaklaşımını birleştiren çok modlu sistemlerin, tek modlu sistemlere kıyasla üstün performans sağladığını göstermektedir.

Vardiya düzenleri ve sirkadiyen ritmin etkilerine ilişkin bulgular, yalnızca vardiya süresi sınırlarına dayanan yorgunluk yönetimi yaklaşımlarının yetersiz olduğunu göstermektedir. Araştırmalar, yorgunluk üzerinde sirkadiyen zamanlama ve vardiya döngüleri boyunca biriken etkilerin, tek bir vardiyada geçirilen çalışma süresinden daha belirleyici olduğunu ortaya koymaktadır [6], [9], [10]. Bu bulgular, izleme teknolojisinin tamamlayıcı bir savunma katmanı sağladığı vardiya tasarımının, birincil yorgunluk yönetimi stratejisi olarak güçlendirilmesini desteklemektedir.

Sosyo-teknik sistemler perspektifi, yorgunluk yönetiminin etkinliğini anlamak için temel bir çerçeve sunmaktadır. Teknolojik yetkinlik tek başına başarılı bir uygulamayı garanti etmemekte; operatör kabulü, örgüt kültürü, yönetim taahhüdü ve uygulamanın mevcut güvenlik yönetim sistemleriyle bütünleşmesi sonuçları belirgin biçimde etkilemektedir [6]. Gözetim aracı olarak algılanan sistemler dirençle karşılaşırken, güvenlik destek mekanizmaları olarak konumlandırılan sistemler kabul ve etkinlik kazanmaktadır.

Alana ilişkin önemli ilerlemelere rağmen kayda değer zorluklar varlığını sürdürmektedir. Çevresel sağlamlık, yanlış alarm oranları ve sistem güvenilirliğiyle ilgili teknik sınırlamalar, sürekli araştırma ve geliştirme gerektirmektedir. Operatör kabulü, veri yönetimi ve operasyonel iş akışlarıyla bütünleşme gibi örgütsel zorluklar, insan faktörlerine ve değişim yönetimine sistematik biçimde odaklanılmasını gerektirmektedir. Uzun vadeli etkinlik, en uygun müdahale stratejileri ve yorgunluğa duyarlılıktaki bireysel farklılıklar gibi konulardaki araştırma boşlukları, daha fazla inceleme gereksinimine işaret etmektedir.

Geleceğe yönelik yönelimler, yapay zekâ destekli tahmin sistemlerini, çok modlu algılayıcı kaynaşımını, kişiselleştirilmiş yorgunluk riski değerlendirmesini ve otonom araç sistemleriyle bütünleşmeyi içermektedir. En az bunlar kadar önemli olan, vardiya planlaması, işyeri kültürü ve kanıta dayalı politika çerçevelerine ilişkin örgütsel yeniliklerdir. Teknolojik gelişmelerin, yorgunluk mekanizmalarının ve sosyo-teknik sistem prensiplerinin daha iyi anlaşılmasıyla birleşmesi, madencilik güvenliği ve operasyonel verimlilikte anlamlı iyileştirmeler vaat etmektedir.

Madencilikte etkili yorgunluk yönetimi endüstri, araştırmacılar, düzenleyiciler ve teknoloji geliştiricilerinden sürekli bir bağlılık gerektirmektedir. Maden endüstrisi, gelişmiş izleme teknolojilerini kanıta dayalı örgütsel uygulamalar ve iş gücüyle etkileşimli iş birliğiyle birleştirdiğinde, işçi sağlığı ve yorgunlukla ilişkili riskleri önemli ölçüde azaltabilir.

KAYNAKLAR

[1] T. Bauerle, Z. Dugdale and G. Poplin, “Mineworker fatigue: A review of what we know and future decisions”, *Mining engineering*, 70(3), 3, March. 2018.

[2] D. Dawson, A. K. Searle and J. L. Paterson, “Look before you (s) leep: evaluating the use of fatigue detection technologies within a fatigue risk management system for the road transport industry”, *Sleep medicine reviews*, 18(2), 141-152. March. 2014. http://dx.doi.org/10.1016/j.smrv.2013.03.003

[3] W. P. Rogers, J. Marques, E. Talebi and F. A. Drews, “Iot-enabled wearable fatigue-tracking system for mine operators”, *Minerals,* 13(2), 287, Feb. 2023. https://doi.org/10.3390/min13020287

[4] T. Horberry, R. Burgess-Limerick and L. J. Steiner, “Human factors for the design, operation, and maintenance of mining equipment”, *CRC Press*. 2011.

[5] T. J. Bauerle, J. J. Sammarco, Z. J. Dugdale and D. Dawson, “The human factors of mineworker fatigue: An overview on prevalence, mitigation, and what's next”, *American journal of industrial medicine*, 65(11), 832-839. Nov. 2022. http://doi:10.1002/ajim.23301.

[6] D. Dawson and K. McCulloch, “Managing fatigue: it's about sleep”, *Sleep medicine reviews*, *9*(5), 365-380, Oct. 2005.

[7] J. A. Caldwell, J. L. Caldwell, L. A. Thompson and H. R. Lieberman, “Fatigue and its management in the workplace”, *Neuroscience & Biobehavioral Reviews*, *96*, 272-289, 2019. https://doi.org/10.1016/j.neubiorev.2018.10.024

[8] C. James, R. Tynan, D. Roach, L. Leigh, C. Oldmeadow, M. Rahman and B. Kelly, “Correlates of psychological distress among workers in the mining industry in remote Australia: Evidence from a multi-site cross-sectional survey”, *PLoS One*, *13*(12), e0209377, Dec. 2018. https://doi.org/10.1371/journal. pone.0209377

[9] T. Åkerstedt, “Shift work and disturbed sleep/wakefulness”, *Occupational medicine*, 53(2), 89-94, March. 2003.

[10] A. Williamson, D. A. Lombardi, S. Folkard, J. Stutts, T. K. Courtney and J. L. Connor, “The link between fatigue and safety”, *Accident Analysis & Prevention*, *43*(2), 498-515, March. 2011. https://doi.org/10.1016/j.aap.2009.11.011

[11] M. W. Johns, “A new method for measuring daytime sleepiness: the Epworth sleepiness scale”, *sleep*, 14(6), 540-545, Nov. 1991. https://doi.org/10.1093/sleep/14.6.540

[12] P. Sharma, J. C. Justus, M. Thapa and G. R. Poudel, "Sensors and systems for monitoring mental fatigue: A systematic review", *arXiv preprint arXiv:2307.01666*, Sep. 2023. https://doi.org/10.48550/arXiv.2307.01666
[13] B.R. Pratomo and N. B. Puspıtasarı, "Analisis penyebab kelelahan operator haul dumptruck (hd)(studi kasus di pt x rantau nangka kalimantan selatan), *Doctoral dissertation, Diponegoro University*, 2014.
[14] T. Mulyati, P. D. Sentia, A. Maulana and F. Erwan, "Fatigue analysis of high dump truck operators in indonesia's coal mining industry: A case study", *Malaysian Journal of Public Health Medicine*, 20(Special1), 38-44, 2020. https://doi.org/10.37268/mjphm/vol.20/no.Special1/art.666
[15] F. A. Drews, W. P. Rogers, E. Talebi and S. Lee, "The experience and management of fatigue: a study of mine haulage operators", *Mining, Metallurgy & Exploration*, 37(6), 1837-1846, July. 2020. https://doi.org/10.1007/s42461-020-00259-w
[16] D. Dawson and K. McCulloch, "Managing fatigue: it's about sleep", *Sleep medicine reviews*, *9*(5), 365-380, Oct. 2005. https://doi.org/10.1016/j.smrv.2005.03.002
[17] E. Sun, A. Nieto and Q. Li, "The drive fatigue pattern monitor for haul truck drivers in surface mining operations", *12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD)* (pp. 1329-1334). IEEE, August. 2015. 10.1109/FSKD.2015.7382136
[18] E. Sun, "The fuzzy neural network based haul truck driver fatigue detection in surface mining", *13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)* (pp. 1522-1526). IEEE, July. 2017. 10.1109/FSKD.2017.8392991
[19] M. Kowal, S. Kubal, P. Piotrowski, K. Staniec and M. Nowicki, "Wireless platform for the fatigue measurement of workers employed in underground mines", *Elektronika ir Elektrotechnika*, *23*(4), 24-28, 2017. https://doi.org/10.5755/j01.eie.23.4.18718
[20] Y. Jiang, W. Chen, X. Zhang, X., Zhang and G. Yang, "Real-time monitoring of underground miners' status based on mine IoT system", *Sensors*, *24*(3), 739, Jan. 2024. https://doi.org/10.3390/s24030739
[21] Y. Wang, Z. He and L. Wang, "Truck driver fatigue detection based on video sequences in open-pit mines", *Mathematics*, *9*(22), 2908, Oct. 2021. https://doi.org/10.3390/math9222908
[22] E. Sun, A. Nieto, Z. Li, Z and V. Kecojevic, "An integrated information technology assisted driving system to improve mine trucks-related safety", *Safety science*, 48(10), 1490-1497, Dec. 2010. https://doi.org/10.1016/j.ssci.2010.07.012
[23] Y. Dong, Z. Hu, K. Uchimura and N. Murayama, "Driver inattention monitoring system for intelligent vehicles: A review", *IEEE transactions on intelligent transportation systems*, 12(2), 596-614, Dec. 2011. 10.1109/TITS.2010.2092770
[24] T. Li, B. Chang, Y. Sun, Y. Zhou, B. Zhou and L. Zhang, "Development and application of a generalized fatigue driving monitoring system for open-pit mining vehicles: A case study in a myanmar copper mine", *Results in Engineering*, *26*, 105206, June. 2025. https://doi.org/10.1016/j.rineng.2025.105206
[25] Q. Ji, Z. Zhu and P. Lan, "Real-time nonintrusive monitoring and prediction of driver fatigue", *IEEE transactions on vehicular technology*, *53*(4), 1052-1068. July. 2004. 10.1109/TVT.2004.830974
[26] D. R. Bongers, "An EEG-Based technique for operator fatigue determination", *The Australasian Institute of Mining and Metallurgy*. Jan. 2005.
[27] L. Hartley, T. Horberry, N. Mabbott and G. P. Krueger, "Review of Fatigue Detection and Prediction Technologies", *National Road Transport Commission, Melbourne, Australia*, ISBN: 0 642 54469 7, s, 41. Sep. 2000.
[28] N. Mabbott, "ARRB Pro-Active Fatigue Management System*"*, *Queensland Mining Industry Health & Safety Conference*, ARRB Transport Research Ltd. 2002.
[29] Y. Wang, F. Jiang and Y. Zhang, "Prediction of Safe Working Hours of Pumping Station Drivers in Coal Mine", *Journal of Engineering Science & Technology Review*, 15(1), 162-167, 2022. https://doi:10.25103/jestr.151.20
[30] K. R. Espinoza Ugarte and O. A. Muñoz Sánchez, "*Control de fatiga y posicionamiento de flota de acarreo mediante el sistema Wombat – minería superficial",* (Bachelor's thesis, Universidad Privada del Norte, Peru), 2016. https://hdl.handle.net/11537/10684
[31] M. S. L. Paulo, H. W. S. Cabral, J. Kill, W. F. Soares, E. S. B. Medeiros, T. C. Jesus and A. Lemos, "Computerized Assessment of Readiness (Focos/Prontos System) and Its Multifactor Screening for Worker's Fatigue and Productivity: A Case-Control and Cross-Sectional Study", *Journal of Advances in Medicine and Medical Research*, 36(5), 157–176, April. 2024. https://doi.org/10.9734/JAMMR/2024/v36i55426
[32] A. M. Hernández, J. A. M. Covarrubias and D. A. P. Guzmán, "Sistema inteligente para detección de fatiga y distracción en conductores de camion de acarreo pesado en minería de Cielo Abierto", *Pistas Educativas*, *38*(122), 290-303, 2016.
[33] S. H. Lee, "*Exploring fatigue management of haul truck drivers through a socio-technical perspective"* (Master's thesis, The University of Utah), 2022.
[34] N. Mabbott and B. Lloyd, "Driver fatigue through nightshifts in succession", *In rail achieving growth: conference proceedings* (pp. 145-152). Melbourne, Vic.: RTSA, Jan. 2006.
[35] P. Gander, L. Hartley, D. Powell, P. Cabon, E. Hitchcock, A. Mills and S. Popkin, "Fatigue risk management: Organizational factors at the regulatory and industry/company level", *Accident Analysis & Prevention*, *43*(2), 573-590, March. 2011. 10.1016/j.aap.2009.11.007
[36] T. Horberry, J. Harris, D. Cliff, P. Dodshon, J. Lee, N. Lim and K. Way, "Fatigue management in mining: current practices and future directions", *Journal of work health and safety regulation*, *2*(2), 145-155, March. 2024. DOI: 10.57523/jaohlev.oa.23-004

[37] Z. Dugdale, B. Eiter, C. Chaumont Menendez, I. Wong and T. Bauerle, “Findings from a systematic review of fatigue interventions: What's (not) being tested in mining and other industrial environments”, *American journal of industrial medicine*, 65(4), 248-261, April. 2022. https://doi.org/10.1002/ajim.23334
[38] K. Ker, P. J. Edwards, L. M. Felix, K. Blackhall and I. Roberts, “Caffeine for the prevention of injuries and errors in shift workers”, *Cochrane Database of Systematic Reviews*, 12(5), May. 2010. https://doi.org/10.1002/14651858.cd008508
[39] G. Maisey, M. Cattani, A. Devine, J. Lo and I. C. Dunican, “The sleep of shift workers in a remote mining operation: methodology for a randomized control trial to determine evidence-based interventions”, *Frontiers in neuroscience*, *14*, 579668, 1-11, January. 2021. https://doi.org/10.3389/fnins.2020.579668
[40] I. F. Akyildiz, Z. Sun and M. C. Vuran, “Signal propagation techniques for wireless underground communication networks”, *Physical Communication*, *2*(3), 167-183, Sep. 2009. https://doi.org/10.1016/j.phycom.2009.03.004
[41] I. F. Akyildiz and M. C. Vuran, “Wireless Underground Sensor Networks”, *In Wireless Sensor Networks*, 305–340. Wiley, 2010. DOI: 10.1002/9780470515181.ch17
[42] A. Sahayadhas, K. Sundaraj and M. Murugappan, “Detecting driver drowsiness based on sensors: a review”, *Sensors*, 12(12), 16937-16953, Dec. 2012. https://doi.org/10.3390/s121216937
[43] T. Horberry, “The health and safety benefits of new technologies in mining: A review and strategy for designing and deploying effective user-centred systems”, *Minerals*, *2*(4), 417-425, Oct. 2012. https://doi.org/10.3390/min2040417

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/icmar/home

Effect of Cutting Parameters on Temperature in Machining of AISI P20+S (DIN 1.2312) Steel

Muhammet Enver GÖKDEMİR*, Yunus Kaplan [2], Tayfun ÇETİN [3] and Mehmet TOPUZ [4]

[1 3]*Department of Electricity and Energy, Yuksekova Vocational High School, Hakkari University, Turkey*
[2] *Selçuk University, Faculty of Technology, Department of Mechanical Engineering, Konya*
[4]*Department of Mechanical Engineering, Van Yuzuncu Yil University, Turkey*

(menvergokdemir@hakkari.edu.tr)

Abstract – AISI P20+S steel is a sulfur-alloyed, pre-hardened plastic mold steel widely utilized in the plastic injection molding, tool making, and die-casting industries due to its pre-hardened structure and mechanical strength. As a modified version of standard AISI P20 steel (industrially designated as DIN 1.2312), sulfur (S) is incorporated into this material to significantly enhance its machinability. Sulfur inclusion facilitates higher cutting speeds during machining, providing substantial time savings, particularly in the production of large and complex molds. However, during the dry machining process, this material generates significant heat at the cutting zone, which can negatively affect tool life and the surface roughness of the material. In this study, the effects of feed rate (0.2 and 0.4 mm/rev), depth of cut (0.3 and 0.6 mm), and cutting speed (50 and 120 m/min) on the cutting temperature during the dry turning of AISI P20+S steel were investigated using a 2^3 full factorial design, supported by ANOVA and Taguchi's Signal-to-Noise (S/N) ratio methodology based on the "Smaller is better" quality characteristic. Based on the experimental and statistical analyses, the most influencing parameter on cutting temperature was measured to be depth of cut. Sample 7 generated significantly higher temperature compared to other samples. Except Sample 7, the effect of cutting parameters on temperature was found to be limited. The temperature variation was less than 10% for all samples when all parameters were evaluated separately, indicating that the effect of cutting parameters on AISI P20+S steel is limited.

Keywords – AISI P20+S, DIN 1.2312, turning, cutting parameters, surface roughness

I. INTRODUCTION

Turning is a fundamental material removal process to fabricate precise cylindrical components [1]. During the turning operation, high temperatures occur due to friction at the tool-workpiece interface [2]. This heat generation is highly localized and concentrated primarily at the primary shear zone, where plastic deformation occurs, and the secondary shear zone, which forms the tool-chip interface [3].

The machining performance is directly affected by the temperature, particularly in terms of dimensional accuracy and surface quality [4]. In addition, excessive heat generation at the tool tip accelerates tool wear through mechanisms such as diffusion, adhesion, and crater wear, thereby reducing tool life [5]. Moreover, elevated cutting temperatures can induce thermal damage to the machined workpiece surface, leading to microstructural alterations, tensile residual stresses, and poor surface integrity [6]. To dissipate heat and provide lubrication to the cutting zone, cutting fluids are preferred. However, extra costs, environmental regulations, and health hazards limit the usage of fluids and force

manufacturers to shift toward alternatives such as dry machining [7]. The complete elimination of cutting fluids requires precise temperature control to prevent a reduction in the quality of the products.

Cutting parameters, particularly cutting speed, feed rate, and depth of cut, influence the tendency toward built-up edge (BUE) formation due to cutting forces, which significantly affect the final product [8]. An increase in cutting parameters elevates the temperature generated at the tool-workpiece interface and may have a negative effect on the machined parts. Korkut and Donertas [9] investigated the cutting forces, surface roughness, and tool-chip contact length during the face milling of AISI 1020 and AISI 1040 steels. Their study revealed that increasing the cutting speed significantly reduced the surface roughness for both materials. However, AISI 1040 exhibited higher roughness values than AISI 1020. Furthermore, in alignment with the mechanics of BUE formation, they observed that lower and intermediate cutting speeds increased the built-up edge tendency, which consequently reduced the tool-chip contact length and altered the resulting cutting forces. Rao et al. [10] investigated the effects of cutting speed, feed rate, and depth of cut on cutting force and surface roughness during the dry turning of hardened AISI 1050 steel using mixed ceramic tools (Al_2O_3+TiC). Taguchi L27 analysis of variance (ANOVA) results indicated that the feed rate was the most effective parameter influencing both cutting force and surface roughness. Depth of cut was found to be effective on cutting force, yet the effect on the surface roughness was found to be insignificant. Zha et al. **Hata! Başvuru kaynağı bulunamadı.** studied the influence of cutting feed rate on the surface integrity of Ti-6Al-4V alloy. They analyzed the cutting forces, cutting temperature, surface roughness, hardness, residual stress, and microstructure of the samples. In addition, they simulated the cutting force, temperature, and residual stress and compared them with experimental measurements, resulting in less than a 10% average error between the experiments and simulations. Specifically, an increase in feed rate was shown to increase the surface roughness, with a rapid increase observed above 0.1 mm/rev.

Although various studies in the literature have focused on the machining characteristics, such as cutting forces and surface roughness of different steel types and alloys, research evaluating the thermal behavior of AISI P20+S (DIN 1.2312) plastic mold steel under dry turning conditions remains limited. Understanding the temperature distributions at the tool-workpiece interface for this pre-hardened material is essential to optimize tool life and achieve high dimensional accuracy without using hazardous cutting fluids. To address this gap, this study investigates the individual and combined effects of cutting speed (50 and 120 m/min), feed rate (0.2 and 0.4 mm/rev), and depth of cut (0.3 and 0.6 mm) on the cutting temperature during the dry turning of AISI P20+S steel. The turning operations were performed using a CCMT-09T308-304 cutting tool with a 0.8 mm nose radius, and a full factorial experimental design was employed to map the thermal characteristics of the process.

II. Materials and Method

CCMT-09T308-304 cutting tool was preferred on AISI P20+S steel of 200 mm length for turning operation. Temperature sensor used for temperature measurement. Eight samples for two different feed rates (0.2 and 0.4 mm/rev), depth of cut (0.3 and 0.6 mm) and cutting speed (50 and 120 m/min) was employed. The temperature values were compared via fixing two parameters to obtain the effect of feed rate, depth of cut and cutting speed individually.

Table 1. Material content of the AISI 5115 P20+S steel

wt.%	C	Cr	Mn	S	Si	Mo	P
AISI P20+S	0.35-0.45	1.80-2.00	1.40-1.60	0.05-0.10	0.30-0.50	0.15-0.25	≤ 0.030 (Max.)

Dry machining refers to machining operations without using coolant or lubricant supply, avoiding extra costs for coolant and chip recycling costs [11]. Under controlled temperature setting via optimized cutting parameters, dry turning is widely preferred on many turning applications due to environmentally friendly, improving workplace safety and operator health benefits, and cost-effective properties compared to wet turning [12]. The dry turning experimental methodology is shown in the Figure 1.

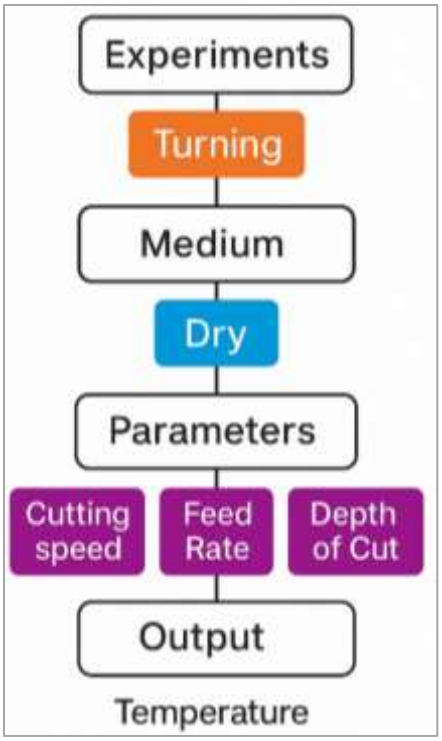


Figure 1. Experimental methodology

Table 2. Cutting Parameters

Sample	Feed Rate (mm/rev)	Depth of Cut (mm)	Cutting Speed (m/min)	Temperature (°C)
1	0.2	0.3	50	365.6
2	0.2	0.3	120	365
3	0.2	0.6	50	372.6
4	0.2	0.6	120	363.6
5	0.4	0.3	50	361.5
6	0.4	0.3	120	364.9
7	0.4	0.6	50	415.3
8	0.4	0.6	120	379.8

III. RESULTS

Experimental results indicate that increase in feed rate have no significant effect on temperature as shown in Table 2. Even with higher depth of cut and cutting speed, the temperature change is in the range of 0-11% and 3.69% in average. However, a significant peak in temperature is recorded in Sample 7, which can be attributed to the non-homogeneous distribution of MnS inclusions in the AISI P20+S steel and may cause instantaneous hard-spot formations during shearing [12]. Furthermore, this behavior may stem from friction at micro-scale chip clogging or the early onset of tool wear at that specific cutting combination, which accelerates plastic deformation and causes excessive heat [13]. A similar trend is observed in the analysis of the depth of cut. When Sample 7 is excluded, the temperature variation remains within a narrow range of 0.3–4% while keeping other parameters constant. However, when all tests are evaluated collectively, the average temperature change is found to be approximately 5%. This significant difference indicates that the depth of cut has only a slight effect on the cutting temperature within the investigated range. The effect of cutting speed on the temperature variation is remarkably

lower, ranging between 0% and 8% with an overall average of 2.54%, despite a 2.4-fold increase in the cutting speed. This is attributed to the self-lubricating behavior of MnS inclusions in the AISI P20+S steel. Elevated cutting speeds deform MnS and form a solid lubricating film on the tool-chip interface, which drastically reduces the friction coefficient. In addition, at higher speeds, a major portion of the generated heat is rapidly carried away from the surface of the sample by the detached chips before it can conduct into the bulk of the workpiece and maintains a stable temperature profile. When the sample 7 was excluded, the temperature difference between the lowest and highest was found to be 5%.

The temperature values of each samples exhibit slight variations except Sample 7. As shown in Figure 2, the cutting parameters have no significant effect on the temperature of samples.

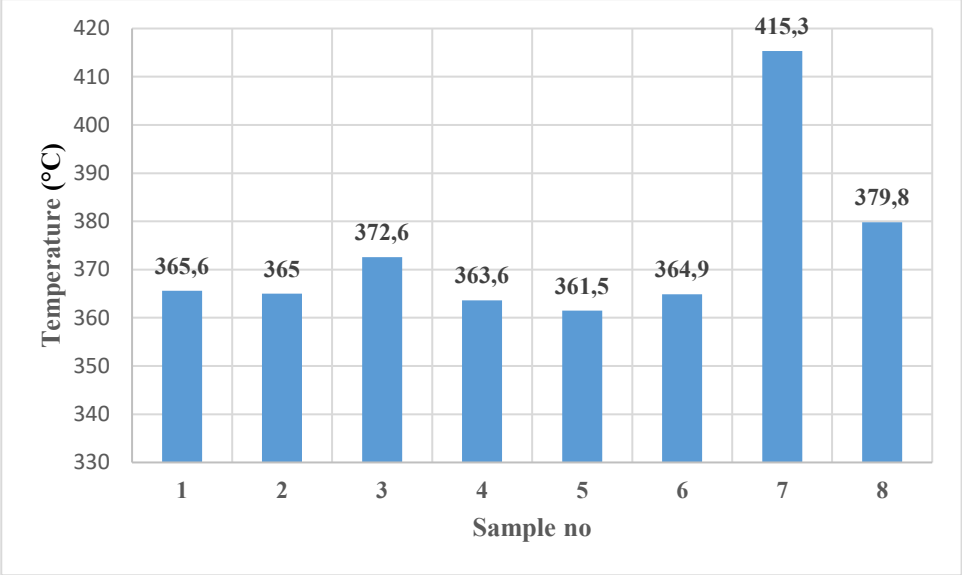


Figure 2. Variation of cutting temperature depending on cutting parameters

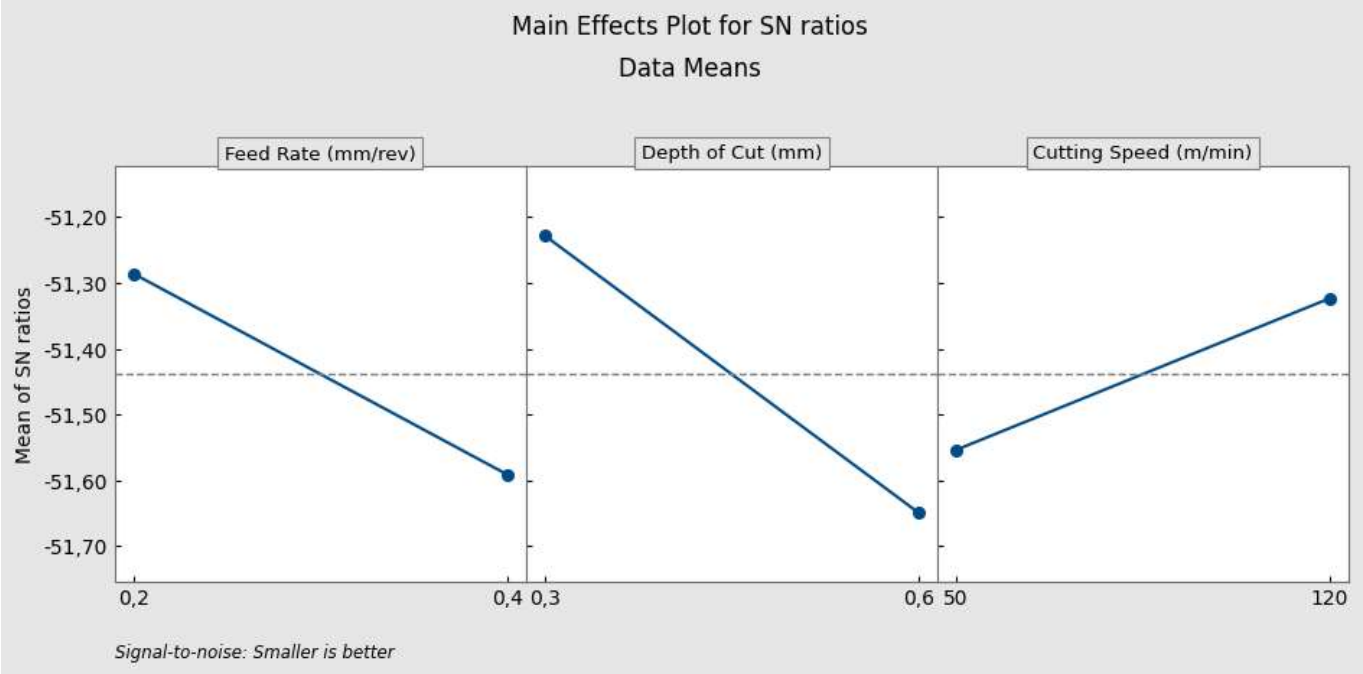


Figure 3. S/N ratio graph for temperature

The main effects plot for S/N ratios visually represents the influence of each machining parameter on temperature variation under the "Smaller is better" criterion, where higher S/N values correspond to lower process temperatures in Figure 3. An increase in feed rate and depth of cut decreases the mean of S/N ratios and resulting in a downward slope, which indicates an increase in temperature. Conversely, an

increase in cutting speed shows an upward slope, validating the temperature drop. As a result, the optimum cutting parameters for minimum temperature were determined to be 0.2 mm/rev for feed rate, 0.3 mm for depth of cut, and 120 m/min for cutting speed.

Table 3. Statistical evaluation of the effect of parameters on temperature (Analysis of Variance)

Source	DF	Seq SS	Adj SS	Adj MS	F	P	Contribution
Feed Rate (mm/rev)	1	374.01	374.01	374.01	1.56	0.279	16.71
Depth of Cut (mm)	1	690.06	690.06	690.06	2.88	0.165	30.83
Cutting Speed (m/mm)	1	217.36	217.36	217.36	0.91	0.394	9.71
Residual Error	4	956.92	956.92	239.23	null	null	42.75
Total	7	2238.36	null	null	null	null	100.00
Model Summary	S = 15.46	R^2: 57.25%		$R^2_{(adj)}$: 25.19%		$R^2_{(pred)}$: -71%	

An Analysis of Variance (ANOVA) was performed to evaluate the statistical significance of the cutting parameters and determine their relative percentage contributions to the variation in cutting temperature as shown in Table 3. Depth of Cut was found to be the most influential factor on temperature, accounting for 30.83% of the variation (F = 2.88), followed by Feed Rate at 16.71% (F = 1.56) and Cutting Speed at 9.71% (F = 0.91). Because all p-values exceed the 0.05 threshold, none of the individual main effects show statistical significance at a 95% confidence level, primarily due to the limited degrees of freedom (d_f = 4 for error) and a substantial unmodeled residual error contribution of 42.75%. The model achieves a coefficient of determination (R^2) of 57.25% (R^2_{adj} = 25.18%) with a standard error S = 15.47 °C.

Table 4. Response table for Signal-to-Noise Ratios of for temperature

Level	Feed Rate (mm/rev)	Depth of Cut (mm)	Cutting Speed (mm)
1	-51.29	-51.23	-51.55
2	-51.59	-51.65	-51.32
Delta	0.31	0.42	0.23
Rank	2	1	3

The Taguchi response table prioritizes the control factors based on their Delta values, which measure thermal sensitivity under the Smaller is better criterion as shown in Table 4. Depth of Cut is ranked first with a Delta of 0.42 dB, confirming it as the most influential factor in temperature control. Feed Rate ranks second with a Delta of 0.31 dB, while Cutting Speed ranks third with a Delta of 0.23 dB, indicating the least effect on thermal output.

IV. DISCUSSION

The experimental test results showed the thermal variation of the AISI P20+S steel during machining is minimal. With all cutting parameters, the temperature values were found similar (except sample 7) which is less than 5% for each individual parameter. Considering all samples, the temperature difference between highest and lowest values was found about 14%. When compared to AISI P20 steel, the sulfur addition in AISI P20+S (DIN 1.2312) improves machinability primarily by facilitating chip formation and reducing localized friction. The slight temperature variation validates the benefits of sulfur addition in maintaining thermal conditions, even when cutting parameters are varied. This provides a clear advantage in tool life and process predictability for industrial mold manufacturing.

V. CONCLUSION

In this study, the effect of cutting parameters on the machining of AISI P20+S (DIN 1.2312) steel were evaluated. The main conclusions of the study are as follows.

- The sulfur addition AISI P20+S provided more stable machining performance by maintaining stable thermal behavior during different cutting conditions with different parameters.
- The temperature variation when sample 7 was excluded was found to be 5%, however, when all samples are included, the variation was about 14% between the lowest and highest temperature.
- The effect of cutting parameters on AISI P20+S was measured limited effect on temperature.

- The overall thermal stability reinforces the suitability of AISI P20+S steel for manufacturing large and complex plastic molds efficiently.
- The ANOVA confirms Depth of Cut as the most influencing factor on temperature, which is consistent with the experimental results.
- Taguchi S/N ratio analysis ranked Depth of Cut as the most sensitive parameter on thermal output (Delta = 0.42 dB), followed by Feed Rate (Delta = 0.31 dB) and Cutting Speed (Delta = 0.23 dB).
- Based on Taguchi's "Smaller is better" criterion, the optimal cutting parameter combination to achieve minimum cutting temperature was determined as Feed Rate = 0.2 mm/rev, Depth of Cut = 0.3 mm, and Cutting Speed = 120 m/min.

REFERENCES

[1] He, C. L., Zong, W. J., & Zhang, J. J. (2018). Influencing factors and theoretical modeling methods of surface roughness in turning process: State-of-the-art. International Journal of Machine Tools and Manufacture, 129, 15-26, https://doi.org/10.1016/j.ijmachtools.2018.02.001.

[2] Abukhshim, N. A., Mativenga, P. T., & Sheikh, M. A. (2005). Investigation of heat partition in high speed turning of high strength alloy steel. *International Journal of Machine Tools and Manufacture*, *45*(15), 1687-1695, https://doi.org/10.1016/j.ijmachtools.2005.03.008.

[3] Sun, J., Ke, Q., & Chen, W. (2019). Material instability under localized severe plastic deformation during high speed turning of titanium alloy Ti-6.5 AL-2Zr-1Mo-1V. Journal of Materials Processing Technology, 264, 119-128, https://doi.org/10.1016/j.jmatprotec.2018.09.002.

[4] Suhail, A. H., Ismail, N., Wong, S. V., & Jalil, N. A. (2010). Optimization of cutting parameters based on surface roughness and assistance of workpiece surface temperature in turning process. American journal of engineering and applied sciences, 3(1), 102-108, https://doi.org/10.3844/ajeassp.2010.102.108.

[5] Demirpolat, H., Binali, R., Patange, A. D., Pardeshi, S. S., & Gnanasekaran, S. (2023). Comparison of Tool Wear, Surface Roughness, Cutting Forces, Tool Tip Temperature, and Chip Shape during Sustainable Turning of Bearing Steel. Materials, 16(12), 4408. https://doi.org/10.3390/ma16124408.

[6] Elsheikh, A.H., Shanmugan, S., Muthuramalingam, T. et al. A comprehensive review on residual stresses in turning. Adv. Manuf. 10, 287–312 (2022). https://doi.org/10.1007/s40436-021-00371-0.

[7] Adler, D. P., Hii, W. W. S., Michalek, D. J., & Sutherland, J. W. (2006). EXAMINING THE ROLE OF CUTTING FLUIDS IN MACHINING AND EFFORTS TO ADDRESS ASSOCIATED ENVIRONMENTAL/HEALTH CONCERNS. Machining Science and Technology, 10(1), 23–58. https://doi.org/10.1080/10910340500534282.

[8] Kuntoğlu, M., Binali, R., Demirpolat, H. et al. Built-up-edge Formation and its Effect on Surface Topography and Machinability Indicators in Sustainable Minimum Quantity Lubrication Turning of Al2024-T6. J. of Materi Eng and Perform 35, 9066–9077 (2026). https://doi.org/10.1007/s11665-025-12313-z.

[9] Korkut, I., & Donertas, M. A. (2007). The influence of feed rate and cutting speed on the cutting forces, surface roughness and tool–chip contact length during face milling. *Materials & design*, *28*(1), 308-312, https://doi.org/10.1016/j.matdes.2005.06.002.

[10] Rao, C. J., Rao, D. N., & Srihari, P. (2013). Influence of cutting parameters on cutting force and surface finish in turning operation. Procedia Engineering, 64, 1405-1415, https://doi.org/10.1016/j.proeng.2013.09.222.

[11] Zha, X., Qin, H., Yuan, Z. et al. Effect of cutting feed rate on machining performance and surface integrity in cutting process of Ti-6Al-4V alloy. Int J Adv Manuf Technol 131, 2791–2809 (2024). https://doi.org/10.1007/s00170-023-12458-y.

[12] Machado, T. J. (2019). Machinability of Crankshaft Steel: On the Influence of Batch-to-Batch Material Variations [Master's thesis, Chalmers University of Technology]. Chalmers Publication Library, https://odr.chalmers.se/bitstreams/b0e80d52-6fc1-41ea-923d-32c32fe6cf5b/download.

[13] Raval, P., Patel, D., Prajapati, R., Badheka, V., Gupta, M. K., & Khanna, N. (2022). Energy consumption and economic modelling of performance measures in machining of wire arc additively manufactured Inconel-625. Sustainable Materials and Technologies, 32, e00434. https://doi.org/10.1016/j.susmat.2022.e00434

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

TEK SAĞLIK ALANINDAKİ ARAŞTIRMA MAKALELERİNİN BİBLİYOMETRİK ANALİZİ

Mustafa Nal[1*], Veli Durmuş[2] ve İ. Serden Başak[3]

[1] *Sağlık Yönetimi Bölümü, Sağlık Bilimleri Fakültesi, Kütahya Sağlık Bilimleri Üniversitesi, Türkiye*
[2] *Sağlık Yönetimi Bölümü, Sağlık Bilimleri Fakültesi, Kütahya Sağlık Bilimleri Üniversitesi, Türkiye*
[3] *Sağlık Yönetimi Bölümü, Sağlık Bilimleri Fakültesi, Kütahya Sağlık Bilimleri Üniversitesi, Türkiye*
[*] *(mustafa.nal@ksbu.edu.tr)*

Özet – Bu çalışma, Tek Sağlık alanında yayımlanan bilimsel araştırmaları bibliyometrik yöntemlerle analiz ederek literatürün gelişim eğilimlerini, araştırma odaklarını ve uluslararası bilimsel yapısını ortaya koymayı amaçlamaktadır. Araştırma verileri Web of Science Core Collection veri tabanından elde edilmiştir. "One Health" anahtar kelimesi kullanılarak ulaşılan yayınlar VOSviewer programı ile analiz edilmiş; yayın yılları, araştırma alanları, anahtar kelime ağları, yayıncı kuruluşlar ve fon sağlayıcı kurumlar bibliyometrik göstergeler açısından değerlendirilmiştir. Analiz sonucunda Tek Sağlık alanındaki ilk araştırma makalesinin 1981 yılında yayımlandığı belirlenmiştir. Yayın sayısının özellikle 2010'lu yıllardan itibaren belirgin biçimde arttığı ve 2025 yılında en yüksek düzeye ulaştığı saptanmıştır. Araştırmaların en yoğun olduğu alanlar Halk Sağlığı, Çevre ve İş Sağlığı, Bulaşıcı Hastalıklar ve Veterinerlik Bilimleri olmuştur. Anahtar kelime analizinde en sık kullanılan terimler "one health", "antimicrobial resistance" ve "public health" olarak belirlenmiştir. Yayınların önemli bir bölümü Elsevier, Springer Nature ve MDPI gibi uluslararası yayıncı kuruluşlarda yayımlanmış; en fazla fon desteği sağlayan kuruluşlar arasında Birleşik Krallık Araştırma ve İnovasyon Kurumu, Avrupa Birliği ve Amerika Birleşik Devletleri Sağlık ve İnsan Hizmetleri Bakanlığı öne çıkmıştır. Bulgular, Tek Sağlık yaklaşımına yönelik bilimsel ilginin son yıllarda önemli ölçüde arttığını ve alanın disiplinler arası, uluslararası ve stratejik bir araştırma alanı haline geldiğini göstermektedir.

Anahtar Kelimeler – Tek Sağlık, Bibliyometrik Analiz, Vosviewer, Web Of Science, Halk Sağlığı, Çevre Sağlığı.

I. GİRİŞ

Her yıl dünya çapında milyonlarca insan ve hayvan zoonotik hastalıklardan olumsuz yönde etkilenmektedir (CDC, 2025). Tek Sağlık kavramının 2000'li yılların başlarında literatüre girmesinin ardından, bu yaklaşımın küresel, bölgesel ve ulusal düzeylerde hayata geçirilmesine yönelik girişimler belirgin biçimde artış göstermiştir (Khan et al., 2018; Mwatondo et al., 2023). Tek Sağlık, insanların sağlığının hayvanların ve paylaştığımız çevrenin sağlığıyla yakından bağlantılı olduğunu kabul eden bir yaklaşımdır. Tek Sağlık yeni bir kavram değildir, ancak son yıllarda daha da önem kazanmıştır. Bunun nedeni, insanlar, hayvanlar, bitkiler ve çevremiz arasındaki etkileşimleri değiştiren birçok faktörün ortaya çıkmasıdır (CDC, 2025). İnsan, hayvan ve çevre sağlığını birlikte ele alan Tek Sağlık (One Health) yaklaşımı, çevresel kaynaklı sağlık tehditlerinin önlenmesi ve yönetilmesinde giderek daha yaygın biçimde benimsenen bütüncül bir yaklaşım haline gelmiştir (Sweileh, 2020; Zhao et al., 2022). Tek Sağlık, insan, hayvan ve ekosistemlerin sağlığını sürdürülebilir bir şekilde dengelemeyi ve optimize etmeyi amaçlamaktadır (Milazzo et al., 2025).

Bu yaklaşım bütüncül bir bakış açısıyla iklim değişikliği, biyolojik çeşitlilik kaybı, hava ve su kirliliği, zoonotik hastalıklar ve antimikrobiyal direnç gibi farklı problemlerin tek bir sektörün gayretiyle üstesinden gelinemeyeceğini net bir şekilde ortaya koymaktadır (Cascajares et al., 2021; Mishra et al., 2023; Özen Çınar & Özkaya Bozkurt, 2026). Birleşmiş Milletler Gıda ve Tarım Örgütü (FAO), Dünya Hayvan Sağlığı Örgütü (WOAH) ve Dünya Sağlık Örgütü (DSÖ) tarafından 2010 yılında oluşturulan Üçlü İttifak, insan, hayvan ve çevre sağlığı arasındaki etkileşimlerin etkin biçimde yönetilebilmesi için kurumsal düzeyde iş birliğinin gerekliliğini vurgulamıştır (Mwatondo et al., 2023). Mart 2022'de Üçlü İş

Birliği mekanizmasına Birleşmiş Milletler Çevre Programı'nın (UNEP) katılmasıyla oluşturulan Dörtlü İşbirliği (Quadripartite), Tek Sağlık yaklaşımı kapsamında insan, hayvan, bitki ve ekosistem sağlığı arasındaki etkileşimlerin koordineli biçimde ele alınmasını destekleyen küresel bir yönetişim çerçevesi oluşturmuştur (WHO, 2022). Tek Sağlık politikaları, iklim değişikliğine uyum stratejilerinin geliştirilmesi, zoonotik hastalıkların önlenmesi, çevre kirliliğinin azaltılması, gıda güvenliğinin sağlanması ve sürdürülebilir kent planlamasının desteklenmesi gibi birçok alanda Tek Sağlık yaklaşımının uygulamaya geçirilmesini mümkün kılan temel politika araçları arasında yer almaktadır (Miao et al., 2022; Raphael et al., 2024).

Literatürde Tek Sağlık yaklaşımını ele alan sınırlı sayıda sistematik incelemeler bulunmaktadır. Şimşir ve Mete (2022), Web of Science verilerini SCIMAT bilim haritalama programı ile analiz ederek 2021 yılına kadar olan yayınların tematik gelişimini incelemiştir (Şimşir & Mete, 2022). Yalçın ve Yalçın (2026), Web of Science veri tabanında yer alan Ankara Üniversitesi Veteriner Fakültesi Dergisi yayınlarını R Studio ortamında Bibliometrix/Biblioshiny paketi kullanarak bibliyometrik yöntemlerle analiz etmiş ve araştırma temalarının zaman içinde geleneksel veterinerlik konularından enfeksiyon hastalıkları, moleküler tanı yöntemleri, çevre sağlığı ve sürdürülebilirlik odaklı alanlara yöneldiğini ortaya koymuştur (Yalçın & Yalçın, 2026). Mwatondo ve ark. (2023) tarafından gerçekleştirilen sistematik inceleme, COVID-19 sonrası dönemde Tek Sağlık ağlarına yönelik yatırımların arttığını, ancak bu ağların coğrafi ve sektörel dağılımının dengeli olmadığını ortaya koymuştur. Çalışma, Tek Sağlık iş birliklerinin daha kapsayıcı ve eşitlikçi biçimde yapılandırılması gerektiğini vurgulamaktadır (Mwatondo et al., 2023). Milazzo (2025) tarafından gerçekleştirilen sistematik inceleme, Tek Sağlık yaklaşımının insan, hayvan ve çevre sağlığını bütüncül biçimde ele alan çok sektörlü bir iş birliği modeli olduğunu ortaya koymakta; mevcut uygulamaların ise büyük ölçüde insan ve hayvan sağlığına odaklandığını göstermektedir (Milazzo et al., 2025).

Bu doğrultuda çalışmanın amacı, Tek Sağlık alanında yayımlanan bilimsel araştırmaların bibliyometrik analizini gerçekleştirerek literatürün mevcut yapısını ve zaman içindeki gelişimini ortaya koymaktır. Bu kapsamda yayınların yıllara göre değişimi, öne çıkan araştırma temaları ile alana en fazla katkı sağlayan yazarlar, kurumlar ve ülkeler analiz edilmiştir.

II. MATERYAL VE YÖNTEM

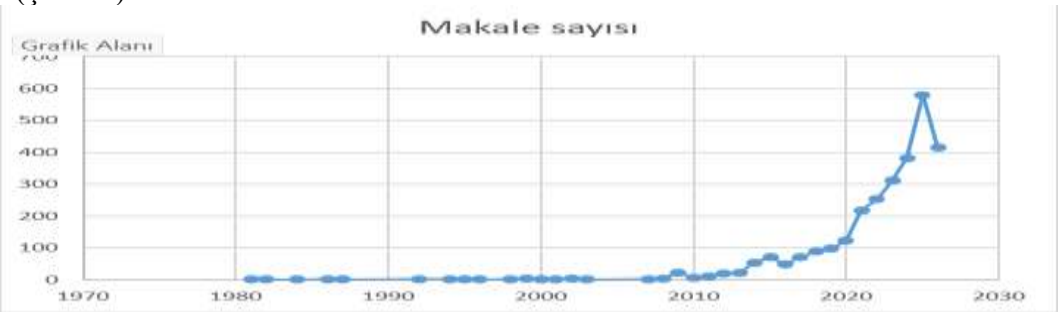
Mevcut çalışmada, bibliyometrik analiz tekniğinden faydalanılmıştır. Veri analiz sürecinde VOSviewer 1.6.18 yazılımı kullanılmıştır. Araştırma kapsamında veriler, Web of Science veri tabanında "başlık" seçeneği kullanılarak "One Health" anahtar kelimesi taratılmıştır. Yapılan tarama sonucunda toplam 2857 araştırma makalesine ulaşılmıştır (Web of Science, 2026). İlgili veriler 12 Temmuz 2026 tarihinde derlenmiştir. Analiz sürecinde ise kümeleme, yoğunluk ve ağ haritalama teknikleri kullanılmıştır.

III. BULGULAR

Bu çalışmaya ait bulgular aşağıda sunulmuştur.

Yıllara Göre Eser Sayıları

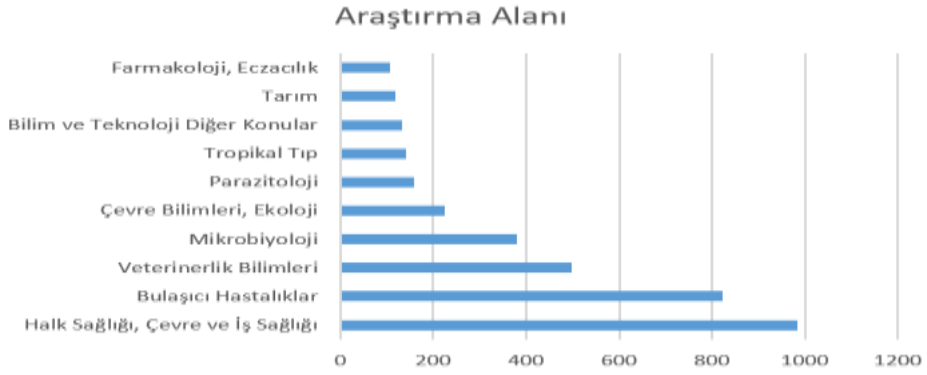
Web of Science veri tabanında yayımlanan eserler incelendiğinde Tek Sağlık konusundaki ilk araştırma makalesinin 1981 yılında yayımlandığı belirlenmiştir. Alandaki yayın sayısının uzun yıllar bir artış olmamasına rağmen 2010'lu yıllarda artışa geçtiği ve 2025 (f = 580) yılında en yüksek seviyeye çıktığı belirlenmiştir (Şekil 1).



Şekil 1. Yıllara göre makale sayısı

Eser Araştırma Alanı

Tek Sağlık alanındaki araştırma makalelerinin 110 farklı bilim alanında gerçekleştiği belirlenmiştir. En fazla yayına sahip ilk 10 araştırma alanına ilişkin dağılım Şekil 2'te gösterilmiştir. Tek Sağlık konusundaki yayınların en çok Halk Sağlığı, Çevre ve İş Sağlığı (f = 983), Bulaşıcı Hastalıklar (f = 822) ile Veterinerlik Bilimleri (f = 497) alanlarında olduğu belirlenmiştir.



Şekil 2. Makalelerin araştırma alanı

Anahtar Kelime Ağ Analizi

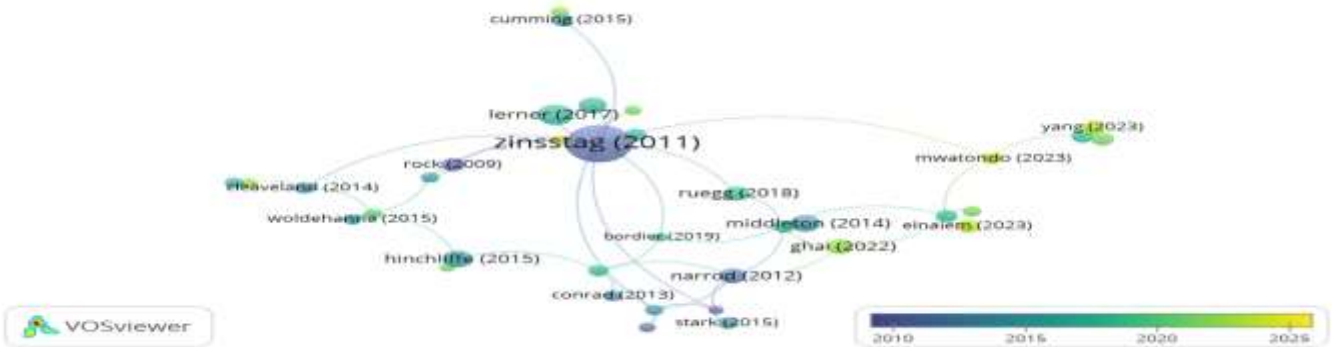
Tek Sağlık konulu çalışmalarda öne çıkan anahtar kelimelere ilişkin ağ haritası analizi bulguları Şekil 3'te sunulmaktadır. Analiz kapsamında, tekrar sayısı 10 ve üzeri olan anahtar kelimeler dikkate alınmıştır. Toplam 2223 anahtar kelime arasında 40 ağ bağlantısı olduğu tespit edilmiştir (Şekil 4). Bu alandaki eserlerde en çok kullanılan anahtar kelimelerin sırasıyla; "one health" (f=503), "antimicrobial resistance" (f = 68), ve "public health" (f=65) olduğu belirlenmiştir.



Şekil 4. Anahtar kelime ağ haritası

Eser Atıf Analizi

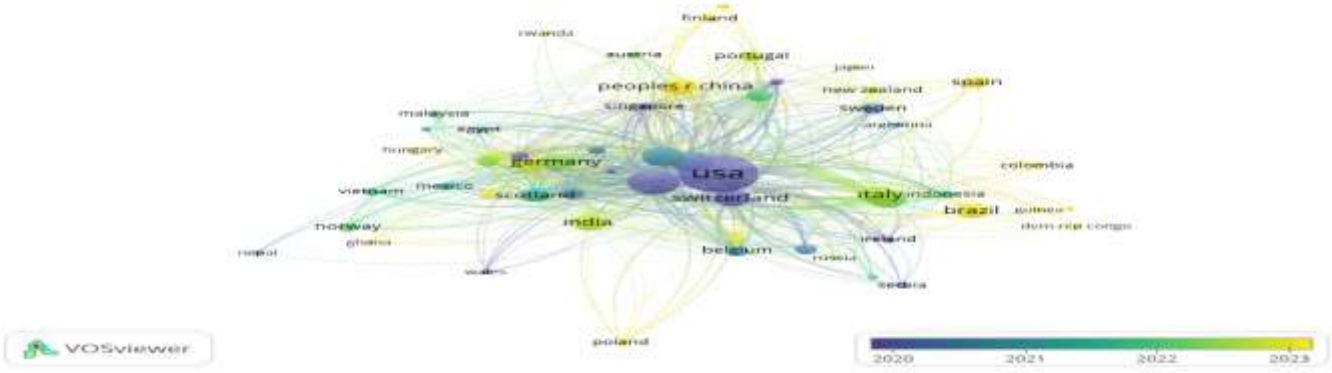
Zinsstag et al. (2011) çalışması ağın merkezinde yer almakta ve çok sayıda çalışma ile bağlantı göstermektedir. Bu durum, söz konusu çalışmanın Tek Sağlık literatüründe temel referanslardan biri olduğunu göstermektedir. Ayrıca Johnson et al. (2018), (Rüegg et al., 2018), Narrod et al. (2012) ve Middleton et al. (2014) gibi çalışmalar farklı araştırma kümeleri arasında bağlantı sağlayan köprü çalışmalar olarak dikkat çekmektedir.



Şekil 5. Eser atıf ağ analizi

Ülke Bazlı Atıf Analizi

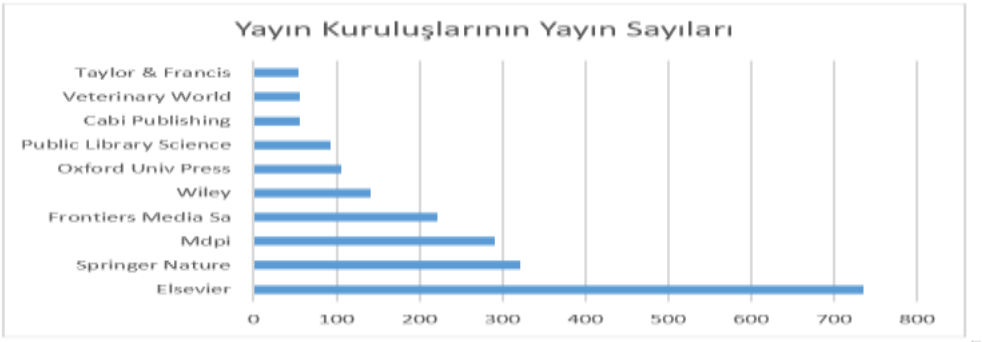
Şekil 6'da ülkeler arasındaki atıf ağına ilişkin örtüşüm görselleştirmesi sunulmaktadır. Ağ yapısı incelendiğinde Amerika Birleşik Devletleri'nin (USA) en büyük düğüme sahip olduğu ve ağın merkezinde yer aldığı görülmektedir. ABD'nin İsviçre, Almanya, İtalya, Çin, Hindistan ve İskoçya başta olmak üzere çok sayıda ülke ile yoğun bağlantılara sahip olması, uluslararası atıf ağında merkezi bir konumda bulunduğunu göstermektedir.



Şekil 6. Ülke ağ haritası

En Çok Esere Sahip Yayıncı Kuruluşlar

Tek Sağlık alanındaki makalelerin 231 farklı yayıncı kuruluş tarafından yayımlandığı belirlenmiştir. En fazla araştırma makalesi yayımlayan yayıncı kuruluşlar sırasıyla; Elsevier (f = 735), Springer Nature (f = 321) ve MDPI (f = 290) olmuştur (Şekil 6).



Şekil 6. En çok eser yayına sahip yayıncı kuruluşlar

Araştırmalara Fon Sağlayan Kuruluşlar

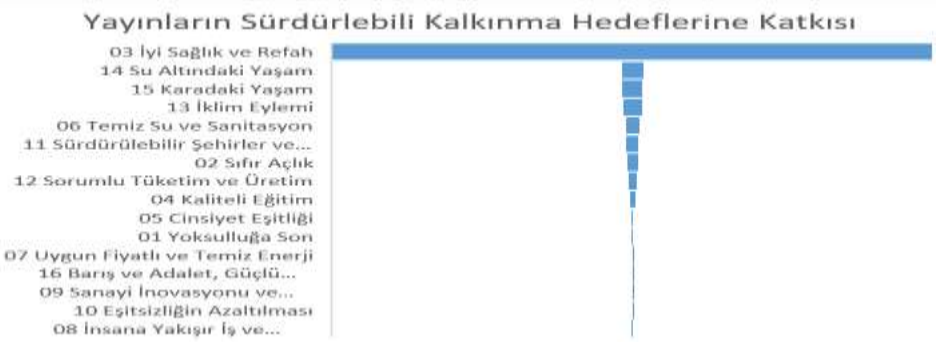
Tek Sağlık konusunda yayınlanan araştırma makalelerine fon desteği sağlayan toplam 2599 kuruluş olduğu tespit edilmiştir. Bu kuruluşlar arasında en çok fon desteği sağlayan ilk 10 kuruluşa ait destek sayıları Tablo 1'de gösterilmiştir.

Tablo 1. Finansman sağlayan kurumlar

Sıra	**Fon Sağlayan Kuruluşlar**	**Makale Sayısı**
1.	Birleşik Krallık Araştırma ve İnovasyon Kurumu (Birleşik Krallık)	91
2.	Avrupa Birliği (Çok uluslu)	90
3.	Amerika Birleşik Devletleri Sağlık ve İnsan Hizmetleri Bakanlığı (ABD)	82
4.	Çin Ulusal Doğa Bilimleri Vakfı (Çin)	74
5.	ABD Ulusal Sağlık Enstitüleri (ABD)	61
6.	Wellcome Vakfı (Birleşik Krallık)	50
7.	Ulusal Bilimsel ve Teknolojik Gelişim Konseyi (Brezilya)	49
8.	Bilim ve Teknoloji Vakfı (Portekiz)	47
9.	Tıbbi Araştırma Konseyi (Birleşik Krallık)	41
10.	Yükseköğretim Personelini Geliştirme Koordinasyonu (Brezilya)	40

Eserlerin Sürdürülebilir Kalkınma Hedeflerine Katkısı

Tek Sağlık konusunda yayınlanan eserlerin toplamda 16 farklı sürdürülebilir kalkınma hedefine katkı sağladığı belirlenmiştir (Şekil 8). Bu alanda yayın yapılan eserlerin en çok katkı sağladığı sürdürülebilir kalkınma hedefleri, İyi Sağlık ve Refah (f = 2510), 14 Su Altındaki Yaşam (f = 96) ve 15 Karadaki Yaşam (f = 90).



Şekil 8. Sürdürülebilir Kalkınma Hedefleri (WoS, 2026)

IV. TARTIŞMA

Bu çalışmada, Web of Science veri tabanında Tek Sağlık alanında yayımlanan bilimsel çalışmalar VOSviewer programı kullanılarak bibliyometrik yöntemlerle analiz edilmiş ve literatürün gelişim eğilimleri, tematik yapısı ve iş birliği ağları değerlendirilmiştir.

Bu çalışmada, Tek Sağlık alanındaki ilk yayının 1981 yılında yayımlandığı (Ashton, 1981) ve özellikle 2010'lu yıllardan itibaren yayın sayısında belirgin bir artış yaşandığı görülmüştür. Yayın üretiminin 2025 yılında en yüksek düzeye ulaşması, Tek Sağlık yaklaşımına yönelik bilimsel ilginin son yıllarda hızla arttığını göstermektedir. Bu eğilim, küresel salgınlar, zoonotik hastalıklar ve çevresel sağlık sorunlarının giderek daha fazla önem kazanmasıyla Tek Sağlık yaklaşımının araştırma gündeminde öncelikli bir alan haline geldiğini düşündürmektedir. Benzer şekilde, Aguilar-Elena et al. (2026) bibliyometrik analizinde araştırma üretiminin uzun yıllar sınırlı düzeyde seyrettiğini, ancak özellikle 1990'lardan sonra hız kazandığını ve 2010-2025 döneminde en yüksek yayın düzeyine ulaştığını bildirmiştir. Bu paralellik, Tek Sağlık ile ilişkili araştırma alanlarının zaman içinde bilimsel olarak olgunlaştığını ve küresel sağlık öncelikleri doğrultusunda giderek daha fazla ilgi gördüğünü göstermektedir.

Bu çalışmada, Tek Sağlık araştırmalarının 110 farklı bilim alanına yayıldığı ve yayınların ağırlıklı olarak Halk Sağlığı, Çevre ve İş Sağlığı, Bulaşıcı Hastalıklar ile Veterinerlik Bilimleri alanlarında yoğunlaştığı belirlenmiştir. Bu dağılım, Tek Sağlık yaklaşımının insan, hayvan ve çevre sağlığını bütüncül biçimde ele alan disiplinler arası yapısını yansıtmaktadır. Elde edilen bulgular, Qiang et al. (2022) tarafından gerçekleştirilen bibliyometrik analizle büyük ölçüde örtüşmektedir. Qiang et al. (2022), Tek Sağlık yayınlarının en yaygın olarak halk sağlığı, çevre ve iş sağlığı, bulaşıcı hastalıklar ve veterinerlik alanlarında toplandığını ve yayınların önemli bir bölümünün birden fazla bilim alanına atandığını bildirmiştir. Her iki çalışmada da özellikle halk sağlığı ve bulaşıcı hastalıkların ön plana çıkması, zoonotik hastalıklar, antimikrobiyal direnç ve küresel sağlık güvenliği gibi sorunların Tek Sağlık araştırmalarının temel odak noktalarını oluşturduğunu göstermektedir. Bulgular, Tek Sağlık araştırmalarının başlangıçta kavramsal çerçeve oluşturma ve insan-hayvan etkileşimine odaklanan çalışmalardan, günümüzde daha spesifik halk sağlığı sorunlarına ve ortaya çıkan zoonotik risklere yönelik uygulamalı araştırmalara doğru evrildiğini göstermektedir. Özellikle COVID-19 pandemisi sonrası dönemde Tek Sağlık yaklaşımı, salgın hazırlığı, zoonotik hastalıkların önlenmesi ve sürdürülebilir sağlık politikalarının geliştirilmesi bağlamında daha fazla önem kazanmıştır.

Bu çalışmada gerçekleştirilen anahtar kelime ağ analizi, Tek Sağlık literatüründe en sık kullanılan kavramların "one health", "antimicrobial resistance" ve "public health" olduğunu göstermiştir. Özellikle antimikrobiyal direnç ve halk sağlığının öne çıkması, Tek Sağlık araştırmalarının zoonotik hastalıkların kontrolü, sağlık güvenliği ve dirençli patojenlerle mücadele gibi küresel önceliklere odaklandığını göstermektedir. Benzer şekilde, Maulana et al. (2023) tarafından VOSviewer kullanılarak gerçekleştirilen anahtar kelime eş-oluşum analizinde de "one health", "zoonoses", "antimicrobial resistance", "COVID-19", "surveillance" ve "public health" kavramlarının en sık kullanılan anahtar kelimeler arasında yer

aldığı bildirilmiştir. Her iki çalışmada ortak olarak belirlenen “one health”, “antimicrobial resistance” ve “public health” anahtar kelimeleri, Tek Sağlık literatürünün temel araştırma eksenlerinin halk sağlığı, zoonotik hastalıklar ve antimikrobiyal direnç etrafında şekillendiğini ve bu temaların alanın bilimsel gelişimini yönlendiren başlıca konular olduğunu göstermektedir.

Bu çalışmada elde edilen ülkeler arası atıf ağı bulguları, Tek Sağlık literatürünün güçlü bir uluslararası bilgi paylaşımı ve iş birliği yapısına sahip olduğunu göstermektedir. Özellikle Amerika Birleşik Devletleri’nin ağın merkezinde yer alması ve çok sayıda ülkeyle yoğun atıf ilişkileri kurması, bu ülkenin Tek Sağlık araştırmalarının yönlendirilmesinde ve bilginin yayılmasında önemli bir rol üstlendiğini düşündürmektedir. İsviçre, Almanya, Çin, İtalya ve Hindistan’ın da güçlü bağlantılarla öne çıkması, Tek Sağlık yaklaşımının farklı kıtalarda etkin biçimde benimsendiğini ve küresel ölçekte disiplinler arası bir araştırma ağı oluşturduğunu göstermektedir. Ayrıca, İtalya, Brezilya, İspanya, Portekiz ve Finlandiya gibi ülkelerin son yıllarda ağda daha görünür hâle gelmesi, Tek Sağlık araştırmalarının coğrafi olarak genişlediğine ve yeni araştırma merkezlerinin uluslararası literatürde daha etkin rol almaya başladığına işaret etmektedir. Bu bulgular, zoonotik hastalıklar, antimikrobiyal direnç ve çevresel sağlık tehditleri gibi sınır aşan sorunların çözümünde uluslararası bilimsel iş birliğinin Tek Sağlık yaklaşımının temel unsurlarından biri olduğunu desteklemektedir.

Bu çalışmada, Tek Sağlık alanındaki yayınların önemli bir bölümünün Elsevier, Springer Nature ve MDPI gibi uluslararası yayıncı kuruluşlarda yayımlandığı belirlenmiştir. Bu bulgu, Tek Sağlık araştırmalarının uluslararası düzeyde faaliyet gösteren yayıncı kuruluşlar bünyesindeki dergilerde yoğunlaştığını ve alanın disiplinler arası niteliği nedeniyle farklı bilim alanlarına hitap eden yayın platformlarında yer bulduğunu göstermektedir. Özellikle halk sağlığı, çevre bilimleri, veterinerlik ve enfeksiyon hastalıkları gibi Tek Sağlık ile ilişkili alanlarda geniş dergi portföyüne sahip bu yayıncı kuruluşların öne çıkması, Tek Sağlık yaklaşımının farklı disiplinleri bir araya getiren araştırmalar için önemli bir yayın ortamı sunduğunu düşündürmektedir.

Bu çalışmada, Tek Sağlık alanındaki araştırmaların çok sayıda ulusal ve uluslararası kuruluş tarafından desteklendiği, özellikle Birleşik Krallık Araştırma ve İnovasyon Kurumu, Avrupa Birliği, ABD Sağlık ve İnsan Hizmetleri Bakanlığı ve Çin Ulusal Doğa Bilimleri Vakfı’nın öne çıktığı belirlenmiştir. Bu bulgu, Tek Sağlık araştırmalarının büyük ölçüde güçlü araştırma fonlama mekanizmalarına sahip ülkeler ve uluslararası kuruluşlar tarafından desteklendiğini göstermektedir. Farklı ülkelerden kamu kurumları, araştırma konseyleri ve uluslararası fon sağlayıcılarının ilk sıralarda yer alması, Tek Sağlık yaklaşımının küresel sağlık, zoonotik hastalıklar ve çevresel sürdürülebilirlik gibi sınır aşan sorunlarla ilişkili olması nedeniyle uluslararası araştırma gündeminde önemli bir öncelik haline geldiğini düşündürmektedir.

SONUÇLAR

Bu bibliyometrik çalışma, Tek Sağlık alanındaki bilimsel üretimin zaman içinde belirgin biçimde arttığını ve özellikle 2010’lu yıllardan itibaren hızlanan yayın eğiliminin 2025 yılında en yüksek düzeye ulaştığını göstermiştir. Bulgular, Tek Sağlık araştırmalarının ağırlıklı olarak halk sağlığı, bulaşıcı hastalıklar ve veterinerlik bilimleri ekseninde yoğunlaştığını; anahtar kelime analizinde ise “one health”, “antimicrobial resistance” ve “public health” kavramlarının alanın temel araştırma odaklarını oluşturduğunu ortaya koymuştur. Ayrıca yayınların önde gelen uluslararası yayıncı kuruluşlarda ve güçlü araştırma fonu sağlayıcıları tarafından desteklenmesi, Tek Sağlık yaklaşımının küresel ölçekte stratejik ve disiplinler arası bir araştırma alanı haline geldiğini göstermektedir.

Bu çalışma, Tek Sağlık literatürünün son yıllarda hızla büyüyen, disiplinler arası ve uluslararası nitelik taşıyan bir araştırma alanına dönüştüğünü göstermektedir. Araştırmaların özellikle halk sağlığı, bulaşıcı hastalıklar ve veterinerlik bilimlerinde yoğunlaşması, Tek Sağlık yaklaşımının sağlık yönetimi ve sağlık politikaları açısından stratejik bir önem kazandığını ortaya koymaktadır. Bu doğrultuda, insan, hayvan ve çevre sağlığını birlikte ele alan entegre yönetim modellerinin geliştirilmesi, kurumlar arası veri paylaşımı ve koordinasyon mekanizmalarının güçlendirilmesi ile Tek Sağlık yaklaşımının ulusal sağlık politikalarına daha etkin biçimde entegre edilmesi önerilmektedir.

Bu çalışma, Web of Science veritabanı ve VOSviewer programı kullanılarak gerçekleştirilen güncel bibliyometrik analiziyle Türkçe literatüre katkı sunmakta ve Tek Sağlık araştırmalarının gelişim dinamiklerini kapsamlı biçimde ortaya koymaktadır. Elde edilen bulguların, araştırmacılar ve politika

yapıcılar için gelecekteki araştırma önceliklerinin belirlenmesine, disiplinler arası iş birliklerinin güçlendirilmesine ve daha etkili Tek Sağlık politikalarının geliştirilmesine katkı sağlayacağı düşünülmektedir. Bununla birlikte, çalışmanın yalnızca Web of Science veritabanındaki yayınlarla sınırlı olması önemli bir kısıttır. Gelecekte yapılacak araştırmalarda Scopus, PubMed ve diğer uluslararası veri tabanlarının da analize dahil edilmesi, Tek Sağlık literatürünün daha kapsamlı biçimde değerlendirilmesine olanak sağlayacaktır.

Bu çalışmanın önemli güçlü yönlerinden biri, Tek Sağlık alanındaki bilimsel üretimi uzun bir zaman dilimini kapsayacak şekilde değerlendirmesidir. Web of Science veritabanından elde edilen güncel verilerin VOSviewer programı kullanılarak analiz edilmesi, yayın eğilimlerinin, tematik yapıların ve uluslararası iş birliği ağlarının bütüncül biçimde ortaya konulmasına olanak sağlamıştır. Ayrıca, Türkçe literatürde Tek Sağlık alanına ilişkin güncel bibliyometrik değerlendirmelerin sınırlı olması nedeniyle çalışma, alandaki mevcut bilgi birikimini güncelleyen ve tamamlayan bir katkı sunmaktadır.

Bununla birlikte, çalışmanın bazı sınırlılıkları bulunmaktadır. Araştırma yalnızca Web of Science veri tabanında indekslenen yayınlarla sınırlandırılmış olup Scopus, PubMed, Dimensions veya diğer uluslararası veri tabanlarında yer alan çalışmalar analize dahil edilmemiştir. Bu durum, özellikle bölgesel yayınlar ve farklı disiplinlerde yayımlanan bazı çalışmaların kapsam dışında kalmasına neden olmuş olabilir. Ayrıca bibliyometrik analizler, yayın ve atıf verilerine dayalı nicel değerlendirmeler sunduğundan, çalışmaların metodolojik kalitesi veya bilimsel içeriklerinin derinlemesine değerlendirilmesine olanak tanımamaktadır. Bu nedenle elde edilen bulgular, Tek Sağlık literatüründeki eğilimleri ve yapısal özellikleri yansıtmakta olup alanın tüm bilimsel çıktılarının eksiksiz bir temsilini sağlamamaktadır.

KAYNAKLAR

Aguilar-Elena, R., Rodríguez-Escolar, I., Collado-Cuadrado, M., González-Mohino, E. I., Balmori-de la Puente, A., Gil-Abad, A., & Morchón, R. (2026). Mapping Global Trends in Dirofilaria immitis Research Within the One Health Framework (1945–2025): A Bibliometric Perspective. *Animals*, *16*(6), 988. https://doi.org/10.3390/ani16060988

Ashton, J. R. (1981). Characteristics of special groups of abortion patients from one health district. *Journal of Biosocial Science*, *13*(1), 63–69. https://doi.org/10.1017/S0021932000013195

Cascajares, M., Alcayde, A., Salmerón-Manzano, E., & Manzano-Agugliaro, F. (2021). The Bibliometric Literature on Scopus and WoS: The Medicine and Environmental Sciences Categories as Case of Study. *International Journal of Environmental Research and Public Health*, *18*(11), 5851. https://doi.org/10.3390/ijerph18115851

Johnson, I., Hansen, A., & Bi, P. (2018). The challenges of implementing an integrated One Health surveillance system in Australia. *Zoonoses and Public Health*, *65*(1). https://doi.org/10.1111/zph.12433

Khan, M. S., Rothman-Ostrow, P., Spencer, J., Hasan, N., Sabirovic, M., Rahman-Shepherd, A., Shaikh, N., Heymann, D. L., & Dar, O. (2018). The growth and strategic functioning of One Health networks: a systematic analysis. *The Lancet Planetary Health*, *2*(6), e264–e273. https://doi.org/10.1016/S2542-5196(18)30084-6

Maulana, F. K., Maghfiroh, L., Sari, S. S. N., & Rantam, F. A. (2023). A bibliometric analysis of worldwide research on One Health from 2012 to 2022. *International Journal of One Health*, *9*(2), 159–165. https://doi.org/10.14202/IJOH.2023.159-165

Miao, L., Li, H., Ding, W., Lu, S., Pan, S., Guo, X., Zhou, X., & Wang, D. (2022). Research Priorities on One Health: A Bibliometric Analysis. *Frontiers in Public Health*, *10*(889854), 1–12. https://doi.org/10.3389/fpubh.2022.889854

Middleton, D., Pallister, J., Klein, R., Feng, Y.-R., Haining, J., Arkinstall, R., Frazer, L., Huang, J.-A., Edwards, N., Wareing, M., Elhay, M., Hashmi, Z., Bingham, J., Yamada, M., Johnson, D., White, J., Foord, A., Heine, H. G., Marsh, G. A., … Wang, L.-F. (2014). Hendra Virus Vaccine, a One Health Approach to Protecting Horse, Human, and Environmental Health. *Emerging Infectious Diseases*, *20*(3). https://doi.org/10.3201/eid2003.131159

Milazzo, A., Liu, J., Multani, P., Steele, S., Hoon, E., & Chaber, A.-L. (2025). One Health implementation: A systematic scoping review using the Quadripartite One Health Joint Plan of Action. *One Health*, *20*, 101008. https://doi.org/10.1016/j.onehlt.2025.101008

Mishra, M., Desul, S., Santos, C. A. G., Mishra, S. K., Kamal, A. H. M., Goswami, S., Kalumba, A. M., Biswal, R., da Silva, R. M., dos Santos, C. A. C., & Baral, K. (2023). A bibliometric analysis of sustainable development goals (SDGs): a review of progress, challenges, and opportunities. *Environment, Development and Sustainability*, *26*(5), 11101–11143. https://doi.org/10.1007/s10668-023-03225-w

Mwatondo, A., Rahman-Shepherd, A., Hollmann, L., Chiossi, S., Maina, J., Kurup, K. K., Hassan, O. A., Coates, B., Khan, M., Spencer, J., Mutono, N., Thumbi, S. M., Muturi, M., Mutunga, M., Arruda, L. B., Akhbari, M., Ettehad, D., Ntoumi, F., Scott, T. P., … Dar, O. (2023). A global analysis of One Health Networks and the proliferation of One Health collaborations. *The Lancet*, *401*(10376), 605–616. https://doi.org/10.1016/S0140-6736(22)01596-3

Narrod, C., Zinsstag, J., & Tiongco, M. (2012). A One Health Framework for Estimating the Economic Costs of Zoonotic Diseases on Society. *EcoHealth*, *9*(2), 150–162. https://doi.org/10.1007/s10393-012-0747-9

Özen Çınar, İ., & Özkaya Bozkurt, E. (2026). General, Conceptual, and Social Structure of Environmental Health Literacy Studies: Bibliometric Analysis. *Public Health Nursing*, *43*(2), 367–379. https://doi.org/10.1111/phn.70056

Qiang, N., Gu, S.-Y., Wang, X.-Y., Zhang, X.-X., Xia, S., Zheng, J.-X., Gong, W.-F., Bergquist, R., Ran, J.-J., & Han, L.-F. (2022). A One Health information database based on standard bibliometric analysis. *Science in One Health*, *1*(2022), 1–8. https://doi.org/10.1016/j.soh.2023.100012

Raphael, R., AS, D., Karnaver, P., Devy, R. A., M, D. P., & Jose, J. (2024). Mapping Research Trends in One Health: A Comprehensive Review. *Cureus*, *16*(9), e70047. https://doi.org/10.7759/cureus.70047

Rüegg, S. R., Nielsen, L. R., Buttigieg, S. C., Santa, M., Aragrande, M., Canali, M., Ehlinger, T., Chantziaras, I., Boriani, E., Radeski, M., Bruce, M., Queenan, K., & Häsler, B. (2018). A Systems Approach to Evaluate One Health Initiatives. *Frontiers in Veterinary Science*, *5*. https://doi.org/10.3389/fvets.2018.00023

Şimşir, İ., & Mete, B. (2022). Examination of One Health Studies with Science Mapping Technique. *Online Türk Sağlık Bilimleri Dergisi*, *7*(3), 425–431. https://doi.org/10.26453/otjhs.1063769

Sweileh, W. M. (2020). Bibliometric analysis of peer-reviewed literature on climate change and human health with an emphasis on infectious diseases. *Globalization and Health*, *16*(44), 1–17. https://doi.org/10.1186/s12992-020-00576-1

U.S. Centers for Disease Control and Prevention. (2025, June 27). *About One Health*. Https://Www.Cdc.Gov/One-Health/about/Index.Html. (Erişim Tarihi: 07.08.2026).

Web of Science. (2026). *Web of Science Core Collection*. Https://Www.Webofscience.Com/Wos/Woscc/Summary/00c9c8e2-812f-49e0-9c32-73f760eeb629-01be577edf/Relevance/1. Erişim Tarihi: 12.07.2026.

WHO. (2022, May 29). *Quadripartite Memorandum of Understanding (MoU) signed for a new era of One Health collaboration*. Https://Www.Who.Int/News/Item/29-04-2022-Quadripartite-Memorandum-of-Understanding-(Mou)-Signed-for-a-New-Era-of-One-Health-Collaboration. (Erişim Tarihi:07.07.2026).

Yalçın, S., & Yalçın, S. S. (2026). Mapping One Health and Sustainability from 2007 to 2024: Multi-Period Evolution with Bibliometric and Content Insights from Türkiye's Oldest Veterinary Journal. *Publications*, *14*(1), 12. https://doi.org/10.3390/publications14010012

Zhao, L., Zhao, Y., Du, J., Desloge, A., Hu, Z., & Cao, G. (2022). Mapping the Research on Health Policy and Services in the Last Decade (2009–2018): A Bibliometric Analysis. *Frontiers in Public Health*, *10*(2022), 1–12. https://doi.org/10.3389/fpubh.2022.773668

Zinsstag, J., Schelling, E., Waltner-Toews, D., & Tanner, M. (2011). From "one medicine" to "one health" and systemic approaches to health and well-being. *Preventive Veterinary Medicine*, *101*(3–4), 148–156. https://doi.org/10.1016/j.prevetmed.2010.07.003

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Goal-Conditioned Point Navigation of a Differential-Drive Robot Using Deep Deterministic Policy Gradient in MuJoCo

Ufuk Demircioğlu [1*], Halit Bakır[2]

[1] *Department of Mechanical Engineering, Faculty of Engineering and Natural Sciences, Sivas University of Science and Technology, Sivas, Türkiye*

[2]*Department of Computer Engineering, Faculty of Engineering and Natural Sciences, Sivas University of Science and Technology, Sivas, Türkiye*

[*]*(udemircioglu@sivas.edu.tr) Email of the corresponding author*

Abstract – This study presents a Deep Deterministic Policy Gradient (DDPG) controller for continuous point-goal navigation of a differential-drive robot in MuJoCo. Training employed actor–critic networks, slowly updated target networks, uniform experience replay, randomized goal locations, and a distance-based reward with a terminal success bonus. The recorded run comprised approximately 1,100 episodes and exhibited substantial return variability. However, episodes with returns approaching the 10,000-point terminal bonus became more frequent after approximately 850–900 episodes, and the 50-episode moving average increased from negative values to approximately 3,000 near the end of training. Post-training evaluation included eight point-goal trials covering seven distinct target coordinates. The resulting trajectories were smooth and goal-dependent, while repeated trials for the same target produced nearly coincident paths. In a separate waypoint-sequence test, the same policy was used to execute a multi-stage route that broadly followed the prescribed waypoint geometry while smoothing abrupt changes in direction. These findings provide qualitative evidence that the learned policy responds to goal-relative information and can be composed for sequential waypoint navigation. Because the evidence is based on a single training run and qualitative trajectory plots, a multi-seed evaluation reporting success rate, terminal error, cross-track error, path length, and control effort is required to establish robustness and comparative performance.

Keywords – deep reinforcement learning, differential-drive robot, mobile robot navigation, continuous control

I. INTRODUCTION

Reinforcement learning (RL) provides a complementary formulation in which a control policy is optimized through repeated interaction with an environment and reward-based feedback [1]. Deep Deterministic Policy Gradient (DDPG) extends deterministic policy gradients with nonlinear function approximation and combines an actor, a critic, slowly updated target networks, and experience replay for continuous-action learning [2]. These characteristics make DDPG a natural candidate for direct wheel-level control. Nevertheless, off-policy actor–critic learning remains sensitive to function-approximation error, the replay distribution, reward scaling, and exploration. Twin Delayed DDPG (TD3) subsequently demonstrated that critic error can lead to value overestimation and unstable policy updates, motivating clipped double critics, delayed actor updates, and target-policy smoothing [3]. A precise account of the implemented baseline is therefore necessary before performance can be attributed to a particular design choice. This study considers an obstacle-free point-goal task for a simulated differential-drive robot. At each control step, the policy receives a compact state vector and produces two bounded motor commands. The target changes between episodes, encouraging a goal-conditioned control law rather than memorization of a single destination. MuJoCo provides rigid-body dynamics and contact simulation [4], whereas Gymnasium provides the environment interface and explicit termination semantics [5].

A. Literature Review

Recent mobile-robot studies have extended off-policy actor–critic methods through modifications to replay, exploration, state representation, and reward design. Zhang *et al.* combined DDPG with improved state processing and planning mechanisms for mapless navigation among static and dynamic obstacles [6]. Xue *et al.* introduced bidirectional obstacle-avoidance enhancement and cross-attention to emphasize informative state–action relationships in unknown dynamic environments [7]. Wei *et al.* partitioned experience into multiple pools and incorporated guided rewards for navigation in unknown environments [8]. Deshpande *et al.* coupled DDPG exploitation with differential-gaming exploration, illustrating the importance of the behavior policy used to populate replay memory [9]. Studies published in 2025 continued this emphasis on auxiliary mechanisms. Xue *et al.* improved decision evaluation for path planning in high-density environments [10], while Zhang *et al.* incorporated self-attention, adversarial training, and prioritized experience replay into DDPG-based mobile-robot planning [11]. Nasti *et al.* adopted TD3 and combined latent representation learning, prediction-based intrinsic motivation, and regularization for adaptive mapless navigation [12]. Collectively, these studies indicate that navigation performance often depends on several interacting mechanisms beyond the nominal actor–critic update. Recent work has also investigated learning-based control as a complement to, rather than a replacement for, classical control. Demircioğlu *et al.* integrated Pure Pursuit reference generation and an inner-loop PID controller with a residual TD3 policy for differential-drive target tracking under chassis- and wheel-mass uncertainty. Their analysis additionally examined replay capacity, minibatch size, optimizer selection, and network architecture [13]. In a related study, a TD3 policy constructed from proportional, integral, and derivative error components provided adaptive PID-like control of an uncertain mass–spring–damper system [14]. These approaches derive interpretability and robustness from explicit control structure, whereas the present study evaluates a simpler end-to-end point-goal policy. Although these extensions can improve navigation performance, their simultaneous use complicates attribution of measured gains to individual design choices [8], [9], [11]. A compact physics-based baseline is therefore valuable for controlled analysis [15]. The present study intentionally omits obstacle perception and addresses the narrower question of whether DDPG can map low-dimensional, goal-relative state information directly to continuous wheel commands

B. Motivation

The motivation of this study is to investigate whether a standard DDPG algorithm can learn continuous point-goal navigation for a differential-drive robot using a compact state representation. Continuous wheel commands avoid the limitations of discrete action spaces, while randomized target locations encourage the policy to learn goal-directed behavior instead of memorizing a single trajectory. MuJoCo provides a realistic physics-based simulation environment, making it suitable for evaluating reinforcement learning algorithms under continuous control.

C. Novelty and Contributions

The proposed method does not modify the original DDPG algorithm. Instead, the contribution of this work is its application to goal-conditioned navigation of a differential-drive robot in MuJoCo. The main contributions are:

1. a compact nine-dimensional goal-conditioned observation for continuous control;
2. a complete DDPG implementation including the environment, reward function, actor–critic networks, and training procedure; and
3. qualitative evaluation through training rewards, point-goal navigation, and waypoint-following experiments.

II. MATERIALS AND METHOD

D. Implementation and Simulation

The experiment was implemented in Python using PyTorch for function approximation and gradient-based optimization, NumPy for numerical operations, Gymnasium for the reinforcement-learning interface, and MuJoCo for rigid-body simulation. The custom WheelRobotEnv class extends Gymnasium's MujocoEnv and returns the observation, reward, termination flag, truncation flag, and an information dictionary at each environment step. The training program assigns the neural networks and minibatch tensors to a CUDA device when one is available and otherwise uses the CPU. The environment was instantiated with human rendering enabled and consequently renders after each control step. This configuration supports qualitative inspection but adds computational overhead. Exact software versions, hardware specifications, execution time, and pseudorandom seeds were not recorded in the training program; therefore, the available artifacts permit reconstruction of the algorithmic workflow but not exact numerical replication of the reported run.

E. Robot Model and Point-Goal Task

The MuJoCo scene contains a planar ground surface, a free-moving chassis, two actuated wheels, a tracked spherical body located centrally within the chassis assembly, and a non-colliding spherical target. Figure 1 shows the simulated platform and identifies the two independently actuated wheels and the target object. The yellow rectangular body represents the chassis, the green cylinders represent the left and right wheels, and the red sphere is the tracked body whose center-of-mass position is used in the observation, reward, and termination calculations. The black sphere represents the goal and is excluded from contact interactions. Table 1 summarizes the geometric, inertial, actuator, and task parameters defined by the XML model and environment code. MuJoCo specifies box sizes as half-lengths along the body-fixed axes and cylinder sizes as radius and half-length; the corresponding full dimensions are therefore reported explicitly to avoid ambiguity.

.

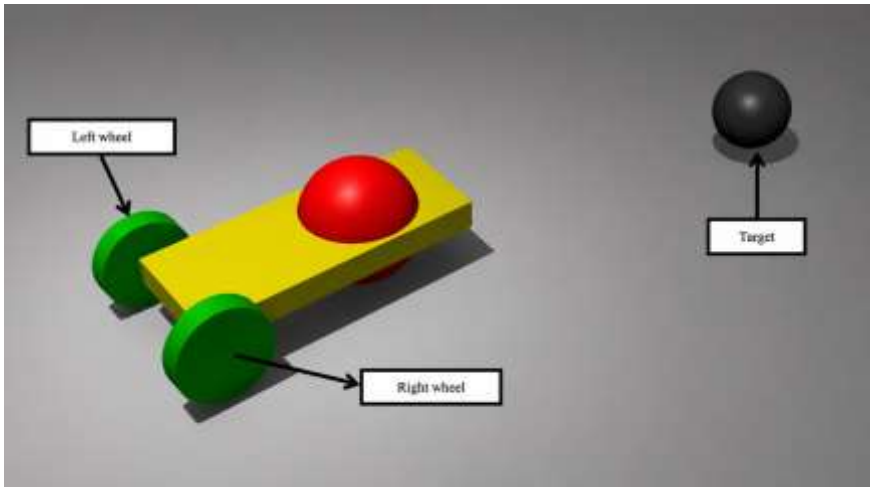


Fig. 1 Differential-drive robot model in the MuJoCo environment.

The left and right wheel hinges are driven by independent MuJoCo motors, and their bounded control inputs are supplied directly by the policy. Each action is maintained for five internal simulation steps through frame_skip=5. At the beginning of an episode, the parent environment restores the model-defined initial robot state. The custom reset routine then samples the horizontal target coordinates independently according to

$$x_g \sim \mathcal{U}(0.5,4.5), \qquad y_g \sim \mathcal{U}(0.5,4.5),$$

sets $z_g = 0.2\,\mathrm{m}$, and calls mj_forward to propagate the modified target position through the MuJoCo state. Thus, the robot begins from a common model-defined state, whereas the goal varies across episodes within the positive x–y quadrant.

Table 1. Robot, Actuator, and Simulation Parameters

Element	Parameter	Implemented value
Environment	Gravity	$(0,0,-9.81)\ \mathrm{m/s^2}$
Chassis	Geometry	Box
	XML half-sizes	$(0.05,0.20,0.50)$ m
	Full body-frame dimensions	$(0.10,0.40,1.00)$ m
	Mass	2.0 kg
	Initial chassis position	$(0,0,0.20)$ m
Tracked body	Geometry	Sphere connected by a ball joint
	Radius; mass	0.20 m; 5.0 kg
	Position relative to chassis	$(0,0,0.20)$ m
Wheels	Geometry	Cylinders
	Radius; full width	0.20 m; 0.10 m
	Mass of each wheel	0.50 kg
	Relative wheel centers	$(0,+0.30,-0.50)$ and $(0,-0.30,-0.50)$ m
	Nominal wheel separation	0.60 m
Wheel motors	Control range; gear	$[-1,1]$; 1.0
Target	Geometry; radius	Sphere; 0.20 m
	Collision properties	Disabled (contype=0, conaffinity=0)
	Reset coordinates	$x_g, y_g \in [0.5,4.5]$ m, $z_g = 0.20$ m
Simulation	Action frame skip	5 internal physics steps

F. State and Action Representations

The control problem is represented as a discounted Markov decision process $\langle \mathcal{S}, \mathcal{A}, P, r, \gamma \rangle$. Let $\dot{\mathbf{q}}_{t,1:3} \in \mathbb{R}^3$ denote the first three entries of MuJoCo's generalized-velocity vector, $\mathbf{p}_{b,t} \in \mathbb{R}^3$ the center-of-mass position of the tracked body, and $\mathbf{p}_g \in \mathbb{R}^3$ the target position. The nine-dimensional observation returned to the agent is

$$\mathbf{s}_t = [\dot{\mathbf{q}}_{t,1:3} \quad \mathbf{p}_{b,t} \quad \mathbf{p}_g - \mathbf{p}_{b,t}]^\top \in \mathbb{R}^9.$$

Table 2 details the ordering and physical interpretation of the observation components. The final three components explicitly provide the three-dimensional goal displacement, so the policy receives both the direction and distance to the active target without having to reconstruct them from two independent absolute-position inputs.

Table 2 Composition of the Nine-Dimensional Observation Vector

Component	Vector entries	Dim.	Physical interpretation
Free-base translational velocity	$\dot{\mathbf{q}}_{t,1:3}$	3	First three generalized-velocity entries associated with translational motion of the free base.
Tracked-body position	$\mathbf{p}_{b,t}$	3	Cartesian center-of-mass coordinates (x_b, y_b, z_b) of the tracked central body.
Goal-relative displacement	$\mathbf{p}_g - \mathbf{p}_{b,t}$	3	Cartesian displacement $(x_g - x_b, y_g - y_b, z_g - z_b)$ from the tracked body to the target.

The continuous action is

$$\mathbf{a}_t = [u_{L,t} \quad u_{R,t}]^\top \in [-1,1]^2,$$

where $u_{L,t}$ and $u_{R,t}$ are the left- and right-wheel motor commands. No intermediate kinematic steering law converts the policy output into actuator commands; the two actor outputs are applied directly to the MuJoCo motors.

G. Transition, Reward, and Episode Logic

After applying $\mathbf{a}_t$ for five internal physics steps, the environment obtains the next observation and computes the post-transition Euclidean distance

$$d_{t+1} = \|\mathbf{p}_{b,t+1} - \mathbf{p}_g\|_2.$$

The scalar reward associated with the transition is

$$r_t = -d_{t+1} + 10{,}000\,\mathbb{I}(d_{t+1} < 0.005),$$

where $\mathbb{I}(\cdot)$ is the indicator function. The dense component penalizes absolute distance rather than the change in distance between consecutive states. A terminal bonus of $10{,}000$ is added when the tracked body enters a 5 mm radius around the target.

The environment sets the termination flag according to

$$z_t = \mathbb{I}(d_{t+1} < 0.005 \ \vee \ d_{t+1} > 5.0).$$

Consequently, an episode ends either upon target acquisition or when the target distance exceeds 5 m. The truncation flag is always false, and no independent maximum episode length is imposed by the custom environment

H. DDPG Function Approximators and Updates

The controller follows the DDPG actor–critic formulation [2]. The actor $\mu_\theta: \mathbb{R}^9 \to [-1,1]^2$ contains two fully connected hidden layers with 400 and 300 units, respectively. Rectified linear units are used in both hidden layers, and the two-dimensional output is bounded by a hyperbolic tangent:

$$\mu_\theta(\mathbf{s}) = a_{\max}\tanh(W_3\sigma(W_2\sigma(W_1\mathbf{s} + b_1) + b_2) + b_3),$$

where $a_{\max} = 1$. The critic $Q_\phi(\mathbf{s}, \mathbf{a})$ concatenates the nine-dimensional state and two-dimensional action before its first layer. It then applies hidden layers of 400 and 300 rectified linear units and produces a single scalar estimate of the action value.

Figure 2 summarizes the implemented online actor and critic architectures. The target actor and target critic use identical network topologies and differ only in their parameter values.

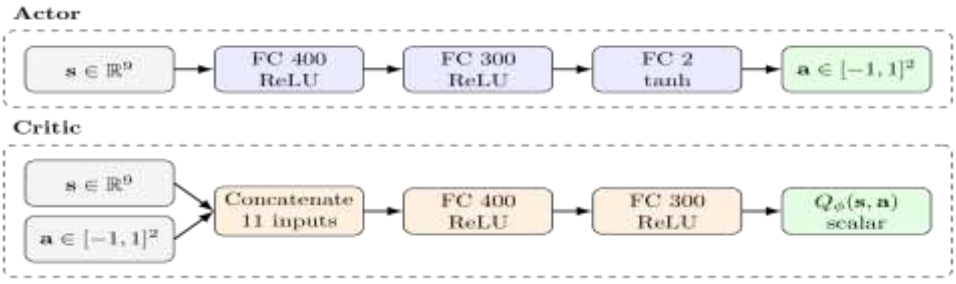


Fig. 2 Actor and critic network architectures of the DDPG agent

Separate actor and critic target networks, denoted by $\mu_{\bar{\theta}}$ and $Q_{\bar{\phi}}$, are initialized by copying the online-network parameters. For a replayed transition $(\mathbf{s}_i, \mathbf{a}_i, r_i, \mathbf{s}_{i+1}, z_i)$, the target value is

$$y_i = r_i + (1 - z_i)\gamma Q_{\bar{\phi}}(\mathbf{s}_{i+1}, \mu_{\bar{\theta}}(\mathbf{s}_{i+1})).$$

The critic parameters are optimized by minimizing the mean-squared temporal-difference error

$$\mathcal{L}_Q(\phi) = \frac{1}{B}\sum_{i=1}^{B}\big(Q_\phi(\mathbf{s}_i, \mathbf{a}_i) - y_i\big)^2,$$

whereas the actor minimizes

$$\mathcal{L}_\mu(\theta) = -\frac{1}{B}\sum_{i=1}^{B} Q_\phi\left(\mathbf{s}_i, \mu_\theta(\mathbf{s}_i)\right).$$

Adam optimization is used with learning rates of 10^{-4} and 10^{-3} for the actor and critic, respectively. Following every gradient update, both target networks are adjusted by Polyak averaging:

$$\bar{\psi} \leftarrow \tau\psi + (1-\tau)\bar{\psi}, \qquad \psi \in \{\theta, \phi\},$$

with $\tau = 0.001$. The discount factor is $\gamma = 0.99$.

i. Experience Replay and Training Procedure

Experience is stored as $(\mathbf{s}_t, \mathbf{a}_t, r_t, \mathbf{s}_{t+1}, z_t)$ in a bounded double-ended queue with a configured maximum capacity of 10^8 transitions. At each environment step, the new transition is appended to the buffer and one training call is made. Optimization is skipped until the buffer contains at least $B = 128$ transitions. Thereafter, each update uses a uniformly sampled minibatch drawn without replacement by Python's random.sample. The implementation does not use prioritized replay, multiple replay pools, reward normalization, state normalization, gradient clipping, or explicit regularization. The behavior action is the deterministic actor output $\mathbf{a}_t = \mu_\theta(\mathbf{s}_t)$, without Ornstein–Uhlenbeck, Gaussian, or parameter-space exploration noise. Thus, diversity in the collected experience arises from randomized target positions and changes in the learned network parameters rather than from an explicitly stochastic behavior policy. Training is configured for at most 5000 episodes. For episode e, the program accumulates the episodic return $G_e = \sum_t r_t$ and records both G_e and the cumulative mean $e^{-1}\sum_{j=1}^{e} G_j$. After every episode, the six most recent returns are inspected. Training terminates early if all six are strictly greater than zero. When this condition is met, the online actor and critic, both target networks, the episode index, the final six returns, and the complete reward history are saved. The program contains no unconditional checkpoint operation after the maximum episode count; therefore, checkpoint creation depends on satisfaction of the early-stopping criterion. Table 3 summarizes the principal implementation settings.

Table 3 DDPG Training Configuration

Parameter	**Value**
Maximum training episodes	5000
Observation dimension	9
Action dimension	2
Action bounds	$[-1,1]$
Actor hidden layers	400, 300
Critic hidden layers	400, 300
Actor learning rate	10^{-4}
Critic learning rate	10^{-3}
Optimizer	Adam
Discount factor γ	0.99
Target update rate τ	0.001
Minibatch size B	128
Replay capacity	10^8
Simulation frame skip	5
Goal sampling range	$x_g, y_g \in [0.5,4.5]$ m
Success distance	0.005 m
Divergence distance	5.0 m
Terminal success bonus	10,000
Exploration noise	None
Independent time limit	None
Early-stop window	6 episodes
Early-stop return threshold	0

III. RESULTS

Figure 3 shows the episodic returns, cumulative mean return, and 50-episode moving average over approximately 1,100 training episodes. Although episodic returns remained highly variable throughout training, the moving average exhibited a clear upward trend after approximately 850 episodes, reaching about 3,000 by the end of training. Episodes achieving rewards close to the 10,000 terminal bonus also became more frequent, indicating an increasing success rate and shorter trajectories. In contrast, the cumulative mean remained negative because it was influenced by the large negative returns observed during the early stages of learning. Overall, the results suggest that the DDPG agent learned an effective goal-conditioned navigation policy, although the persistent variability indicates that additional multi-seed evaluations are required to assess robustness and convergence.

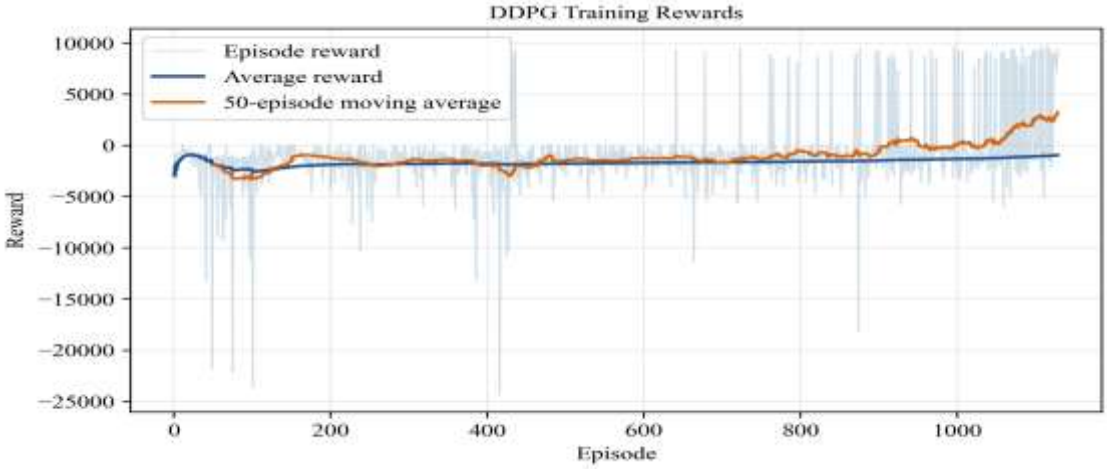


Fig. 3 Training reward history of the DDPG controller.

The trained actor was evaluated from a common initial condition using eight point-goal trials in the x–y plane. As shown in Fig. 4, the controller generated smooth, goal-directed trajectories toward all target locations. The trajectory shapes varied according to the target position, indicating that the policy effectively exploited the goal-relative state information rather than reproducing a single memorized path. The nearly identical trajectories obtained for the repeated target $(2, 3)$ (trials T_3 and T_8) demonstrate the deterministic and consistent behavior of the learned policy. Although these qualitative results indicate successful point-goal navigation, quantitative evaluation using metrics such as terminal error, path length, travel time, and control effort is required for a comprehensive performance assessment.

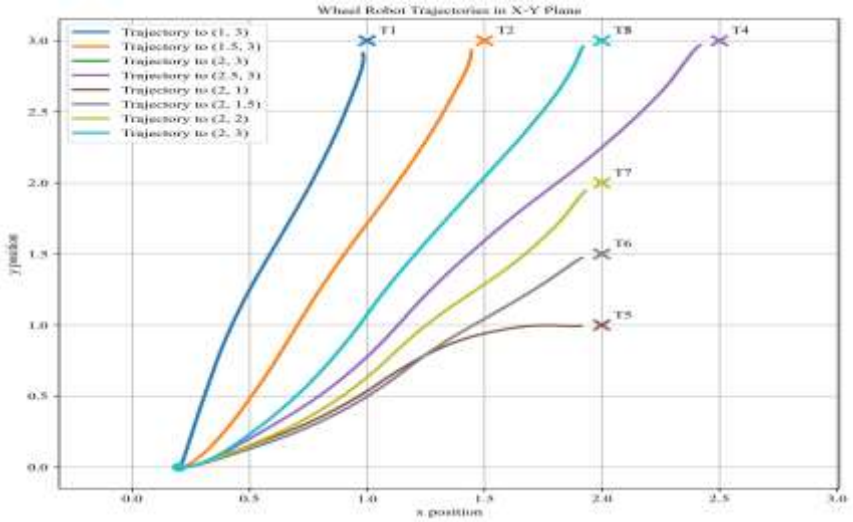


Fig. 4 Planar trajectories generated by the trained DDPG actor for eight point-goal trials.

To evaluate waypoint-based navigation, the trained actor was tested using a sequence of predefined waypoints. As shown in Fig. 5, the policy generated a smooth trajectory that closely followed the

waypoint sequence without an explicit path-tracking controller. The trajectory naturally smoothed directional changes, reflecting the continuous control policy learned by the DDPG agent. These results suggest that a point-goal policy can be effectively reused for multi-stage navigation by sequentially updating the target. However, quantitative metrics such as waypoint-reaching error, cross-track error, and completion time are required for a comprehensive evaluation.

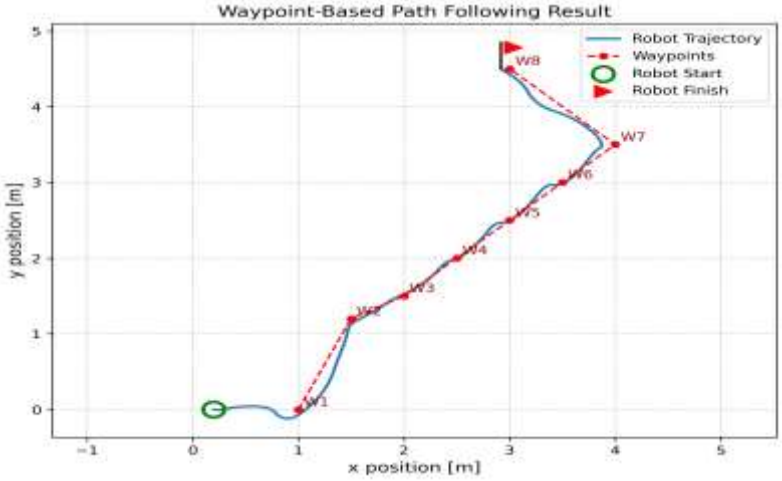


Fig. 5 Waypoint-sequence evaluation using the trained DDPG actor.

IV. DISCUSSION

The results demonstrate that a standard DDPG agent can learn a goal-conditioned navigation policy for a differential-drive robot using a compact nine-dimensional observation. The increasing training reward and the successful point-goal and waypoint evaluations indicate that the learned policy effectively maps goal-relative information to continuous wheel commands. Although the obtained trajectories were smooth and consistent, the evaluation was limited to an obstacle-free simulation environment and a single training run. Therefore, additional multi-seed experiments, quantitative performance metrics, and comparisons with more advanced algorithms such as TD3 and SAC are needed to establish the robustness and generalization capability of the proposed approach.

V. CONCLUSION

This study presented a Deep Deterministic Policy Gradient (DDPG)-based controller for continuous point-goal navigation of a differential-drive robot in the MuJoCo simulation environment. A compact nine-dimensional observation space was used to generate bounded left- and right-wheel motor commands through an end-to-end actor–critic architecture. Training with randomized goal locations enabled the policy to learn goal-conditioned behavior without relying on predefined trajectories or intermediate steering controllers. The experimental results demonstrated that the learned policy successfully generated smooth and consistent trajectories toward multiple target locations. In addition, the waypoint-sequence evaluation showed that the learned point-goal policy can be extended to multi-stage navigation by sequentially updating the target position. Although the evaluation was limited to an obstacle-free simulation environment and qualitative trajectory analysis, the results demonstrate that a standard DDPG framework provides an effective baseline for goal-conditioned navigation of differential-drive robots in a physics-based simulation environment.

REFERENCES

[1] A. H. Abdul Hafez, U. Demircioğlu, H. Bakır, and R. Alzubi, "A comprehensive review of reinforcement learning algorithms for control systems: Applications, challenges, and future directions," *IEEE Access*, vol. 14, pp. 42809-42837, 2026, doi: 10.1109/ACCESS.2026.3674539.

[2] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning," in *Proc. Int. Conf. Learning Representations (ICLR)*, 2016, doi: 10.48550/arXiv.1509.02971.

[3] S. Fujimoto, H. van Hoof, and D. Meger, "Addressing function approximation error in actor-critic methods," in *Proc. 35th Int. Conf. Machine Learning (ICML)*, vol. 80, pp. 1587-1596, 2018.

[4] E. Todorov, T. Erez, and Y. Tassa, "MuJoCo: A physics engine for model-based control," in *Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS)*, Vilamoura, Portugal, 2012, pp. 5026-5033, doi: 10.1109/IROS.2012.6386109.
[5] M. Towers, A. Kwiatkowski, J. Terry, J. U. Balis, G. D. Cola, T. Deleu, M. Goulão, A. Kallinteris, M. Krimmel, A. KG, R. Perez-Vicente, A. Pierré, S. Schulhoff, J. J. Tai, H. Tan, and O. G. Younis, "Gymnasium: A standard interface for reinforcement learning environments," *arXiv preprint*, arXiv:2407.17032, 2024, doi: 10.48550/arXiv.2407.17032.
[6] S. Zhang, W. Tang, P. Li, and F. Zha, "Mapless path planning for mobile robot based on improved deep deterministic policy gradient algorithm," *Sensors*, vol. 24, no. 17, Art. no. 5667, 2024, doi: 10.3390/s24175667.
[7] J. Xue, S. Zhang, Y. Lu, X. Yan, and Y. Zheng, "Bidirectional obstacle avoidance enhancement-deep deterministic policy gradient: A novel algorithm for mobile-robot path planning in unknown dynamic environments," *Advanced Intelligent Systems*, vol. 6, no. 4, Art. no. 2300444, 2024, doi: 10.1002/aisy.202300444.
[8] L. Wei, Q. Xu, and Z. Hu, "Mobile robot path planning based on multi-experience pool deep deterministic policy gradient in unknown environment," *International Journal of Machine Learning and Cybernetics*, vol. 15, no. 12, pp. 5823-5837, 2024, doi: 10.1007/s13042-024-02281-6.
[9] S. V. Deshpande, H. R, B. S. K. K. Ibrahim, and M. D. S. Ponnuru, "Mobile robot path planning using deep deterministic policy gradient with differential gaming (DDPG-DG) exploration," *Cognitive Robotics*, vol. 4, pp. 156-173, 2024, doi: 10.1016/j.cogr.2024.08.002.
[10] J. Xue, M. He, J. Chen, B. Dong, and Y. Zheng, "Improved DDPG based on enhancing decision evaluation for path planning in high-density environments," *Expert Systems with Applications*, vol. 279, Art. no. 127378, 2025, doi: 10.1016/j.eswa.2025.127378.
[11] L. Zhang, Z. Leng, J. Liu, and Y. Wang, "Mobile robot path planning based on the improved DDPG algorithm," *Transactions of the Institute of Measurement and Control*, 2025, doi: 10.1177/01423312251375111.
[12] S. M. Nasti, Z. A. Najar, and M. A. Chishti, "Adaptive mapless mobile robot navigation using deep reinforcement learning based improved TD3 algorithm," *Frontiers in Robotics and AI*, vol. 12, Art. no. 1625968, 2025, doi: 10.3389/frobt.2025.1625968.
[13] U. Demircioğlu, H. Bakır, B. Almarri, and A. H. Abdul Hafez, "A hybrid Pure Pursuit-PID and TD3 control framework for differential-drive mobile robots," *IEEE Access*, vol. 14, pp. 80817-80838, 2026, doi: 10.1109/ACCESS.2026.3691502.
[14] U. Demircioğlu, H. Bakır, B. Almarri, and A. H. Abdul Hafez, "TD3-based reinforcement learning for adaptive PID-like control of uncertain dynamical systems," *Mathematics*, vol. 14, no. 10, Art. no. 1744, 2026, doi: 10.3390/math14101744.
[15] U. Demircioğlu and H. Bakır, "Reinforcement learning–driven proportional–integral–derivative controller tuning for mass–spring systems: Stability, performance, and hyperparameter analysis," *Engineering Applications of Artificial Intelligence*, vol. 162, Art. no. 112692, Dec. 2025, doi: 10.1016/j.engappai.2025.112692.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Dynamic Stiffness Enhancement of Cross-Laminated Timber Observation Tower through MgO Surface Coating

Sertaç Tuhta*[1]

[1]Department of Civil Engineering, Ondokuz Mayıs University, Turkey

*(stuhta@omu.edu.tr) Email of the corresponding author

Abstract – Natural disasters, particularly earthquakes and dynamic environmental loads, continue to pose significant threats to the safety and serviceability of civil engineering structures. Enhancing the dynamic performance of sustainable structural systems has therefore become an important research focus. Cross-laminated timber (CLT) structures have gained widespread attention due to their high strength-to-weight ratio, dimensional stability, and favorable structural performance. Nevertheless, further improvement of their dynamic stiffness and stability remains an important engineering challenge, particularly for slender structures such as observation towers. Among the available strengthening techniques, coating using magnesium oxide (MgO) has emerged as a promising approach because of its high stiffness, favorable mechanical properties, and strong bonding characteristics. Therefore, this study numerically investigates the influence of MgO surface coating on the dynamic stiffness of a CLT observation tower using the finite element method. Initially, a three-dimensional finite element model of the CLT observation tower was developed, and its dynamic characteristics were evaluated through modal analysis. Subsequently, the structural members were modified by incorporating an MgO surface coating layer, and the same numerical procedure was repeated. The dynamic responses of the existing and MgO-coated models were compared in terms of natural frequencies and mode shapes. The numerical results revealed that MgO coating significantly improved the dynamic stiffness of the CLT observation tower by increasing its natural frequencies and enhancing structural stiffness under static loading. These findings demonstrate that MgO surface coating is an effective strengthening strategy for improving the vibration performance and dynamic stability of CLT observation towers, providing valuable insights into the development of high-performance and sustainable timber structures.

Keywords – FEM, CLT, Modal Analysis, MgO Coating, Dynamic Stability

I. INTRODUCTION

Natural hazards such as earthquakes, strong winds, floods, landslides, and tsunamis continue to threaten the safety, serviceability, and long-term performance of buildings and civil engineering structures worldwide. Among these hazards, earthquakes are particularly critical because they generate substantial inertial forces that can induce excessive deformation, structural damage, economic losses, and human casualties. The seismic and dynamic performance of a structure is directly associated with its stiffness, mass distribution, damping properties, and energy dissipation capacity. When these parameters are not appropriately balanced, excessive lateral displacements, stress concentrations, and resonance effects may develop under dynamic excitation. Therefore, enhancing the dynamic stiffness and stability of structures has become one of the principal objectives of contemporary structural engineering ([1], [4], [8], [12], [21]).

Timber structures have received increasing attention in recent years because of their renewable nature, low embodied carbon, favorable strength-to-weight ratio, and efficient construction characteristics. The

development of engineered wood products has considerably expanded the application of timber in large-scale and multi-story structures by providing more uniform and reliable mechanical properties than conventional solid wood. Among these engineered products, cross-laminated timber (CLT) has become one of the most widely investigated structural materials due to its high in-plane and out-of-plane stiffness, dimensional stability, prefabrication potential, and efficient load-transfer capability. CLT systems are increasingly employed in residential, commercial, institutional, and specialized structures because they provide a combination of structural efficiency and environmental sustainability. Their relatively low self-weight can also reduce seismic demands compared with conventional reinforced concrete systems, making CLT an attractive material for lightweight structures in earthquake-prone regions ([2], [3], [5], [7], [14]).

Despite these advantages, lightweight CLT structures may remain vulnerable to dynamic excitation, particularly when they are designed as slender and vertically elongated structures such as observation towers. Earthquake- and wind-induced vibrations can result in excessive lateral displacement, localized stress concentrations, and reductions in structural serviceability when sufficient stiffness and damping are not provided. The dynamic characteristics of timber structures, including natural frequencies and mode shapes are strongly governed by the stiffness and mass characteristics of the structural members and their connections. Consequently, improving the stiffness characteristics of CLT components through innovative strengthening and surface modification techniques has become an important research area in advanced timber engineering ([13], [16], [20]).

Fiber-reinforced polymers, nanomaterials, mineral-based treatments, and advanced surface modification methods have recently emerged as promising approaches for improving the mechanical and durability performance of timber elements. In particular, surface engineering and nanotechnology provide significant opportunities to enhance the stiffness, hardness, dimensional stability, and environmental resistance of wood-based materials without introducing a substantial increase in structural weight. Functional surface coatings can modify the mechanical characteristics of timber members while simultaneously improving their resistance to moisture, temperature variations, biological deterioration, and environmental degradation. Such improvements may contribute not only to the local performance of timber components but also to the overall dynamic response and stiffness characteristics of timber structures subjected to seismic and wind-induced loading ([9], [11], [22]).

Among various mineral-based coating materials, magnesium oxide (MgO) has attracted considerable interest because of its favorable mechanical properties, chemical stability, thermal resistance, and potential compatibility with construction materials. MgO-based materials have been investigated in cementitious systems, protective coatings, composite materials, and construction applications because of their capacity to contribute to mechanical performance and environmental durability. When applied as a surface coating or strengthening layer, MgO may improve the stiffness and hardness of the treated surface while providing additional resistance against environmental effects. These characteristics make MgO surface treatment a potentially suitable approach for improving the mechanical and dynamic performance of CLT components used in lightweight structural systems ([24], [25], [26]).

Recent developments in engineered timber construction have encouraged researchers to investigate advanced surface treatments and coating technologies for improving the structural performance of wood-based systems. Although numerous studies have examined the effects of fiber-reinforced polymers, adhesive systems, mechanical reinforcement, and composite strengthening methods on timber members, comparatively limited attention has been given to the influence of mineral-based surface coatings on the global dynamic behavior of complete CLT structures. In particular, research concerning the numerical evaluation of MgO-coated CLT systems and their influence on the dynamic stiffness of slender timber structures remains limited. Understanding how MgO surface modification affects modal characteristics, stiffness distribution, and vibration response is essential for developing high-performance observation towers capable of satisfying modern structural safety and serviceability requirements ([18], [23], [28]).

The finite element method (FEM) has become one of the most effective numerical approaches for evaluating the structural behavior of complex timber systems subjected to static and dynamic loading. By appropriately representing structural geometry, material properties, connection conditions, boundary constraints, and interaction between structural components, FEM enables detailed assessment of structural

behavior before physical construction or strengthening applications are implemented. Modal analysis based on finite element models provides important information regarding natural frequencies, mode shapes, and dynamic stability. These parameters are particularly important for evaluating the vibration characteristics and serviceability of slender CLT structures such as observation towers ([15], [17], [19], [27]).

The dynamic stiffness of a structure is closely related to its natural vibration characteristics and its response to external dynamic excitation. During an earthquake or strong wind event, resonance may develop when the dominant excitation frequencies approach the natural frequencies of the structure, leading to amplified vibrations, increased displacement demands, and potential serviceability or structural problems. Therefore, modifications that increase structural stiffness and appropriately alter the dynamic characteristics can improve structural performance by increasing natural frequencies, reducing deformation demands, and decreasing the likelihood of unfavorable resonance conditions. For CLT observation towers, evaluating these parameters before and after the application of strengthening or coating provides valuable information regarding the effectiveness of the proposed technique ([6], [10]).

In this study, the dynamic stiffness of a CLT observation tower enhanced through MgO coating is numerically investigated using the finite element method. An existing CLT observation tower and MgO-coated CLT observation tower are developed and comparatively evaluated under identical geometric, material, and boundary conditions. The influence of MgO coating on the dynamic characteristics of the tower, including natural frequencies and mode shapes, is examined in detail. In addition, the changes in structural stiffness under static loading are evaluated to determine the contribution of the coating treatment to the overall structural performance. The numerical results aim to provide a better understanding of the effectiveness of MgO coating in enhancing the dynamic stiffness and vibration characteristics of CLT observation towers and to demonstrate the potential of mineral-based surface modification as a practical approach for developing safer, more durable, lightweight, and sustainable timber structures.

II. Materials and Method

A numerical modal analysis was performed to investigate the influence of magnesium oxide (MgO) coating on the dynamic stiffness and vibration characteristics of a Cross-laminated timber (CLT) observation tower. For this purpose, two finite element models with identical geometric configurations were developed: an existing CLT observation tower and an MgO-coated CLT observation tower. To ensure that the differences observed in the dynamic response were attributed exclusively to the applied surface modification, the structural geometry, CLT panel dimensions, support conditions, material properties of the timber components, and other modeling assumptions were kept identical in both numerical models. Accordingly, the MgO surface coating was considered the only variable throughout the numerical investigation. The finite element analyses were carried out using the SAP2000 program. The dynamic characteristics of the two structural configurations were comparatively evaluated based on their natural frequencies and corresponding mode shapes, while the changes in structural stiffness under static loading were also considered to assess the effectiveness of the MgO coating. This approach enabled the influence of the MgO coating on the dynamic stiffness and stability of the CLT observation tower to be systematically examined. The adopted numerical methodology was established in accordance with conventional finite element modeling procedures commonly applied in the analysis of timber structures, CLT systems, and structural performance under static and dynamic loading.

A. Description of CLT Observation Tower

The CLT observation tower model is a six-storey timber structure with two equal spans (3 m) in the x and y directions and a storey height of 3 m. Each storey area is 36 square meter. Column cross-sections are 25x25 cm, beam cross-sections are 20x20 cm, diagonal elements are 15x15 cm and slab thickness is 15 cm. The finite element model of the CLT structure was created using the SAP2000 program. The finite element model of the CLT observation tower is given in Figure 1. Mechanical properties of CLT are given in Table 1, where L, T, and R denote the longitudinal, tangential, and radial directions, respectively.

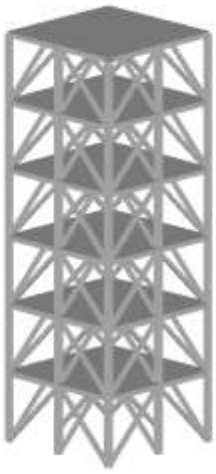


Fig. 1 3D Finite element model of the CLT observation tower

Table 1. Mechanical properties of CLT

Material	**Modulus of Elasticity (kN/m^2)**	**Poisson's Ratio**	**Weight per Unit Volume (kN/m^3)**
CLT	$9x10^6$ (L)	0.40 (LR)	4
	$0.30x10^6$ (T)	0.40 (LT)	
	$0.30x10^6$ (R)	0.01 (RT)	

L: Longitudinal, T: Tangential, R: Radial

B. Description of Coated CLT Observation Tower

The columns of the existing CLT observation tower are coated with 5 mm thick MgO. Thus, a coated model was created. The finite element model of the coated CLT observation tower is given in Figure 2. Mechanical properties of MgO are given Table 2.

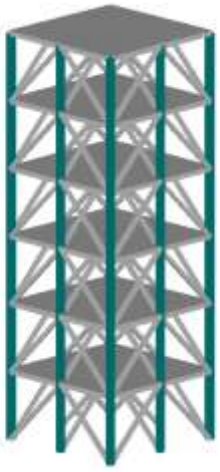


Fig. 2 3D Finite element model of the coated CLT observation tower

Table 2. Mechanical properties of MgO

Material	Modulus of Elasticity (kN/m^2)	Poisson's Ratio	Weight per Unit Volume (kN/m^3)
MgO	$330x10^6$	0.36	36

III. RESULTS

A. Modal Analysis Results of CLT Observation Tower

The modal analysis of the CLT observation tower model was performed to evaluate its dynamic characteristics under free-vibration conditions. The first five vibration modes were examined, as they typically capture the predominant dynamic behavior of low- to mid-rise timber structures and play a significant role in their seismic response. The natural frequencies and corresponding mode shapes obtained from the finite element analysis provide important insights into the stiffness distribution and overall structural behavior of the tower. The results of the modal analysis are presented in Figures 3-7, which illustrate the first five natural frequencies and their associated mode shapes, together with a discussion of their structural significance.

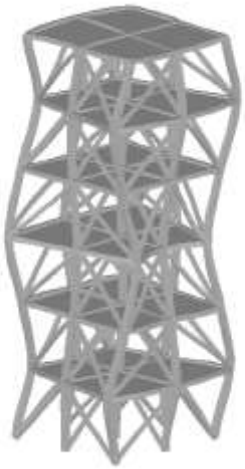


Fig. 3 1st Mode shape (Frequency value = 5.093 Hz)

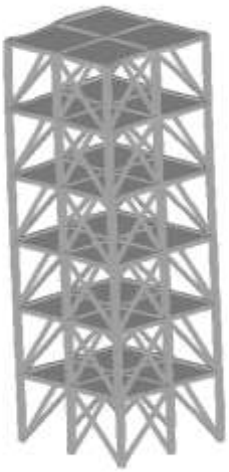


Fig. 4 2nd Mode shape (Frequency value = 6.024 Hz)

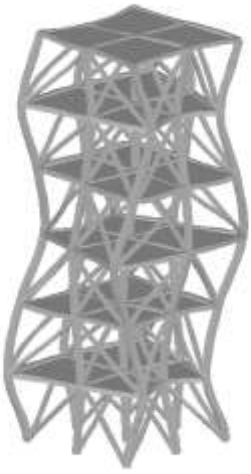


Fig. 5 3rd Mode shape (Frequency value = 8.214 Hz)

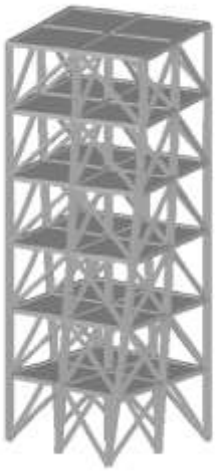


Fig. 6 4th Mode shape (Frequency value = 8.421 Hz)

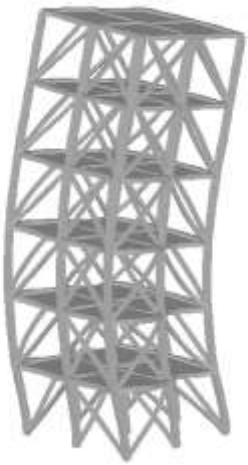


Fig. 7 5th Mode shape (Frequency value = 13.164 Hz)

B. Modal Analysis Results of Coated CLT Observation Tower

The modal analysis of the coated CLT observation tower model was conducted to assess the effect of the coating material on the dynamic characteristics of the structure. The first five vibration modes were investigated, as they primarily characterize the global dynamic response of the tower under free-vibration conditions. The introduction of the coating alters the stiffness distribution of the structural components, which in turn influences both the natural frequencies and the associated mode shapes. The resulting modal parameters provide valuable information for evaluating the contribution of the coating to the dynamic performance and overall structural stability of the CLT observation tower. The first five natural frequencies and their corresponding mode shapes obtained from the analysis are presented in Figures 8-12.

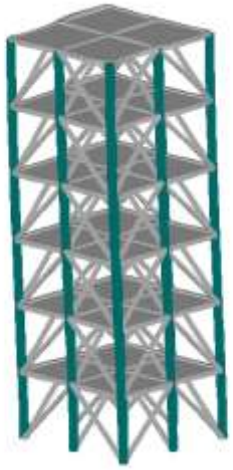


Fig. 8 1st Mode shape (Frequency value = 5.960 Hz)

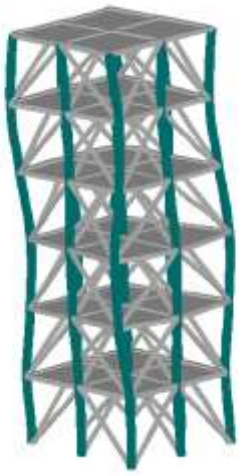


Fig. 9 2nd Mode shape (Frequency value = 6.770 Hz)



Fig. 10 3rd Mode shape (Frequency value = 9.964 Hz)

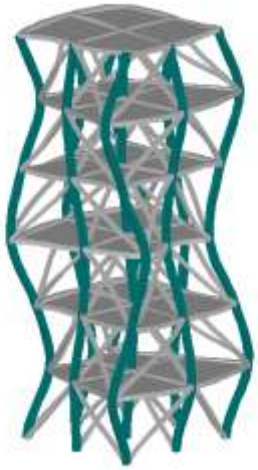


Fig. 11 4th Mode shape (Frequency value = 10.559 Hz)

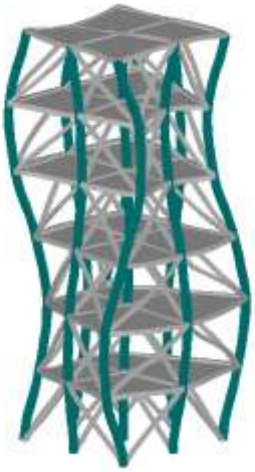


Fig. 12 5th Mode shape (Frequency value = 14.696 Hz)

IV. DISCUSSION

Table 3 presents a comparison of the natural frequencies obtained from the original and coated CLT observation tower models. The results enable a direct assessment of the effect of the coating on the dynamic properties of the structure. Comparing the modal frequencies of the two models makes it possible to evaluate how the application of the coating influences the overall stiffness and vibration characteristics of the building. The observed changes in natural frequencies reflect modifications in the structural dynamic response and provide an indication of the contribution of the coating to the overall improvement of structural performance.

Table 3. Comparison of frequency values

Mode	**1**	**2**	**3**	**4**	**5**
Existing	5.093	6.024	8.214	8.421	13.164
Coated	5.960	6.770	9.964	10.559	14.696
Difference (Hz)	0.867	0.746	1.750	2.138	1.532
Difference (%)	17.023	12.384	21.305	25.389	11.638

V. CONCLUSION

In this study, the effect of MgO coating on the dynamic characteristics of a cross-laminated timber (CLT) observation tower was numerically investigated using the finite element method. Two three-dimensional finite element models, consisting of the existing CLT observation tower and the MgO-coated CLT observation tower, were developed under the same geometric, material, and boundary conditions. A 5 mm thick MgO coating was applied to the structural columns, and the dynamic responses of the uncoated and coated configurations were subsequently evaluated through modal analysis.

The modal analysis results indicated that the application of MgO coating increased the natural frequencies of all five investigated vibration modes. The first five frequencies increased from 5.093, 6.024, 8.214, 8.421, and 13.164 Hz for the existing CLT tower to 5.960, 6.770, 9.964, 10.559, and 14.696 Hz for the coated configuration. The corresponding frequency increases were approximately 17.02%, 12.38%, 21.31%, 25.39%, and 11.64%, respectively. These results demonstrate that the MgO coating modified the stiffness characteristics of the tower and produced a measurable improvement in its global dynamic stiffness.

The increase in natural frequencies suggests that the MgO coating can alter the vibration characteristics of the CLT observation tower and potentially reduce the probability of resonance under dynamic excitation when the dominant excitation frequencies are sufficiently separated from the modified structural frequencies. This effect is particularly relevant for slender timber structures, such as observation towers, where stiffness and vibration characteristics have an important influence on structural serviceability and dynamic stability.

Overall, the numerical findings indicate that MgO coating can be considered a promising strengthening and coating technique for improving the dynamic performance of CLT observation towers. The proposed approach modifies the modal characteristics of the structure while preserving the lightweight characteristics and sustainable advantages associated with CLT construction. The results therefore provide a basis for further investigation of nano-based coatings as an alternative method for enhancing the structural performance of engineered timber systems.

REFERENCES

[1] S. Tuhta, "GFRP Retrofitting effect on the dynamic characteristics of model steel structure," *Steel Compos. Struct.*, vol. 28, no. 2, pp. 223-231, July 2018.

[2] Q. Xu, H. Zou, Z. Wang, Y. He, L. Qi, and J. Wang, "Modal testing and analysis of high-rise laminated timber building," *BioResources*, vol. 19, no. 4, pp. 9616-9630, Oct. 2024.

[3] M. M. Bayramoglu, A. Demir, A. U. Birinci, H. Ozturk, O. İlhan, Z. Candan, and C. Demirkir, "Effect of lumber quality grade on the mechanical properties and product costs of cross-laminated timber panels," *BioResources*, vol. 20, no. 2, pp. 3519-3533, Mar. 2025.

[4] M. A. Hammoud, S. Tuhta, and F. Günday, "Determination of modal parameters of reinforced concrete tunnel retrofitted with CFRP using finite element method," *Int. J. Innov. Eng. Res. Technol.*, vol. 9, no. 4, pp. 10-18, Apr. 2022.
[5] E. D. Gezer, A. U. Birinci, A. Demir, H. Öztürk, O. İlhan, and C. Demirkir, "Determining the optimum layer combination for cross-laminated timber panels according to timber strength classes using Artificial Neural Networks," *BioResources*, vol. 19, no. 3, pp. 4899-4917, Jun. 2024.
[6] S. Tuhta, "Comparative analytical modal analysis of LVL shear-walled structure retrofitted with alumina (Al_2O_3) nanocoating exposed to earthquake effect," *Buildings*, vol. 16, no. 6, pp. 1-21, June 2026.
[7] C. Choi, E. Kojima, K.-J. Kim, M. Yamasaki, Y. Sasaki, and S.-G. Kang, "Analysis of mechanical properties of cross-laminated timber (CLT) with plywood using Korean larch," *BioResources*, vol. 13, no. 2, pp. 2715-2726, Feb. 2018.
[8] F. Bakkour, S. Tuhta, and F. Günday, "Determination of modal parameters of reinforced concrete box culvert retrofitted with GFRP using finite element method," *Int. J. Innov. Eng. Res. Technol.*, vol. 9, no. 4, pp. 1-9, Apr. 2022.
[9] S. Tuhta, "The determination of effect of TiO_2 on dynamic behavior of scaled WPC warehouse by OMA," *Adv. Nano Res.*, vol. 12, no. 1, pp. 65-72, Jan. 2022.
[10] N. Sunar and H. Uysal, "Dynamic characteristics of CFRP-strengthened wooden beams: experimental and numerical study," *Struct. Eng. Mech.*, vol. 89, no. 3, pp. 323-334, Feb. 2024.
[11] S. Tuhta, "The determination of effect of TiO_2 on dynamic behavior of scaled concrete structure by OMA," *Adv. Nano Res.*, vol. 11, no. 6, pp. 641-648, Dec. 2021.
[12] A. A. Kasımzade and S. Tuhta, "OMA of model steel structure retrofitted with CFRP using earthquake simulator," *Earthq. Struct.*, vol. 12, no. 6, pp. 689-697, June 2017.
[13] S. Tuhta, F. Günday, and M. O. Warayth, "The effect of GFRP steel silo on modal parameters using finite element method," *Int. J. Innov. Eng. Res. Technol.*, vol. 8, no. 7, pp. 41-46, July 2021.
[14] M. Kaya, I. Ružiak, R. Bulbul, and V. Cavuş, "Fiber fabric-reinforced laminated veneer lumber (LVL) as insulation material for green buildings," *Materials*, vol. 19, no. 5, pp. 1-17, Feb. 2026.
[15] S. Tuhta, F. Günday, and N. C. Pehlivan, "Investigation of CFRP retrofitting effect on masonry dome on bending moment using finite element method," *Int. J. Innov. Eng. Res. Technol.*, vol. 6, no. 6, pp. 18-22, June 2019.
[16] T. Catterou, Y. Sousseau, S. M. Elachachi, M. Chaplain, and C. Faye, "Non-linear dynamic analysis of timber frame structure with bolted-fastener connections," *Vibration*, vol. 7, no. 4, pp. 1156-1170, Nov. 2024.
[17] S. Tuhta, F. Günday, and H. Aydın, "Dynamic analysis of model steel structures retrofitted with GFRP composites under microtremor vibration," *Int. J. Trend Sci. Res. Dev.*, vol. 3, no. 2, pp. 729-733, Feb. 2019.
[18] D. Buck, X. Wang, O. Hagman, and A. Gustafsson, "Bending properties of cross laminated timber (CLT) with a 45° alternating layer configuration," *BioResources*, vol. 11, no. 2, pp. 4633–4644, Apr. 2016.
[19] S. Tuhta, "OMA of GFRP retrofitted model concrete structure using ambient vibration," *Int. J. Trend Sci. Res. Dev.*, pp. 360-370, vol.3, no. 1, Nov. 2018.
[20] J. Topler, L. Buchholz, J. Lukas, and U. Kuhlmann, "Guidelines for a finite element based design of timber structures and their exemplary application on modelling of beech LVL," *Buildings*, vol. 13, no. 2, pp. 1-27, Feb. 2023.
[21] F. Günday, "OMA of RC industrial building retrofitted with CFRP using SSI," *Int. J. Adv. Eng. Res. Dev.*, vol. 5, no. 5, pp. 759-771, May. 2018.
[22] S. Tuhta and F. Günday, "Investigation of effect of using nanocoating on wooden sheds on dynamic parameters," *Wood Res.*, vol. 66, no. 6, pp. 1006-1014, Dec. 2021.
[23] A. C. İlçe, "Mechanical properties of laminated veneer lumber made from ash and red pine woods," *BioResources*, vol. 13, no. 4, pp. 8653-8661, Oct. 2018.
[24] S. Tuhta and F. Günday, "Investigation of steel chimney retrofitted with nanocoating under earthquake excitation using FEM," *Coatings*, vol. 15, no. 5, pp. 1-20, May 2025.
[25] F. Günday, "Investigation of the effect of coating light steel container houses with nano-TiO_2 on dynamic parameters using OMA," *Buildings*, vol. 15, no. 6, pp. 1-17, Mar. 2025.
[26] S. Tuhta and F. Günday, "Determination of the effect of TiO_2 on the dynamic behavior of scaled concrete chimney by OMA," *Mater. Tehnol.*, vol. 55, no. 3, pp. 459-466, May. 2021.
[27] F. Günday, "Structural health monitoring of steel garage model with stochastic subspace identification covariance variance analysis," *Buildings*, vol. 15, no. 6, pp. 1-18, Mar. 2025.
[28] F. B. T. Imamura, Y. Chen, L. Deng, and Y. H. Chui, "Dimensional changes of cross-laminated timber basement walls under ambient humidity variation," *BioResources*, vol. 20, no. 2, pp. 4568-4589, Apr. 2025.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Yeni İnşaat Birim Fiyat Pozlarının Metin Tabanlı Aile Sınıflandırılması: 2020–2026 Zamansal Genelleme Deneyi

Uğur Kafkas[1*]

[1]*İnşaat Teknolojisi / Kütahya Teknik Bilimler MYO, Kütahya Dumlupınar Üniversitesi, Türkiye*

**Sorumlu yazar: ugur.kafkas@dpu.edu.tr*

Özet – Metin sınıflandırma modelleri çoğunlukla rastgele bölünmüş veri üzerinde değerlendirilir. Rastgele satır bölmeleri, yinelenen veya ilişkili kayıtların ve birebir aynı tariflerin eğitim ile test kümelerine dağılmasına izin verebilir; bu da gelecekteki belgelere genelleme başarımını olduğundan iyimser gösterebilir. Bu çalışma, Çevre, Şehircilik ve İklim Değişikliği Bakanlığı Yüksek Fen Kurulu tarafından yayımlanan inşaat birim fiyat kataloglarının 2020, 2023 ve 2026 sürümlerini kullanarak zamansal bir değerlendirme düzeni kurmaktadır. Modeller yalnız 2020 kataloğundaki 1.509 poz açıklamasıyla eğitilmiş; 2026 kataloğunda ortaya çıkan, poz kodu eğitimde bulunmayan ve normalize açıklaması eğitimde birebir görülmeyen 237 poz üzerinde sınanmıştır. Poz kodunun ilk iki bloğu doğal sınıf etiketi olarak kullanılmış; kod, bölüm başlığı, birim ve fiyat alanları modelden gizlenmiştir. Önceden tanımlanmış dört modelleme hattı karşılaştırılmıştır: çoğunluk sınıfı, sözcük TF-IDF, karakter TF-IDF ve sabit çok dilli cümle gömmesi. Karakter TF-IDF + LinearSVC hattı birincil kümede 0,9831 doğruluk ve gerçek desteği bulunan sınıflar üzerinden 0,9143 Macro-F1 vermiştir; sözcük TF-IDF ile farkı +0,0327, cümle gömmesi ile farkı +0,1348 olup her iki eşlenik bootstrap fark aralığı da sıfırı içermemiştir. 2026'da eklenen 567 yeni kodun 175'i (%30,9) eğitimde hiç görülmeyen yedi aileye aittir; bu kayıtlar için başarı metriği hesaplanmamış, kapalı-küme sınıflandırmanın yapısal sınırı olarak ayrıca raporlanmıştır. Yöntem, model sonuçları görülmeden sabitlenmiş bir protokole bağlanmıştır; nihai deney tek kez yürütülmüş ve sonuçlar görüldükten sonra hiçbir model, hiperparametre veya karar eşiği değiştirilmemiştir. Elde edilen bulgular, karakter TF-IDF tabanlı hattın bu veri ve bu değerlendirme düzeni altındaki dayanıklılığına işaret etmektedir.

Anahtar Kelimeler – metin sınıflandırma, zamansal genelleme, inşaat birim fiyat, karakter n-gram, TF-IDF, cümle gömmesi, Türkçe doğal dil işleme

I. GİRİŞ

İnşaat maliyet belgelerinde iş kalemleri serbest metinle tanımlanır ve bu tanımların standart bir sınıflandırmaya eşlenmesi, metraj, yaklaşık maliyet ve teklif süreçlerinin temelini oluşturur. Türkiye'de kamu yapım işlerinde bu eşleme, Çevre, Şehircilik ve İklim Değişikliği Bakanlığı Yüksek Fen Kurulu tarafından yayımlanan birim fiyat kataloglarındaki poz kodları üzerinden yürütülür. Katalog her yıl güncellenir: yeni pozlar eklenir, bazı pozlar yürürlükten kalkar ve mevcut tariflerin bir bölümü yeniden yazılır.

Metin sınıflandırma çalışmalarının büyük çoğunluğu, veri kümesini rastgele bölerek değerlendirme yapar. Bu düzen iki bakımdan iyimser olabilir. Birincisi, aynı poz kodunun farklı sürümleri eğitim ve test kümelerine dağılabilir. İkincisi, katalog sürümleri arasında tarif metinleri büyük ölçüde değişmediği için, test kümesindeki bir metnin birebir aynısı eğitim kümesinde bulunabilir. Her iki durumda da ölçülen başarım, modelin gelecekteki yeni kayıtlara genelleme yeteneğinden çok, veri kümesindeki tekrarların bir sonucu olur.

Bu çalışma söz konusu tekrarların ikisini de dışarıda bırakan zamansal bir değerlendirme düzeni kurar ve tek bir soruyu sınar: yalnız 2020 kataloğuyla eğitilen metin sınıflandırıcıları, altı yıl sonraki 2026 kataloğunda ortaya çıkan; kodu eğitimde bulunmayan ve normalize açıklaması eğitimde birebir görülmeyen pozları, 2020'de var olan doğru poz ailesine atayabilir mi?

Çalışmanın katkıları şunlardır: (i) üç zaman kesitli, koordinat tabanlı ayrıştırma ve elle kalite kontrolünden geçmiş, denetlenmiş bir Türkçe resmî inşaat poz veri kümesi ve ayrıştırma hattı; (ii) rastgele bölme yerine altı yıllık gelecek sürüm değerlendirmesi; (iii) kod tekrarına ek olarak birebir metin tekrarının da çıkarıldığı sıkı bir test kümesi; (iv) sözcük TF-IDF, karakter TF-IDF ve çok dilli gömme tabanlı üç önceden tanımlanmış temsil–sınıflandırıcı hattının aynı doğal etiketler üzerinde karşılaştırılması; (v) kapalı-küme başarımının, eğitimde görülmeyen aile kapsamından ayrı raporlanması.

II. İLGİLİ ÇALIŞMALAR

A. İnşaat maliyet kalemlerinin sınıflandırılması

İnşaat maliyet belgelerindeki kalem açıklamalarının otomatik sınıflandırılması son yıllarda artan bir ilgi görmektedir. Deza ve ark. [1], Birleşik Krallık'ta 24 büyük altyapı projesinden derlenen 50 binden fazla kalem açıklamasını Uluslararası İnşaat Ölçüm Standardı (ICMS) sınıflarına eşleyen çok sınıflı modeller geliştirmiş ve klasik öznitelik tabanlı yöntemlerin bu görevde güçlü taban çizgileri oluşturduğunu göstermiştir. Dopazo ve ark. [2], altyapı maliyet verilerinin ortak bir yapıya otomatik olarak sınıflandırılması için metin işleme ve birden çok makine öğrenmesi sınıflandırıcısını birleştiren bir yaklaşım geliştirmiştir. Daha yakın tarihli bir çalışma derin topluluk mimarilerini maliyet sınıflandırmasına uygulamıştır [3]. Bu çalışmalar ağırlıklı olarak İngilizce belgelere ve rastgele bölmeye dayanır; sürümler arası zamansal ayrım kurulmaz.

B. Zamansal genelleme ve kavram kayması

Metin sınıflandırıcılarının başarımının zaman içinde düşebildiği yerleşik bir bulgudur. Zhao ve ark. [4], zamansal kavram kaymasının yalnız başarımı değil model açıklamalarını da etkilediğini göstermiştir. Alkhalifa ve ark. [5], üç boylamsal veri kümesinde farklı yıllara ait eğitim ve test kesitlerini karşılaştırmış; zamansal uzaklıkla birlikte başarımın genel olarak düştüğünü, ancak düşüşün büyüklüğünün veri alanının dilsel kararlılığına göre değiştiğini göstermiştir. Bu bulgu, geleceğe dönük kullanım senaryolarında kronolojik değerlendirmenin önemini desteklemektedir. Bu çalışmadaki düzen bu öneriyi izler; ek olarak yalnız kronolojik bölme değil, birebir metin tekrarının da çıkarılması koşulunu getirir.

C. Karakter n-gram ve cümle gömmesi temsilleri

Biçimbilimsel açıdan zengin dillerde karakter n-gram öznitelikleri, sözcük tabanlı temsillere göre sözlük dışı kalma sorununa karşı daha dayanıklıdır. Piskorski ve Jacquet [6], kısa metin sınıflandırmada TF-IDF ağırlıklı karakter n-gram modellerinin çeşitli gömme tabanlı yaklaşımlardan daha yüksek başarım verebildiğini bildirmiştir. Türkçe için yapılan kapsamlı bir çözümlemede de temsil seçiminin göreve göre değiştiği gösterilmiştir [7]. Çok dilli cümle gömmeleri [8] anlamsal yakınlığı yakalayabilse de, kısa ve terim yoğunluklu alan metinlerinde karakter düzeyi örüntülerin taşıdığı ayırt edici bilgiyi tam olarak temsil etmeyebilir. Bu çalışma üç temsili aynı zamansal test üzerinde karşılaştırır.

III. MATERYAL VE YÖNTEM

A. Veri kaynağı ve ayrıştırma

Çalışmada Yüksek Fen Kurulu inşaat birim fiyat kataloglarının 2020, 2023 ve 2026 sürümleri kullanılmıştır [10]. Kapsam, disiplin kısıtı olarak 15. önekli inşaat pozlarıyla sınırlandırılmıştır. Kataloglar PDF biçiminde yayımlandığı için kayıtlar koordinat tabanlı bir ayrıştırıcı ile çıkarılmıştır: poz kodu, tarif, birim ve fiyat alanları sayfa üzerindeki konumlarına göre ayrılmış, bölüm başlıkları tarif kutusunun dışında tutulmuştur. Ayrıştırma sonucunda 2020, 2023 ve 2026 kataloglarından sırasıyla 1.509, 1.560 ve 1.659 kayıt elde edilmiştir (toplam 4.728) ve hiçbir kayıtta çözülmemiş ayrıştırma uyarısı kalmamıştır. 2023 yılına ait iki resmî baskıdan, 1 Ocak–1 Temmuz arasında yürürlükte olan ilk baskı kullanılmıştır.

Bu çalışmada "yeni poz", kodu 2020 kataloğunda bulunmayıp 2026 kataloğunda bulunan kaydı ifade etmektedir; kalemin kapsamının mevzuat bakımından bütünüyle yeni olduğu iddia edilmemektedir.

Ayrıştırma kalitesi iki katmanda denetlenmiştir. Otomatik katman, iç tutarlılık ve sızıntı taramasıdır. Elle katman, sabit tohumla seçilmiş 100 kaydın PDF görüntüsü ile çıkarılan verinin görsel olarak karşılaştırılmasıdır; örneklem rastgele, sayfa sınırı, bölüm başlığı geçişi ve yüksek risk gruplarını birlikte

içerir. Elle incelenen 100 kaydın tamamında poz kodu ve kullanılabilir açıklama doğru bulunmuştur. Kod ve fiyat sızıntısı, örneklemde değil 4.728 kaydın tamamında taranmış ve sıfır bulunmuştur.

Tablo 1. Katalog sürümleri ve zamansal bölmeler

Küme	Yıl	Kayıt	Aile	Rol
Eğitim	2020	1.509	81	eğitim; etiket uzayı
Pilot (D-strict)	2023	39	6	geliştirme; sonuçlara dâhil değil
T-strict	2026	**237**	**14**	**birincil test**
T-exact	2026	155	—	birebir metin tekrarı denetimi
T-open	2026	175	7	etiket uzayı dışı; metrik yok
2026 yeni kod (toplam)	2026	567	—	üç tabakanın birleşimi

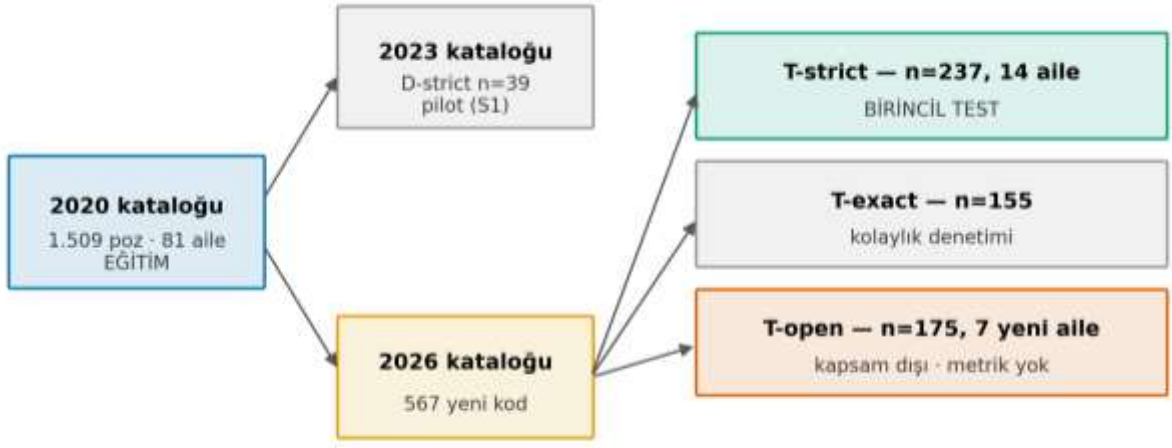
B. Etiket, model girdisi ve gizlenen alanlar

Sınıf etiketi olarak poz kodunun ilk iki bloğu (15.xxx) kullanılmıştır. Bu etiket kataloğun kendi yapısından gelir; ayrıca etiketleme yapılmamıştır. Model girdisi yalnız normalize edilmiş tarif metnidir. Normalizasyon Unicode NFKC, Türkçe farkındalıklı küçük harfe indirme ve boşluk sadeleştirmeden ibarettir; gövdeleme, durak sözcük atma, sayı atma ve ASCII katlama uygulanmamıştır. Türkçeye özgü küçük harf eşlemesi (I→ı, İ→i) bilinçli olarak seçilmiştir: standart küçük harf işlemi İ harfini iki kod noktasına ayırmakta ve karakter n-gram öznitelik uzayını değiştirmektedir.

Poz kodu, aile, eski kod, bölüm başlığı, bölüm yolu, birim, fiyat, yıl ve tüm ayrıştırma üstverisi model girdisinden çıkarılmıştır. Bu kısıt yalnız belgelenmemiş, birim testleriyle denetlenmiştir.

C. Zamansal bölmeler

Eğitim kümesi yalnız 2020 kataloğudur (1.509 kayıt, 81 aile); modelin olası çıktı uzayı bu 81 ailedir. 2023 ve 2026 kayıtları eğitime hiç girmemiştir. 2026'da bulunup 2020'de bulunmayan 567 kod, üç ayrık tabakaya bölünmüştür: ailesi eğitimde bulunanlar ve normalize tarifi eğitimde birebir bulunmayanlar (T-strict, 237 kayıt, 14 aile); ailesi bulunan fakat tarifi eğitimde birebir bulunanlar (T-exact, 155 kayıt); ailesi eğitim etiket uzayında hiç bulunmayanlar (T-open, 175 kayıt, 7 yeni aile). Birincil sonuç T-strict kümesinden gelir. 2023 kataloğundan türetilen benzer yapıdaki küçük bir alt küme (39 kayıt) yalnız önceden tanımlı bir fizibilite pilotunda kullanılmış; bu aşamada hiçbir hiperparametre ayarı yapılmamış ve küme nihai sonuçlara dâhil edilmemiştir.



Şekil 1. Zamansal değerlendirme düzeni. Model uyarlaması yalnız 2020 verisiyle yapılmıştır. 2023 verisi yalnız önceden tanımlanan fizibilite pilotunda kullanılmış ve nihai model seçimi ile test sonuçlarına katılmamıştır. 2026 verisi nihai değerlendirmeye kadar tutulmuştur.

D. Karşılaştırılan yöntemler

Dört yöntem, sonuç ne olursa olsun birlikte raporlanmak üzere önceden belirlenmiştir. M0, eğitim kümesindeki en sık aileyi her kayda atayan çoğunluk taban çizgisidir. M1, sözcük düzeyinde TF-IDF (1–2 gram) öznitelikleri üzerinde doğrusal destek vektör sınıflandırıcısıdır. M2, karakter düzeyinde TF-IDF (char_wb, 3–5 gram) öznitelikleri üzerinde aynı sınıflandırıcıdır. M3, ince ayar yapılmamış çok dilli bir cümle gömmesi [8] üzerinde lojistik regresyondur; gömme modeli sürüm kimliğiyle sabitlenmiştir. M1 ve M2 vektörleştiricileri ile M1–M3 sınıflandırıcıları scikit-learn 1.5.2 [9], M3 gömme çıkarımı ise sentence-transformers 3.3.1 ile gerçekleştirilmiştir. Vektörleştiriciler yalnız 2020 eğitim metnine uyarlanmış, test bölmelerinde yalnız dönüşüm uygulanmıştır.

E. Ölçütler ve belirsizlik

Birincil ölçüt, test kümesinde gerçek desteği bulunan sınıflar üzerinden hesaplanan Macro-F1'dir; yanında doğruluk, ağırlıklı F1 ve ilk-üç doğruluğu her zaman birlikte verilir. Duyarlılık amacıyla, tahmin edilen fakat testte gerçek desteği bulunmayan sınıfları da katan Macro-F1 değeri ayrıca raporlanır. Belirsizlik, gerçek aile içinde tabakalı 2.000 tekrarlı bootstrap ile %95 yüzdelik aralık olarak verilir; model farkları aynı yeniden örnekleme indeksleriyle eşlenik biçimde hesaplanır.

F. Yeniden üretim için ayrıntılar

M1 ve M2 hatlarında TfidfVectorizer min_df=2, sublinear_tf=True ve L2 normalizasyon ile kullanılmış; sınıflandırıcı LinearSVC (C=1,0, class_weight=balanced, squared_hinge, L2, tol=1e-4, max_iter=10000) olarak sabitlenmiştir. M3 hattında sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2 modeli (384 boyut, revision e8f8c211…, ince ayar yok) normalize edilmiş gömmelerle kullanılmış; sınıflandırıcı lojistik regresyondur (C=1,0, class_weight=balanced, lbfgs, tol=1e-4, max_iter=5000). Tüm çalışmada tek bir tohum değeri (20260808) kullanılmıştır. Ortam: Python 3.11.6, scikit-learn 1.5.2, sentence-transformers 3.3.1.

Bölmeler, model tanımları, hiperparametreler, ölçüt tanımları ve karar eşikleri nihai deney çalıştırılmadan önce sabitlenmiş ve yerel SHA-256 manifestolarıyla doğrulanmıştır. Nihai koşu bir kez yürütülmüş; sonuçlar görüldükten sonra hiçbir model, hiperparametre veya eşik değiştirilmemiştir. Mühürlenmiş kod ve yapılandırma, koşu sonrasında yeniden doğrulanmıştır. Katalog metinleri telif nedeniyle yeniden dağıtılmamaktadır.

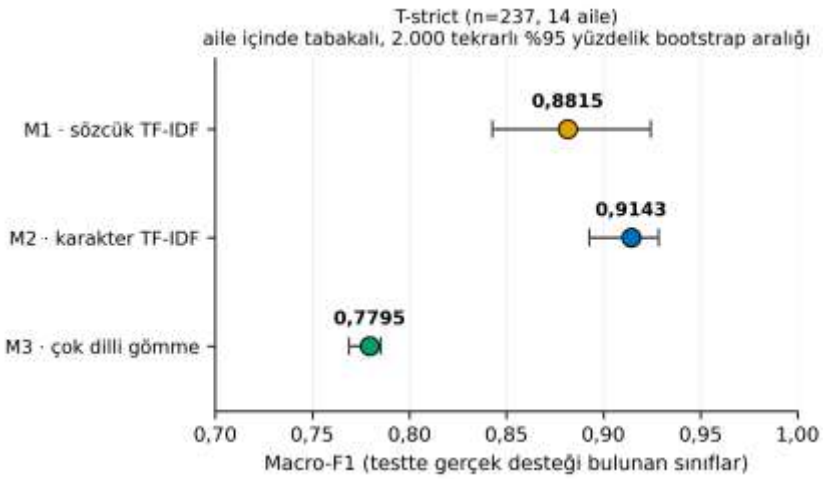
IV. BULGULAR

Birincil test kümesindeki (T-strict, 237 kayıt, 14 aile) sonuçlar Tablo 2'de verilmiştir. Karakter TF-IDF tabanlı M2 modeli 0,9831 doğruluk ve gerçek desteği bulunan sınıflar üzerinden 0,9143 Macro-F1 vermiştir. Sözcük tabanlı M1 0,9747 doğruluk ve 0,8815 Macro-F1; çok dilli cümle gömmesine dayanan M3 ise 0,9283 doğruluk ve 0,7795 Macro-F1 üretmiştir.

Tablo 2. Birincil sonuçlar — T-strict (n=237, 14 aile). Macro-F1, testte gerçek desteği bulunan sınıflar üzerindendir; köşeli parantez, aile içinde tabakalı 2.000 tekrarlı %95 yüzdelik bootstrap aralığıdır.

Model	Doğruluk	Macro-F1	Ağırlıklı F1	İlk-üç
M0 çoğunluk	0,0000	0,0000	0,0000	0,3122
M1 sözcük TF-IDF	0,9747 [0,9620–0,9873]	0,8815 [0,8429–0,9244]	0,9789	0,9831
M2 karakter TF-IDF	**0,9831 [0,9747–0,9916]**	**0,9143 [0,8929–0,9286]**	**0,9865**	**0,9831**
M3 çok dilli gömme	0,9283 [0,9072–0,9451]	0,7795 [0,7689–0,7852]	0,9385	0,9620

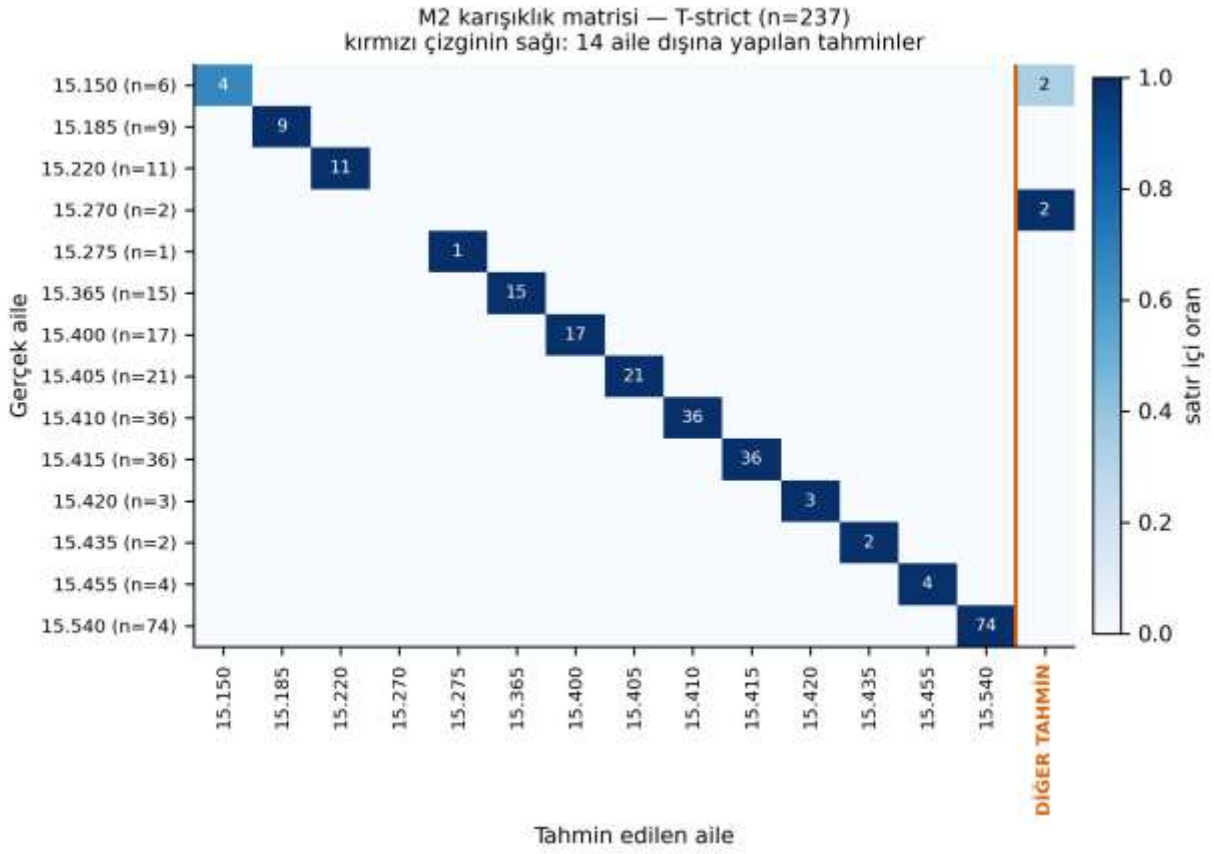
Eşlenik bootstrap ile hesaplanan Macro-F1 farkları şöyledir: M2 ile M1 arasında +0,0327 (%95 aralık [+0,0042, +0,0513]); M2 ile M3 arasında +0,1348 (%95 aralık [+0,1086, +0,1528]). Her iki aralık da sıfırı içermemiştir.



Şekil 2. T-strict üzerinde Macro-F1 ve gerçek aile içinde tabakalı 2.000 tekrarlı %95 yüzdelik bootstrap aralıkları. Macro-F1, testte gerçek desteği bulunan sınıflar üzerindendir.

Çoğunluk taban çizgisi M0, bu kümede 0,0000 doğruluk vermiştir. Bunun nedeni eğitim kümesinin en sık ailesinin (15.530, 205 kayıt) T-strict içinde hiç bulunmamasıdır. M0 bu çalışmada yalnız bir tutarlılık denetimi olarak raporlanmakta, başarı gerekçesi olarak kullanılmamaktadır; yöntem karşılaştırması M1, M2 ve M3 arasında yapılmıştır.

M2 modelinin karışıklık matrisi Şekil 3'te verilmiştir. Dört yanlış sınıflandırmanın tamamı, T-strict içinde gerçek desteği bulunan 14 ailenin dışına yapılmıştır; bu nedenle matriste ayrı bir sütun olarak gösterilmiştir. Hatalar iki aileye toplanmaktadır: 15.150 ailesinden iki kayıt ve 15.270 ailesinden iki kayıt. Her dört kayıtta da doğru aile ilk üç tahmin arasında da yer almamıştır.



Şekil 3. M2 modelinin T-strict karışıklık matrisi. Sağdaki ayrı sütun, testte gerçek desteği bulunan 14 ailenin dışına yapılan tahminleri gösterir. Renk yoğunluğu her satır kendi içinde normalize edilerek verilmiştir; hücrelerdeki sayılar ham kayıt sayılarıdır.

İkincil tabakalar Tablo 3'te özetlenmiştir. Tarifi eğitimde birebir bulunan kayıtlarda (T-exact) M1 ve M2 tam doğruluk vermiştir. Bu beklenen bir sonuçtur ve genelleme kanıtı olarak yorumlanmamalıdır; ilgili metin zaten eğitim kümesinde mevcuttur. Aynı koşulda M3, 155 kaydın 19'unu yanlış sınıflandırmıştır. Bilinen ailelerin tamamında (T-known, 392 kayıt) M2 0,9898 doğruluk ve 0,9455 Macro-F1 vermiştir.

Tablo 3. İkincil tabakalar — aynı koşudan

Tabaka	n	Model	Doğruluk	Macro-F1
T-exact	155	M1	1,0000	1,0000
T-exact	155	M2	1,0000	1,0000
T-exact	155	M3	0,8774	0,8501
T-known	392	M1	0,9847	0,9372
T-known	392	M2	0,9898	0,9455
T-known	392	M3	0,9082	0,7754

Aile bazlı sonuçlar Tablo 4'te verilmiştir. On dört aileden on ikisinde M2 tam doğruluk sağlamıştır. Destek dağılımı oldukça dengesizdir: en büyük aile 74 kayıt içerirken bir aile tek kayıtla temsil edilmektedir. Tek kayıtlı aile tabloda ayrıca işaretlenmiştir.

Tablo 4. Aile bazlı sonuçlar — T-strict, M2 modeli. *Tek kayıtlı aile.

Aile	Destek	Kesinlik	Duyarlılık	F1
15.150	6	1,000	0,667	0,800
15.185	9	1,000	1,000	1,000
15.220	11	1,000	1,000	1,000
15.270	2	0,000	0,000	0,000
15.275*	1	1,000	1,000	1,000
15.365	15	1,000	1,000	1,000
15.400	17	1,000	1,000	1,000
15.405	21	1,000	1,000	1,000
15.410	36	1,000	1,000	1,000
15.415	36	1,000	1,000	1,000
15.420	3	1,000	1,000	1,000
15.435	2	1,000	1,000	1,000
15.455	4	1,000	1,000	1,000
15.540	74	1,000	1,000	1,000

Etiket kapsamı bakımından, 2026'da eklenen 567 yeni kodun 175'i (%30,9) 2020 eğitim kümesinde hiç bulunmayan yedi aileye aittir. Bu kayıtlar için hiçbir başarı ölçütü hesaplanmamıştır; kapalı küme üzerinde çalışan bir sınıflandırıcı bu aileleri yapısal olarak üretemez.

Ölçüt tanımının etkisi de raporlanmıştır. Tahmin edilen fakat testte gerçek desteği bulunmayan sınıfları da katan Macro-F1 değerleri M1 için 0,7714, M2 için 0,8000 ve M3 için 0,4961'dir. Model sıralaması iki tanımda da aynıdır.

V. TARTIŞMA

Bulgular, karakter TF-IDF + LinearSVC hattının bu veri ve bu dondurulmuş zamansal değerlendirme düzeninde diğer iki hattan daha yüksek başarım verdiğini göstermektedir. M1 ile M2 aynı sınıflandırıcıyı kullandığından bu ikisi arasındaki fark doğrudan temsil etkisine bağlanabilir: aynı sınıflandırıcı altında karakter TF-IDF, sözcük TF-IDF'den daha yüksek Macro-F1 vermiştir. M3 ise hem temsili hem sınıflandırıcıyı birlikte değiştirdiği için M2−M3 farkı yalnız temsil farkı olarak okunamaz. Bu sonuç, karakter n-gramlarının biçimbilimsel açıdan zengin dillerde ve terim yoğunluklu kısa metinlerde taşıdığı avantajla tutarlıdır [6], [7]. Elde edilen fark, yöntemlerin genel bir sıralaması olarak okunmamalıdır; farklı bir alan, farklı bir etiket kümesi veya ince ayar yapılmış bir dil modeli farklı sonuç verebilir.

Cümle gömmesine dayanan M3'ün, tarifi eğitimde birebir bulunan kayıtlarda bile 19 hata yapması dikkat çekicidir. Sabit ve ince ayarsız bir gömme uzayında, birbirine çok benzeyen teknik tariflerin komşuluk yapısı aile sınırlarıyla her zaman örtüşmemektedir. Bu gözlem, ölçülen başarımın modelin metni anladığı biçiminde yorumlanmaması gerektiğini de göstermektedir; ölçülen yalnız sınıflandırma başarımıdır.

Hataların yapısı da bilgilendiricidir. Dört hatanın tamamı, test kümesinde gerçek desteği bulunan aileler dışına yapılmıştır. Yalnız gerçek aileleri içeren bir karışıklık matrisi bu hataların hiçbirini göstermeyecekti; bu nedenle matrise kapsam dışı tahminler için ayrı bir sütun eklenmiştir.

Çalışmanın en belirgin sınırı etiket kapsamıdır. Yeni kodların yaklaşık üçte biri, eğitimde hiç görülmeyen ailelere aittir. Kapalı küme varsayımıyla çalışan hiçbir model bu kayıtları doğru sınıflandıramaz. Uygulamada bu, yeni aile tespiti veya açık küme tanıma bileşeni gerektirir; bu çalışmada böyle bir bileşen denenmemiştir.

Diğer sınırlar şunlardır: veri tek ülkeye ve tek kuruma ait üç katalog sürümünden oluşmaktadır; sınıf dağılımı dengesizdir ve bazı aileler çok az kayıtla temsil edilmektedir; değerlendirme resmî katalog açıklamaları üzerinde yapılmıştır, şantiye dilindeki serbest metinler veya kurum dışı metraj belgeleriyle sınanmamıştır. Elle kalite kontrolü tek gözlemcilidir ve körleştirilmemiştir. Ayrıca bölüm başlığı ile tarif metni sözcük düzeyinde örtüşebilmektedir; model bölüm alanını görmese de aile sinyalinin bir bölümü bu ortak sözcüklerden gelebilir. Birincil test kümesindeki 237 kaydın 235'i benzersiz normalize tarife sahiptir; üç kayıt aynı tarifi paylaşmaktadır.

Bootstrap aralıkları, sabitlenmiş modeller altında test kümesi örnekleme belirsizliğini yansıtmaktadır; modelin yeniden eğitilmesinden, aile bileşiminden ve tek kayıtlı sınıflardan kaynaklanan belirsizliği kapsamamaktadır. Tek kayıtlı 15.275 ailesi her yeniden örneklemede aynı kayıtla temsil edildiğinden bu aile için örnekleme belirsizliği oluşmaz.

VI. SONUÇLAR

Bu çalışma, resmî inşaat birim fiyat kataloglarının üç sürümünü kullanarak zamansal bir metin sınıflandırma değerlendirmesi kurmuş; yalnız 2020 verisiyle eğitilen modelleri, 2026'da ortaya çıkan ve kodu ile normalize tarifi eğitimde bulunmayan 237 poz üzerinde sınamıştır. Karakter TF-IDF tabanlı doğrusal sınıflandırıcı bu koşulda 0,9831 doğruluk ve gerçek desteği bulunan sınıflar üzerinden 0,9143 Macro-F1 vermiş; sözcük TF-IDF ve çok dilli cümle gömmesine göre farkların eşlenik bootstrap fark aralıkları sıfırı içermemiştir. Buna karşılık yeni kodların yaklaşık %30,9'u eğitim etiket uzayı dışında kalmış ve kapalı küme sınıflandırmasının yapısal sınırını oluşturmuştur.

Gelecek çalışmalarda, kodsuz harici iş açıklamalarıyla doğrulama, yeni aile tespiti için açık küme yöntemleri ve alan içi ince ayar yapılmış dil modellerinin aynı zamansal düzen altında karşılaştırılması ele alınabilir.

KAYNAKLAR

[1] J. I. Deza, H. Ihshaish ve L. Mahdjoubi, "A machine learning approach to classifying construction cost documents into the International Construction Measurement Standard," arXiv:2211.07705, 2022.

[2] D. A. Dopazo, L. Mahdjoubi, B. Gething ve A.-M. Mahamadu, "An automated machine learning approach for classifying infrastructure cost data," Computer-Aided Civil and Infrastructure Engineering, c. 39, s. 7, ss. 1061–1076, 2024, doi: 10.1111/mice.13114.

[3] H. Sun, L. Qin ve G. Ren, "Deep ensemble learning with CNNs, DPCNNs, and LSTMs for construction cost classification," Buildings, c. 15, s. 23, mak. 4239, 2025, doi: 10.3390/buildings15234239.

[4] Z. Zhao, G. Chrysostomou, K. Bontcheva ve N. Aletras, "On the impact of temporal concept drift on model explanations," Findings of EMNLP 2022, ss. 4039–4054, 2022, doi: 10.18653/v1/2022.findings-emnlp.298.

[5] R. Alkhalifa, E. Kochkina ve A. Zubiaga, "Building for tomorrow: Assessing the temporal persistence of text classifiers," Information Processing & Management, c. 60, s. 2, mak. 103200, 2023, doi: 10.1016/j.ipm.2022.103200.

[6] J. Piskorski ve G. Jacquet, "TF-IDF character n-grams versus word embedding-based models for fine-grained event classification: A preliminary study," Proc. Workshop on Automated Extraction of Socio-political Events from News (AESPEN), LREC 2020, ss. 26–34, 2020.

[7] K. Sarıtaş, C. A. Öz ve T. Güngör, "A comprehensive analysis of static word embeddings for Turkish," Expert Systems with Applications, c. 252, mak. 124123, 2024, doi: 10.1016/j.eswa.2024.124123.

[8] N. Reimers ve I. Gurevych, "Making monolingual sentence embeddings multilingual using knowledge distillation," Proc. EMNLP 2020, ss. 4512–4525, 2020, doi: 10.18653/v1/2020.emnlp-main.365.

[9] F. Pedregosa ve ark., "Scikit-learn: Machine learning in Python," Journal of Machine Learning Research, c. 12, ss. 2825–2830, 2011.

[10] T.C. Çevre, Şehircilik ve İklim Değişikliği Bakanlığı, Yüksek Fen Kurulu Başkanlığı, "Birim Fiyatlar: İnşaat ve Tesisat Birim Fiyatları (2020, 2023/1 ve 2026)," [Çevrimiçi]. Erişim adresi: https://yfk.csb.gov.tr/birim-fiyatlar-100468 [Erişim tarihi: 9 Ağustos 2026].

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Assessing Long-Term Changes in Precipitation Frequency and Intensity Indices in Pınarbaşı (Kayseri, Türkiye)

Murat Karabulut*

Geography, Kahramanmaraş Sütçü İmam University, Türkiye

**mkarabulut@ksu.edu.tr*

Abstract – This study investigated long-term changes in precipitation frequency and intensity extremes in Pınarbaşı District (Kayseri, Türkiye) using daily precipitation records covering the period 1970–2020. Eleven ETCCDI precipitation indices were calculated with the RClimDex software. Long-term trends were evaluated using the non-parametric Mann–Kendall (MK) trend test, Sen's slope estimator, and Innovative Trend Analysis (ITA). The results revealed contrasting changes in the precipitation regime. MK and Sen's slope identified a significant decreasing trend in annual total precipitation (PRCPTOT), while ITA confirmed this decline and further highlighted distribution-dependent changes. Although MK detected no significant trends for R95p, R99p, and RX1day, ITA revealed significant increases in the upper-value group, indicating that the most intense precipitation events have become more severe in recent decades. Among the precipitation frequency indices, a significant decrease was observed in the number of days with precipitation ≥10 mm (R10mm), whereas consecutive dry days (CDD) exhibited a significant increase only in the ITA results, suggesting more persistent drought conditions. Conversely, no significant changes were found for CWD, R20mm, R25mm, or RX5day. The combined application of MK, Sen's slope, and ITA demonstrated that ITA effectively identified hidden trends overlooked by conventional monotonic tests.

Keywords – Precipitation, ETCCDI Indices, Pınarbaşı, Trend

I. INTRODUCTION

Climate change is widely recognized as one of the most significant environmental challenges of the present century because of its profound impacts on the fundamental components of the hydrological cycle. Rising global mean temperatures have increased the atmosphere's moisture-holding capacity, resulting in substantial changes in the amount, intensity, duration, and temporal distribution of precipitation [1]. Recent studies have demonstrated that even in regions where annual total precipitation has remained relatively unchanged, the characteristics of precipitation have undergone considerable alterations, particularly through an increase in the frequency of short-duration heavy rainfall events and prolonged dry periods [2], [3]. These changes have intensified the impacts of climate change on flood occurrence, drought severity, soil erosion, and water resources management.

Assessing changes in precipitation regimes requires more than the analysis of annual or seasonal precipitation totals. Even in the absence of significant changes in total precipitation, substantial shifts may occur in its temporal distribution and intensity. Consequently, standardized climate extreme indices developed by the World Meteorological Organization (WMO) and the Expert Team on Climate Change Detection and Indices (ETCCDI) have become widely adopted for evaluating the hydrological impacts of climate change [4], [5], [6]. Indices such as the Consecutive Dry Days (CDD), Consecutive Wet Days (CWD), the number of heavy precipitation days exceeding specific thresholds (R10mm, R20mm, and R25mm), annual maximum 1-day precipitation (RX1day), annual maximum consecutive 5-day precipitation (RX5day), very wet day precipitation (R95p), extremely wet day precipitation (R99p), the Simple Daily Intensity Index (SDII), and annual total wet-day precipitation (PRCPTOT) provide a comprehensive framework for characterizing different components of precipitation variability and assessing the multidimensional impacts of climate change on precipitation regimes [2].

Among the methods used to detect long-term trends in climate extreme indices, the non-parametric Mann–Kendall (MK) trend test is one of the most widely applied techniques [7]. Owing to its independence from the assumption of normal data distribution, the MK test has become a standard tool for trend detection in hydrological and climatological time series. The magnitude of monotonic trends is commonly quantified using Sen's slope estimator [8]. Nevertheless, because the MK test evaluates the entire time series as a single monotonic trend, it may fail to identify changes occurring within specific segments of the data distribution [9]. To overcome this limitation, Şen [9] proposed the Innovative Trend Analysis (ITA) method, which has gained increasing attention in hydrological and climatological research over the past decade. ITA divides the time series into two equal sub-periods and compares the ordered values relative to the 1:1 reference line, enabling independent assessment of trends in low-, medium-, and high-value groups. Consequently, ITA can reveal hidden trends and distribution-based changes that are often overlooked by conventional monotonic trend tests [10], [10]. Several studies have demonstrated that ITA provides more detailed information than the MK test in the analysis of precipitation extremes, particularly by better identifying changes in high-intensity precipitation events [11], [12].

Türkiye is considered one of the regions most vulnerable to climate change because of its location within the Mediterranean Basin [1], [13], [14]. Previous studies have revealed pronounced spatial variability in precipitation regimes across the country and reported increasing tendencies toward aridification, particularly in the Central Anatolia Region [15], [16], [17]. At the same time, several investigations have indicated that decreases in total precipitation are accompanied by increasing frequencies of short-duration heavy rainfall events in some parts of the country [18]. These contrasting tendencies suggest that climate change may simultaneously intensify both drought and flood risks.

Kayseri Province, located within the semi-arid climatic zone of Central Anatolia, is highly sensitive to precipitation variability because of its dependence on agricultural production, extensive rangelands, and limited water resources. Nevertheless, comprehensive studies investigating the long-term variability of extreme precipitation indices in Pınarbaşı District, situated in the northeastern part of Kayseri Province, remain scarce. Existing studies have primarily focused on variations in total precipitation or air temperature, whereas integrated assessments combining precipitation frequency, precipitation intensity, and climate extreme indices are still limited. Furthermore, studies simultaneously employing the Mann–Kendall test, Sen's slope estimator, and the Innovative Trend Analysis (ITA) method are relatively few. This knowledge gap limits a comprehensive understanding of hydroclimatic changes occurring within the region.

The primary objective of this study is to investigate the long-term trends of ETCCDI-based precipitation frequency and intensity indices using daily precipitation records collected in Pınarbaşı District, Kayseri Province, during the 1970–2020 period. Accordingly, the indices CDD, CWD, R10mm, R20mm, R25mm, R95p, R99p, RX1day, RX5day, SDII, and PRCPTOT were calculated, and their temporal trends were evaluated using the Mann–Kendall trend test, Sen's slope estimator, and the Innovative Trend Analysis (ITA) method. By integrating conventional monotonic trend tests with the distribution-based ITA approach, this study not only determines the direction and magnitude of long-term trends but also identifies the portions of the data distribution in which these changes occur. Consequently, the impacts of climate change on the precipitation regime of Pınarbaşı are evaluated in greater detail.

II. Materials and Method

2.1. Study Area

This study was conducted for Pınarbaşı District, located in Kayseri Province in the eastern part of the Central Anatolia Region of Türkiye (Figure 1). Pınarbaşı is situated at an elevation of approximately 1,513 m above mean sea level and represents a transitional zone characterized by a pronounced continental climate. Winters are long and cold, whereas summers are hot and dry. Precipitation occurs predominantly during the winter and spring seasons, while rainfall decreases considerably during summer. Owing to these climatic characteristics, the region constitutes an important hydroclimatic environment for investigating both drought conditions and extreme precipitation events.

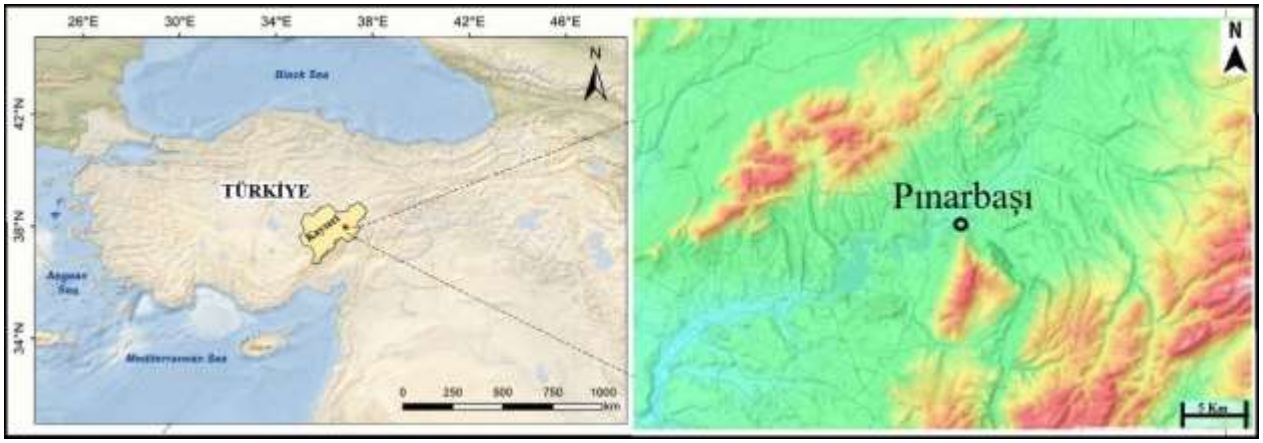


Figure 1. Location map of the study area.

2.2. Data Collection and ETCCDI Climate Indices

Daily precipitation data used in this study were obtained from the archives of the Turkish State Meteorological Service (TSMS), operating under the Ministry of Environment, Urbanization and Climate Change of the Republic of Türkiye. A continuous daily precipitation record covering the period 1970–2020, corresponding to 51 years of observations, was employed in all analyses. Long-term daily observations were selected because the World Meteorological Organization (WMO) recommends a minimum observation period of 30 years for reliable assessments of climate variability and long-term climate change [19]. The dataset underwent quality-control procedures and consisted of a continuous time series without missing observations. Standard precipitation extreme indices recommended by the World Meteorological Organization (WMO), the Climate Variability and Predictability (CLIVAR) Programme, and the Expert Team on Climate Change Detection and Indices (ETCCDI) were calculated using the daily precipitation records [4], [5].

The precipitation indices were generated using the RClimDex 1.0 software package. RClimDex is an open-source program that implements the standardized ETCCDI algorithms and has been widely used for climate variability and climate extreme analyses. The software performs quality-control procedures on daily meteorological observations and computes internationally standardized climate extreme indices [20]. In this study, a total of 11 precipitation indices representing precipitation frequency and intensity were calculated (Table 1).

Table 1. ETCCDI precipitation indices used in this study.

Index	**Description**	**Unit**
CDD	Maximum number of consecutive dry days (< 1 mm)	days
CWD	Maximum number of consecutive wet days (≥ 1 mm)	days
R10mm	Number of days with precipitation ≥ 10 mm	days
R20mm	Number of days with precipitation ≥ 20 mm	days
R25mm	Number of days with precipitation ≥ 25 mm	days
R95p	Annual total precipitation from very wet days (>95th percentile)	mm
R99p	Annual total precipitation from extremely wet days (>99th percentile)	mm
RX1day	Annual maximum 1-day precipitation	mm
RX5day	Annual maximum consecutive 5-day precipitation	mm
SDII	Simple Daily Intensity Index	mm day^{-1}
PRCPTOT	Annual total wet-day precipitation	mm

The indices CDD, CWD, R10mm, R20mm, and R25mm characterize changes in precipitation frequency, whereas R95p, R99p, RX1day, RX5day, SDII, and PRCPTOT describe precipitation intensity and the characteristics of extreme precipitation events [2].

2.3. Mann–Kendall Trend Analysis

Long-term monotonic trends in the precipitation indices were evaluated using the non-parametric Mann–Kendall (MK) trend test [13]. The MK test is one of the most widely applied methods in hydrological and climatological studies because it does not require the assumption of normally distributed data and is robust against outliers.

The Mann–Kendall statistic (S) was calculated by comparing all possible pairs of observations:

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}(x_j - x_k)$$

where x_i and x_j denote the observation values, and n is the total number of observations.

The calculated S statistic was transformed into the standardized normal test statistic (Z), and statistical significance was assessed at the 95% confidence level ($\alpha = 0.05$). Positive Z values indicate increasing trends, whereas negative values indicate decreasing trends.

2.4. Sen's Slope Estimator

The magnitude of monotonic trends was estimated using Sen's slope estimator [8]. The slope between any two observations was calculated as:

$$Q_i = \frac{x_j - x_k}{j - k}$$

The median of all pairwise slope estimates was considered the overall trend magnitude. Positive Sen's slope values indicate increasing trends over time, whereas negative values indicate decreasing trends. Owing to its non-parametric nature and low sensitivity to outliers, Sen's slope estimator is widely applied together with the Mann–Kendall test in hydrological and climatological trend analyses (Sen, 1968).

2.5. Innovative Trend Analysis (ITA)

The Mann–Kendall test evaluates the entire time series as a single monotonic trend. In contrast, the Innovative Trend Analysis (ITA) proposed by Şen [9] enables independent evaluation of trends occurring within different portions of the data distribution. Consequently, ITA has become increasingly popular in recent years for the analysis of hydrological and climatological extremes [9].

In the ITA approach, the chronological time series was divided into two equal halves. Each sub-series was independently sorted in ascending order. The ordered values of the first half were plotted on the horizontal axis, whereas those of the second half were plotted on the vertical axis. If the data points are distributed along the 1:1 reference line, no trend is present. Points located above the 1:1 line indicate an increasing trend, whereas those below the line indicate a decreasing trend.

In this study, the ITA slope (SITA) was calculated using the following equation:

$$S_{ITA} = \frac{2(\bar{Y}_2 - \bar{Y}_1)}{n}$$

Where $\bar{Y}_1$ and $\bar{Y}_2$ represent the arithmetic means of the first and second halves of the time series, respectively, and n denotes the number of observations.

The principal advantage of the ITA method is its ability to evaluate trends separately within low-, medium-, and high-value groups, thereby revealing hidden trends that cannot be detected by conventional monotonic trend tests [10] [11]. Nevertheless, previous studies have emphasized that the assumptions and limitations of the ITA method should also be carefully considered when interpreting its results.

For all statistical analyses, the significance level was set at $\alpha = 0.05$. Trends with $p < 0.05$ obtained from the Mann–Kendall and ITA analyses were considered statistically significant. To improve the robustness and reliability of the results, the direction and magnitude of trends were jointly evaluated using the Mann–Kendall test, Sen's slope estimator, and the Innovative Trend Analysis. This integrated approach enabled the identification of both monotonic trends and distribution-dependent changes occurring within different portions of the precipitation records.

III. RESULTS AND DISCUSSION

3.1. Trend Analysis of Precipitation Intensity Indices

The long-term trends of precipitation intensity indices for the Pınarbaşı (Kayseri) meteorological station during the 1970–2020 period was evaluated using the Mann–Kendall (MK) trend test, Sen's slope estimator, and Innovative Trend Analysis (ITA) (Table 2, Figure 2).

Table 2. Results of precipitation intensity indices

Index	MK (Z)	p-value	Sen's Slope	ITA Slope (SITA)	ITA (Z)	ITA p-value
R95p	-0.902	0.367	-0.359	0.250	4.158	<0.001
R99p	0.000	1.000	0.000	0.327	2.464	0.013
RX1day	-0.585	0.558	-0.042	0.068	2.375	0.017
RX5day	-1.779	0.075	-0.214	-0.064	-1.442	0.149
SDII	-1.352	0.176	-0.010	-0.004	-3.786	<0.001
PRCPTOT	-2.388	0.016	-2.152	-0.883	-3.360	<0.001

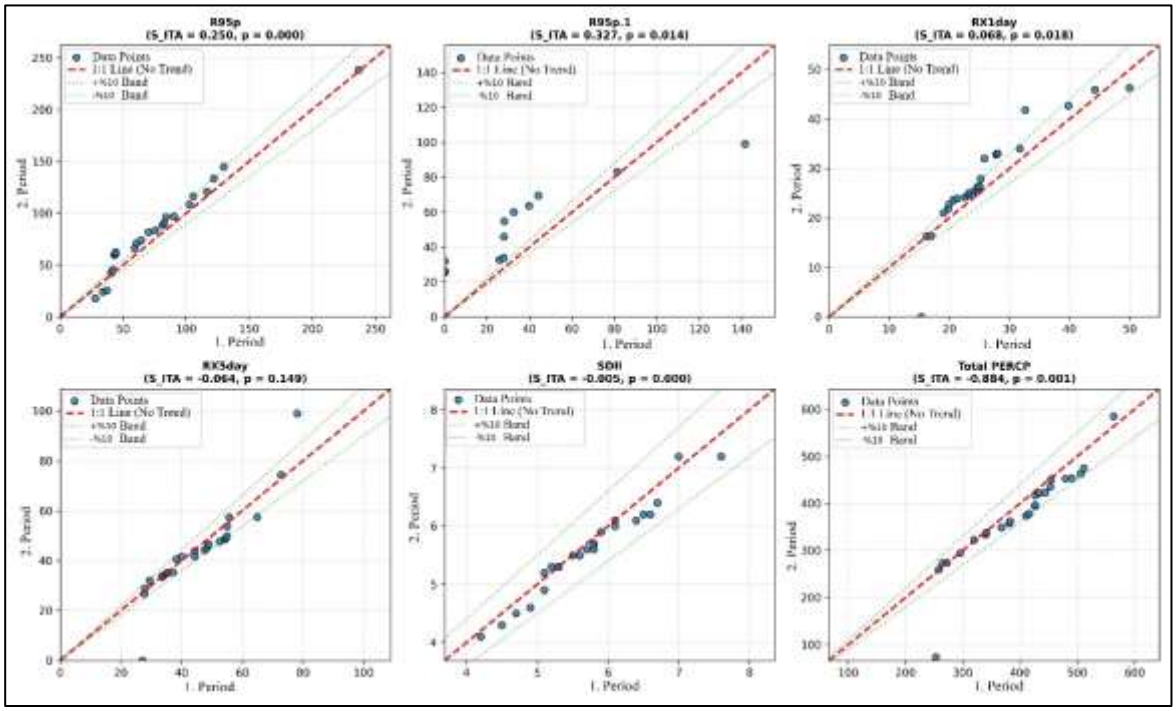


Figure 2. ITA Results of precipitation intensity indices.

For the R95p index, the MK test indicated a negative but statistically insignificant trend (S = -112, Z = -0.902, p = 0.367). Likewise, Sen's slope (-0.359 mm year^{-1}) suggested a slight decline in the annual total precipitation contributed by very wet days. In contrast, the ITA analysis produced a markedly different result. The positive ITA slope (SITA = 0.2504) together with its highly significant test statistic (Z = 4.158, p < 0.001) indicates a pronounced increase in the upper portion of the data distribution. This discrepancy demonstrates that changes in the precipitation regime are not uniformly distributed across the entire dataset. Because the MK test evaluates the time series as a single monotonic trend, decreases in low and moderate values may have offset increases in extreme values. Conversely, the ITA approach clearly reveals that extreme precipitation events became more intense during the second half of the study period. This finding is consistent with recent studies reporting an intensification of extreme precipitation across the Mediterranean Basin under climate change [1].

The MK test detected no significant trend for the R99p index (Z ≈ 0, p = 1.000), a result further supported by the zero Sen's slope estimate. However, the ITA analysis identified a statistically significant positive trend (SITA = 0.3274, Z = 2.464, p = 0.0137). This result suggests that the most extreme precipitation

events have increased only within the upper tail of the distribution, making them undetectable by conventional monotonic trend tests. Previous studies have similarly reported that the MK test may exhibit limited sensitivity for indices dominated by extreme values such as R99p, whereas ITA is more effective in detecting changes in the upper distribution [11]. Therefore, the results indicate that exceptionally rare rainfall events in Pınarbaşı have become more intense during recent decades.

For RX1day, the MK test showed a negative but statistically insignificant trend ($Z = -0.585$, $p = 0.558$), while Sen's slope (-0.042 mm year^{-1}) indicated only a negligible long-term decrease. In contrast, the ITA analysis revealed a statistically significant positive trend ($SITA = 0.0682$, $Z = 2.375$, $p = 0.0175$), indicating that annual maximum daily precipitation has increased within the upper-value group. In other words, although the average behaviour of annual maximum daily rainfall has remained relatively stable, the most severe events have become increasingly intense. This finding agrees with numerous global studies demonstrating that climate change primarily amplifies the magnitude of extreme precipitation events [1], [2].

Neither the MK test ($Z = -1.779$, $p = 0.075$) nor the ITA analysis ($Z = -1.442$, $p = 0.149$) detected a statistically significant trend for RX5day. The consistency between the two methods indicates that no systematic long-term change has occurred in the annual maximum consecutive five-day precipitation totals. This suggests that although short-duration extreme rainfall events have intensified, cumulative multi-day precipitation has not yet experienced a comparable change.

For SDII, the MK test indicated a negative but statistically insignificant trend ($Z = -1.352$, $p = 0.176$). In contrast, the ITA analysis identified a strong and statistically significant negative trend ($SITA = -0.0048$, $Z = -3.786$, $p < 0.001$). This finding demonstrates that the average precipitation intensity per wet day has decreased, particularly within the upper-value group. Consequently, the reduction in total annual precipitation appears to result not only from changes in the number of wet days but also from a decline in daily rainfall intensity.

The most pronounced result was obtained for the annual total precipitation index (PRCPTOT). The MK test revealed a statistically significant decreasing trend ($S = -295$, $Z = -2.388$, $p = 0.0169$). Sen's slope (-2.15 mm year^{-1}) indicates that annual precipitation has decreased by approximately 2 mm per year throughout the study period. The ITA analysis fully confirmed this result ($SITA = -0.8835$, $Z = -3.360$, $p < 0.001$). The agreement between the two methods demonstrates that the decline in annual precipitation is evident throughout the entire data distribution and is particularly pronounced during wetter years. This finding is consistent with previously reported aridification trends in Central Anatolia [1], [16].

Overall, the combined evaluation of the MK test, Sen's slope estimator, and ITA indicates that the precipitation regime in Pınarbaşı exhibits two contrasting long-term tendencies. On one hand, the PRCPTOT and SDII indices indicate an overall decline in annual precipitation amount and rainfall intensity. On the other hand, the R95p, R99p, and RX1day indices reveal an increasing tendency in the magnitude of the most intense precipitation events. The fact that ITA detected statistically significant increases for R95p, R99p, and RX1day, whereas the MK test did not, indicates that these changes are confined primarily to the upper portion of the data distribution. This demonstrates that conventional monotonic trend tests may overlook important changes associated with extreme events, whereas ITA is capable of identifying such hidden distribution-dependent trends more effectively [9], [11].

3.2. Trend Characteristics of Precipitation Frequency Indices

The long-term variations in precipitation frequency indices for Pınarbaşı District (Kayseri Province) over the 1970–2020 period were evaluated using the Mann–Kendall (MK) trend test, Sen's slope estimator, and Innovative Trend Analysis (ITA) (Table 3, Figure 3).

Table 3. Results of precipitation frequency indices

Index	MK (Z)	p-value	Sen's Slope	ITA Slope (SITA)	ITA (Z)	ITA p-value
CDD	0.682	0.495	0.176	0.668	2.021	0.043
CWD	-0.834	0.404	0.000	-0.003	-0.650	0.515
R10mm	-2.203	0.027	0.096	-0.039	-1.378	0.168
R20mm	-0.744	0.456	0.000	-0.006	-0.210	0.833
R25mm	-0.274	0.784	0.000	0.006	0.451	0.652

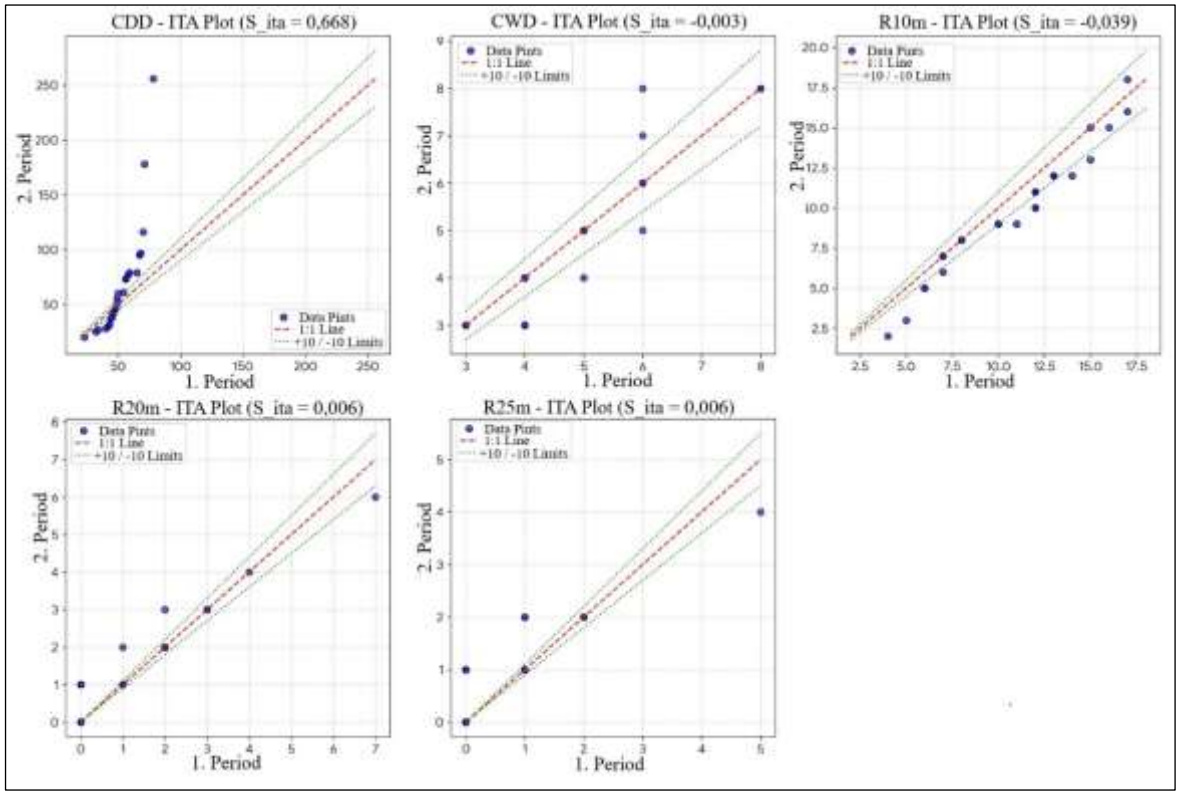


Figure 3. ITA Results of precipitation frequency indices.

For the CDD index, the MK test indicated a positive trend; however, the trend was not statistically significant (Z = 0.682, p = 0.495). Similarly, Sen's slope (0.176 days year^{-1}) suggests that the increase occurred at a relatively low rate. Nevertheless, the ITA plot reveals that data points corresponding to the upper-value group are predominantly distributed above the 1:1 reference line, indicating that prolonged dry periods became more pronounced during the second half of the study period. In other words, because the MK test evaluates the entire time series as a single monotonic trend, the increase was not identified as statistically significant. In contrast, the ITA method demonstrates that this increase is primarily confined to the extreme drought values. This finding is consistent with previous studies highlighting the superiority of ITA in detecting changes occurring within the upper portion of the data distribution [11].

For the CWD index, both the MK analysis (Z = -0.834, p = 0.404) and the ITA results provide consistent evidence. The calculated ITA slope (SITA = -0.003) is very close to zero, and most data points are symmetrically distributed around the 1:1 reference line. These results indicate that no systematic long-term change has occurred in the duration of consecutive wet days during the last five decades. Accordingly, the persistence of wet periods in the study area appears to have remained largely unchanged.

The most remarkable finding was obtained for the R10mm index. The MK test revealed a statistically significant negative trend (Z = -2.203, p = 0.027). Likewise, the ITA plot shows that most of the medium- and high-value observations are located below the 1:1 reference line. The agreement between the two approaches demonstrates a clear decline in the annual number of days with precipitation equal to or greater

than 10 mm over the past 50 years. Although Sen's slope is relatively small, indicating a slow rate of decrease, the persistence of this trend suggests a gradual but continuous reduction in moderate precipitation events. This finding implies that climate change in Pınarbaşı has primarily affected the frequency of moderate rainfall events. Similar reductions in the number of wet days and moderate precipitation events, rather than total annual precipitation, have been widely reported for both the Mediterranean Basin and Türkiye [1], [2], [9].

The R20mm and R25mm indices exhibited a different pattern. Although the MK test identified negative trends for both indices, neither trend was statistically significant ($p > 0.05$). Similarly, the ITA plots show that most observations are clustered around the 1:1 reference line, with only a few deviations among the highest values. The near-zero ITA slope values indicate that no systematic change has occurred in the frequency of heavy and very heavy precipitation events. Therefore, despite the observed decrease in moderate rainfall events, the occurrence frequency of heavy precipitation events has remained largely unchanged during the study period.

The combined interpretation of the MK and ITA results demonstrates that changes in the precipitation regime are not spatially or statistically homogeneous across the data distribution. While the MK test identified a significant trend only for the R10mm index, the ITA approach clearly revealed an increase in the upper-value group of the CDD index. This finding indicates that conventional monotonic trend tests may overlook changes occurring within specific segments of the data distribution. Indeed, Dabanlı et al. (2016) and Güçlü (2020) emphasized that ITA can independently evaluate trends in low-, medium-, and high-value groups, thereby providing a more detailed assessment of hidden trends in hydroclimatological time series than conventional trend analysis methods. Level-2 and level-3 headings can be used to detail main headings.

IV- CONCLUSIONS

This study evaluated the long-term trends of ETCCDI-based precipitation frequency and intensity indices in Pınarbaşı (Kayseri, Türkiye) using daily precipitation records for the period 1970–2020. The combined application of the Mann–Kendall (MK) trend test, Sen's slope estimator, and Innovative Trend Analysis (ITA) provided a comprehensive assessment of both monotonic trends and distribution-dependent changes in the precipitation regime.

The results indicate that the precipitation regime in Pınarbaşı has undergone substantial changes over the last five decades. Annual total precipitation (PRCPTOT) exhibited a statistically significant decreasing trend according to both the MK and ITA methods, demonstrating a progressive reduction in regional water availability. A significant decline was also observed in the Simple Daily Intensity Index (SDII), indicating that the average rainfall intensity per wet day has decreased over time. In addition, the R10mm index revealed a significant reduction in the frequency of moderate rainfall events, suggesting that rainfall events capable of contributing to soil moisture replenishment and groundwater recharge have become less frequent.

In contrast, the behaviour of the most intense precipitation events differed from the general precipitation pattern. Although the MK test did not identify statistically significant trends for R95p, R99p, and RX1day, the ITA method revealed significant positive trends within the upper portion of the data distribution. These findings indicate that the most extreme rainfall events have become more intense during recent decades despite the overall decline in annual precipitation. Such contrasting behaviour suggests that changes in precipitation characteristics are not homogeneous across the data distribution and that conventional monotonic trend tests alone may underestimate changes occurring in extreme events.

Among the precipitation frequency indices, ITA identified a significant increase in the upper-value group of consecutive dry days (CDD), indicating an increasing tendency toward prolonged drought periods. Conversely, no statistically significant long-term changes were detected for consecutive wet days (CWD), the frequency of heavy precipitation events (R20mm and R25mm), or the annual maximum consecutive five-day precipitation (RX5day). These findings suggest that climate change has primarily influenced the duration of dry periods and the occurrence of moderate rainfall events, while the frequency of heavy precipitation events has remained relatively stable.

REFERENCES

[1] IPCC. *Climate Change 2021: The Physical Science Basis.* Cambridge University Press., 2021.

[2] Alexander, L. V., Zhang, X., Peterson, T. C., Caesar, J., Gleason, B., Klein Tank, A. M. G., Haylock,

[3] Donat, M. G., Alexander, L. V., Yang, H., Durre, I., Vose, R., Caesar, J., et al. "Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset", *Journal of Geophysical Research: Atmospheres*, 118, 2098–2118, 2013.

[4] Karl, T.R., Nicholls, N., Ghazi, A. CLIVAR/GCOS/WMO Workshop on Indices and Indicators for Climate Extremes Workshop Summary. In: Karl, T.R., Nicholls, N., Ghazi, A. (eds) *Weather and Climate Extremes*. Springer, Dordrecht, 1999.

[5] Peterson, T. C., Folland, C., Gruza, G., Hogg, W., Mokssit, A., & Plummer, N. *Report on the activities of the Working Group on Climate Change Detection and Related Rapporteurs*. WMO Rep. WCDMP-47, 2001.

[6] Aydın K., Karabulut M., "1970-2020 Yılları Arasında Antakya Kahramanmaraş Grabeninde Ekstrem Sıcaklıkların Trend Analizi", *Avrupa Bilim ve Teknoloji Dergisi*, 52, 46-63, 2023.

[7] Karabulut, M., "Precipitation Trends in Kahramanmaraş along with Gaziantep and Adıyaman during the period of 1963-2005", *Ekoloji,* 18 (71), s. 15-27, 2009.

[8] Sen, P. K. "Estimates of the regression coefficient based on Kendall's tau". *Journal of the American Statistical Association*, 63(324), 1379–1389, 1968.

[9] Şen, Z. "Innovative trend analysis methodology." *Journal of Hydrologic Engineering*, 17(9), 1042–1046, 2012.

[10] Dabanlı, İ., Şen, Z., Yeleğen, M. Ö., Şişman, E., Selek, B., & Güçlü, Y. S. "Trend assessment by the innovative Şen method". *Water Resources Management*, 30, 5193–5203, 2016.

[11] Güçlü, Y. S. "Improved visualization for trend analysis by comparing with classical Mann–Kendall test and innovative trend analysis". *Journal of Hydrology,* 584, 124674, 2020.

[12] Şen, Z. *Innovative Trend Methodologies in Science and Engineering*. Springer Cham., 2017.

[13] Karabulut, M. ve F. Cosun, "Kahramanmaraş ilinde yağışların trend analizi", *Coğrafi Bilimler Dergisi*, 7(1), s. 65-83, (2009).

[14] Giorgi, F. "Climate change hot-spots." *Geophysical Research Letters*, 33, L08707, 2006.

[15] Türkeş, M. "Spatial and temporal analysis of annual rainfall variations in Turkey." *International Journal of Climatology*, 16, 1057–1076., 1996.

[16] Türkeş, M. and Tatlı, H. "Use of the standardized precipitation index (SPI) and a modified SPI for shaping the drought probabilities over Turkey". *Int. J. Climatol*., 29: 2270-2282, 2009.

[17] Demircan, M., Gürkan, H., Eskioğlu, O., Arabacı, H., & Coşkun, M. "Climate Change Projections for Turkey: Three Models and Two Scenarios." *Turkish Journal of Water Science and Management*, 1(1), 22-43, 2017.

[18] Kızılelma, Y., Çelik, M. A. ve Karabulut, M., "İç Anadolu Bölgesinde sıcaklık ve yağışların trend analizi", *Türk Coğrafya Dergisi,* 64, 1-10, 2015.

[19] World Meteorological Organization (WMO). *Guidelines on the Calculation of Climate Normals* (WMO-No. 1203), 2017.

[20] Alashan S. *Combination of modified Mann-Kendall method and Şen innovative trend analysis*. Engineering Reports. 2020; 2:e12131. 2020.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Assessment of Modal Behaviour in SiO_2-Coated Glued Laminated Timber Pole Shed via Finite Element Simulation

Sertaç Tuhta[*1]

[1]*Department of Civil Engineering, Ondokuz Mayıs University, Turkey*

[*]*(stuhta@omu.edu.tr) Email of the corresponding author*

Abstract **–** Natural hazards, particularly earthquakes and environmental actions, represent major challenges to the structural safety of civil engineering systems. Consequently, improving the dynamic performance of sustainable timber structures has become an increasingly important area of research. Glued laminated timber (GLT) structures have attracted considerable interest due to their high strength-to-weight ratio, structural efficiency, and suitability for lightweight construction. However, enhancing their stiffness remains an important engineering concern, especially for slender and relatively flexible structures such as pole sheds. Among potential strengthening methods, silicon dioxide (SiO_2) coating offers a promising alternative due to its favorable mechanical characteristics and potential to improve the structural response. Accordingly, this study numerically evaluates the influence of SiO_2 coating on the modal behaviour and structural performance of a glued laminated timber pole shed using the finite element method. First, a three-dimensional finite element model of the GLT pole shed was established, and its dynamic characteristics were determined through modal analysis. Subsequently, an equivalent numerical model incorporating a SiO_2 coating layer on the structural timber members was developed, and the same analysis procedure was performed. The responses of the existing and SiO_2-coated configurations were then comparatively assessed based on their natural frequencies and mode shapes under static loading. The numerical findings indicated that the application of SiO_2 coating enhanced the stiffness characteristics of the GLT pole shed, resulting in increased natural frequencies and improved structural response. These results indicate that SiO_2 coating can serve as an effective strengthening technique for enhancing dynamic stability of GLT pole shed structures, while also contributing to the development of durable, lightweight, and sustainable timber structural systems.

Keywords – FEA, GLT, Modal Analysis, SiO_2 Coating, Structural Performance

I. INTRODUCTION

Natural hazards, including earthquakes, severe wind events, floods, landslides, and tsunamis, represent persistent challenges to the safety, functionality, and durability of civil engineering structures. Among these hazards, earthquakes and strong winds are particularly important for lightweight and slender structural systems because they can generate significant dynamic effects, including lateral deformation, vibration, stress concentration, and resonance. The structural response to dynamic excitation is governed by several fundamental parameters, such as stiffness, mass distribution, damping characteristics, structural configuration, and energy dissipation capacity. An inadequate balance among these parameters may result in excessive vibration, undesirable deformation, and deterioration of structural serviceability. Consequently, improving the dynamic characteristics and vibration resistance of lightweight structural systems has become an important objective in contemporary structural engineering research ([1], [4], [8], [12], [21]). Timber construction has gained substantial interest as an alternative to conventional structural materials because of its renewable character, relatively low environmental impact, high strength-to-

weight ratio, and suitability for prefabricated construction. Advances in engineered wood products have further expanded the structural applications of timber by providing improved dimensional consistency, mechanical reliability, and manufacturing flexibility compared with conventional solid-sawn wood. Among these engineered timber products, glued laminated timber (GLT) has become an important structural material for beams, columns, frames, roofs, bridges, and specialized lightweight structures. GLT is manufactured by bonding multiple layers of timber laminations, enabling the production of structural members with controlled geometry and relatively uniform mechanical characteristics. Its favorable strength-to-weight ratio and architectural flexibility make GLT particularly suitable for structures in which reduced self-weight, long spans and efficient construction are important design considerations ([2], [3], [5], [7], [14]).

Pole shed structures represent a practical application of timber construction in which vertically oriented structural members, roof elements, and framing components form a relatively lightweight structural system. Such structures can be used for agricultural, storage, recreational, shelter, and other functional purposes where economical construction and rapid assembly are desirable. The structural behavior of pole sheds is strongly influenced by the stiffness of the vertical members, roof system, connections, and overall framing configuration. Because of their relatively low mass and flexible structural characteristics, pole sheds may exhibit noticeable vibration under dynamic environmental actions, particularly under strong wind and seismic excitation. The modal characteristics of these systems, including natural frequencies and mode shapes, therefore constitute important indicators for evaluating their structural behavior and serviceability ([13], [16], [20]).

The dynamic performance of GLT structures is closely associated with the mechanical characteristics of their individual members and the interaction between structural components. Variations in member stiffness, connection rigidity, boundary conditions, and mass distribution can significantly affect the natural vibration characteristics of the complete structural system. For lightweight timber structures, even relatively small modifications to the stiffness of structural components may produce measurable changes in natural frequencies and mode shapes. Accordingly, strengthening and surface modification techniques have increasingly been investigated as potential methods for improving the mechanical and dynamic characteristics of engineered timber structures. Such approaches are particularly relevant for structures in which increasing member dimensions or introducing heavy strengthening components may be undesirable because of additional mass, construction complexity, or architectural limitations ([9], [11], [22]).

Surface coatings and advanced material treatments provide an alternative strategy for modifying the performance of timber components while maintaining the basic structural configuration. Fiber-reinforced polymers, polymeric coatings, mineral-based materials, nanomaterials, and other surface engineering technologies have been investigated for improving the strength, stiffness, durability, moisture resistance, and environmental performance of timber. In addition to protecting wood surfaces against adverse environmental conditions, appropriately selected coatings may provide an additional stiffening effect to structural members. The resulting modification of member stiffness can influence the global dynamic response of the structural system by changing its natural frequencies and modal characteristics. Therefore, the investigation of functional coatings from both material and structural perspectives is important for developing more durable and dynamically stable timber structures ([10], [18], [23]).

Silicon dioxide (SiO_2) is a mineral material that has attracted considerable attention in engineering applications because of its high hardness, chemical stability, thermal resistance, and favorable surface characteristics. SiO_2-based materials have been utilized in protective coatings, composites, construction materials, and surface modification technologies where enhanced mechanical and environmental performance is required. At the surface of a timber element, a SiO_2-based coating may provide additional protection against moisture penetration, surface deterioration, and environmental exposure while potentially contributing to the stiffness and hardness of the treated region. The relatively lightweight nature of thin SiO_2 coatings also makes this approach potentially attractive for timber structures where an increase in structural mass should be minimized. These characteristics suggest that SiO_2 surface treatment may have potential not only for improving the durability of GLT members but also for modifying the modal behavior of complete timber structural systems ([24], [25], [26]).

The application of SiO_2 coating to GLT pole shed systems may therefore provide a multifunctional approach in which surface protection and structural performance enhancement are considered simultaneously. In a lightweight timber system, improvements in member stiffness can affect the global stiffness-to-mass relationship and consequently alter the natural frequencies of the structure. An increase in natural frequency may shift the structural response away from dominant excitation ranges associated with environmental dynamic loads and may reduce the potential for unfavorable resonance conditions. However, the effectiveness of such a coating cannot be evaluated solely through material-level properties. Its influence on the modal response of the entire structural system must also be examined by considering the geometry, member arrangement, support conditions, mass distribution, and interaction among structural components ([6], [10], [18]).

Although numerous studies have investigated the structural strengthening of timber members using mechanical reinforcement, adhesive techniques, fiber-reinforced polymers, and composite materials, relatively fewer studies have focused on the influence of mineral-based surface coatings on the global modal behavior of complete GLT structures. In particular, the numerical assessment of SiO_2-coated GLT pole sheds remains comparatively limited. The relationship between surface coating, structural stiffness, natural frequencies, and mode shapes requires further investigation to determine whether such a treatment can produce meaningful improvements at the structural-system level. This research gap is especially relevant for lightweight timber structures subjected to dynamic environmental actions, where relatively modest changes in stiffness can significantly influence vibration characteristics and serviceability ([19], [23], [28]).

The finite element method (FEM) provides an effective numerical framework for investigating the structural response of engineered timber systems with complex geometries and material configurations. Through appropriate representation of structural members, material properties, support conditions, connections, and coating layers, finite element simulations can be used to evaluate both the original and modified configurations under consistent modelling conditions. In particular, modal analysis enables the determination of natural frequencies and corresponding mode shapes, providing essential information about the dynamic characteristics of a structure. Numerical comparison of these parameters before and after surface treatment can therefore be used to quantify the influence of SiO_2 coating on the vibration behavior of a GLT pole shed ([15], [17], [27]).

Modal behavior is one of the fundamental indicators of the dynamic performance of a structure because natural frequencies reflect the relationship between structural stiffness and mass, while mode shapes describe the spatial distribution of deformation associated with each vibration mode. When the frequency content of an external dynamic load approaches one of the natural frequencies of a structure, resonance may occur and produce amplified vibration and deformation. This phenomenon can be particularly significant in lightweight timber systems because of their relatively low mass and potentially flexible structural configuration. Consequently, a modification that increases structural stiffness without introducing a substantial increase in mass may provide a favorable change in the dynamic characteristics of the system. For GLT pole sheds, evaluating modal parameters before and after SiO_2 treatment can provide a quantitative basis for determining the effectiveness of the proposed surface modification ([6], [10], [21]).

In this study, the modal behavior of a SiO_2-coated glued laminated timber pole shed is numerically assessed using the finite element method. Two three-dimensional structural models, representing the uncoated GLT pole shed and the SiO_2-coated GLT pole shed, are developed under identical geometric, boundary, loading, and modelling conditions to enable a direct comparison of their dynamic characteristics. The influence of the SiO_2 coating on the modal response of the structural system is investigated by examining the natural frequencies and corresponding mode shapes. Particular attention is given to the changes in the first several vibration modes, which provide important information regarding the global dynamic characteristics and stiffness behavior of the pole shed. The obtained numerical results are intended to clarify the potential contribution of SiO_2 surface treatment to the modal performance of GLT pole shed structures and to provide a numerical basis for evaluating mineral-based coating technologies as lightweight approaches for improving the dynamic stability, durability, and structural performance of engineered timber systems.

II. MATERIALS AND METHOD

A numerical modal analysis was conducted to assess the effect of SiO_2 coating on the modal behaviour and dynamic characteristics of a glued laminated timber (GLT) pole shed. For this purpose, two three-dimensional finite element models with the same geometric configuration were established, consisting of an existing GLT pole shed and a SiO_2-coated GLT pole shed. In order to ensure that the differences in the numerical results originated solely from the applied surface treatment, the structural geometry, dimensions and arrangement of the GLT members, support conditions, material properties of the timber components, and other modelling parameters were maintained consistently in both models. Thus, the SiO_2 coating was defined as the sole modification between the two structural configurations. The finite element simulations were performed using the SAP2000 program. The modal characteristics of the two models were comparatively examined in terms of their natural frequencies and corresponding mode shapes to determine the influence of the SiO_2 surface treatment on the dynamic response of the GLT pole shed. This comparative modelling strategy allowed the contribution of the SiO_2 coating to the modal behaviour and dynamic performance of the timber structure to be systematically evaluated.

A. Description of GLT Pole Shed

The GLT pole shed model is a one-storey timber structure with four equal spans (5 m) in the x and six equal spans (5 m) in the y directions and The shed has an eave height of 3 m and a ridge height of 4 m. The shed area is 600 square meter. Column cross-sections are 35x35 cm, beam cross-sections are 25x25 cm, and roof slab thickness is 15 cm. The finite element model of the GLT structure was created using the SAP2000 program. The finite element model of the GLT pole shed is given in Figure 1. Mechanical properties of GLT are given in Table 1, where L, T, and R denote the longitudinal, tangential, and radial directions, respectively.

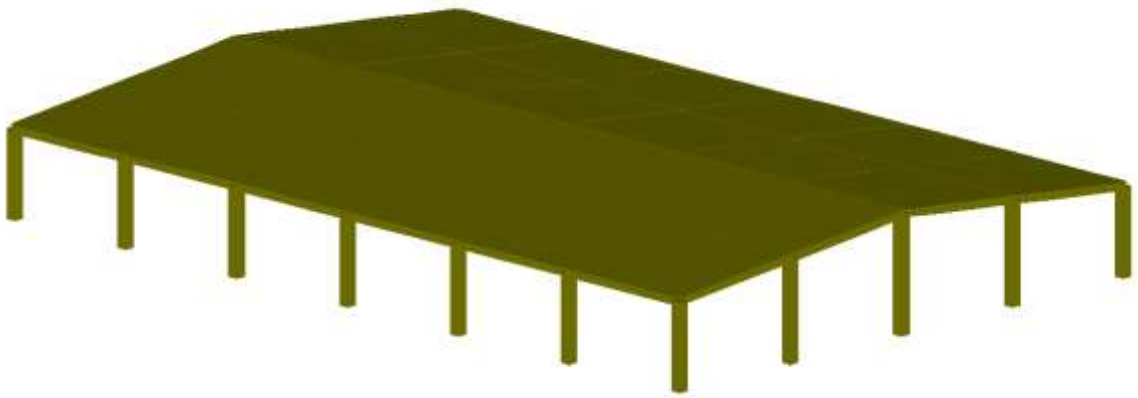


Fig. 1 3D Finite element model of the GLT pole shed

Table 1. Mechanical properties of GLT

Material	**Modulus of Elasticity (kN/m^2)**	**Poisson's Ratio**	**Weight per Unit Volume (kN/m^3)**
GLT	$11.5x10^6$ (L)	0.40 (LR)	4.8
	$0.46x10^6$ (T)	0.40 (LT)	
	$0.46x10^6$ (R)	0.35 (RT)	

L: Longitudinal, T: Tangential, R: Radial

B. Description of Coated GLT Pole Shed

The columns of the existing GLT pole shed are coated with 5 mm thick SiO_2. Thus, a coated model was created. The finite element model of the coated GLT pole shed is given in Figure 2. Mechanical properties of SiO_2 are given Table 2.

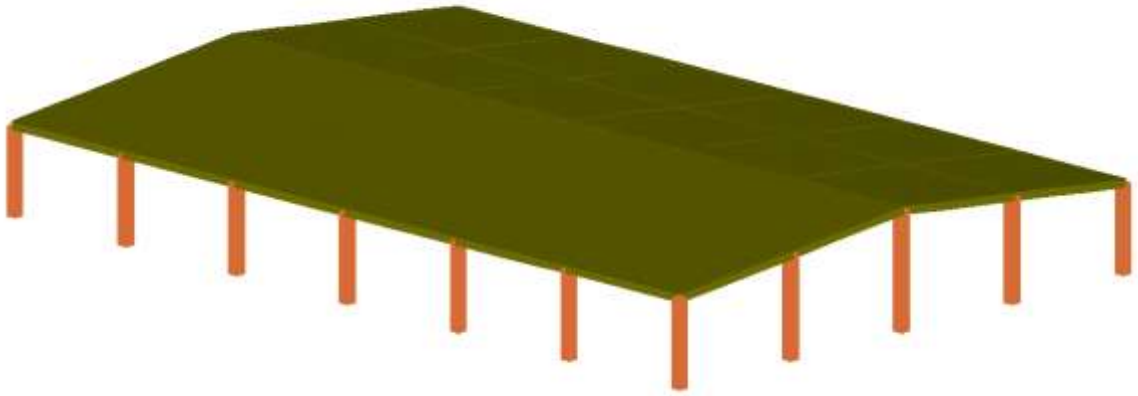


Fig. 2 3D Finite element model of the coated GLT pole shed

Table 2. Mechanical properties of SiO_2

Material	**Modulus of Elasticity (kN/m^2)**	**Poisson's Ratio**	**Weight per Unit Volume (kN/m^3)**
SiO_2	$75x10^6$	0.17	22

III. RESULTS

A. Modal Analysis Results of GLT Pole Shed

The modal characteristics of the GLT pole shed were investigated through a free-vibration modal analysis to assess its dynamic response. The first five vibration modes were considered to characterize the dominant dynamic behavior of the structure and to identify the principal deformation patterns that may influence its seismic performance. The calculated natural frequencies and corresponding mode shapes obtained from the finite element model provide valuable information regarding the global stiffness characteristics and dynamic behavior of the shed. Figures 3-7 present the first five vibration modes along with their corresponding natural frequencies, while the structural characteristics and significance of each mode are discussed in detail.

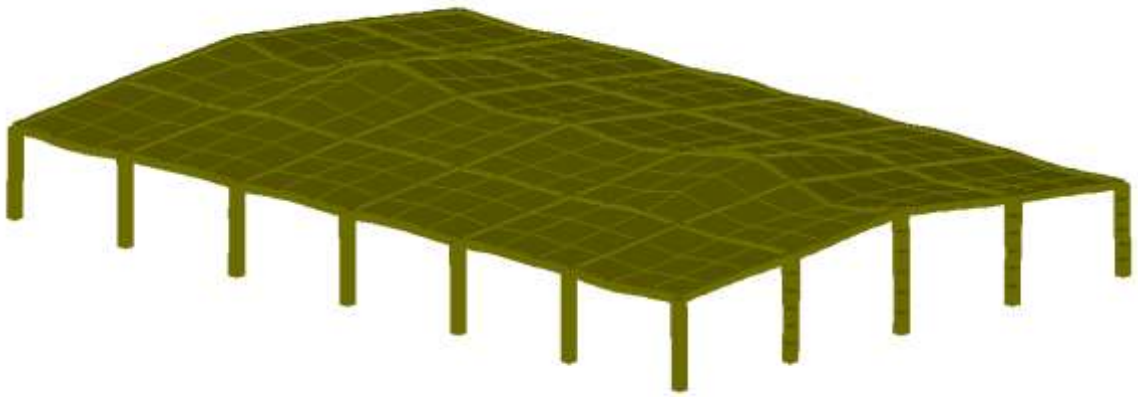


Fig. 3 1st Mode shape (Frequency value = 2.183 Hz)

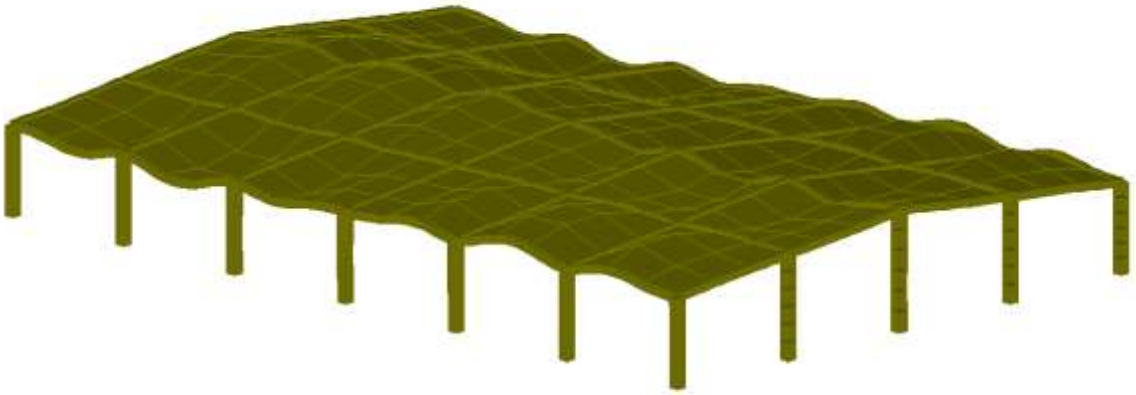


Fig. 4 2nd Mode shape (Frequency value = 2.958 Hz)

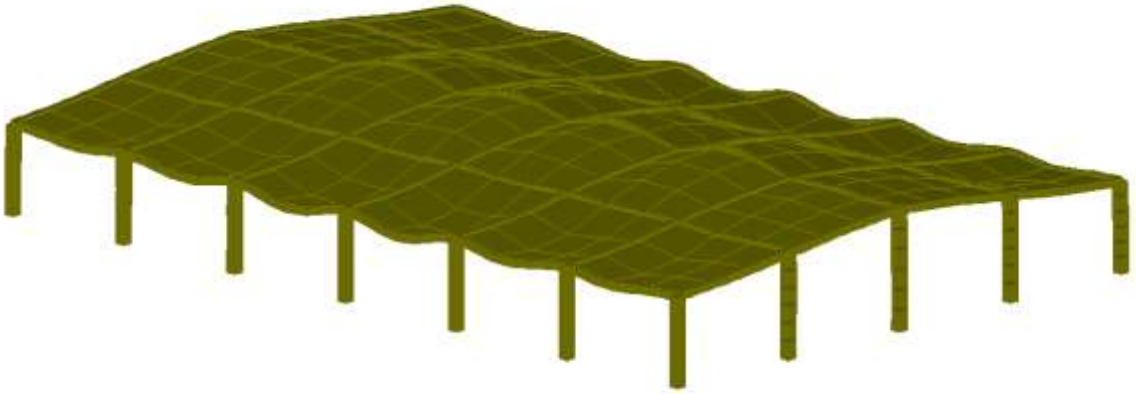


Fig. 5 3rd Mode shape (Frequency value = 4.371 Hz)

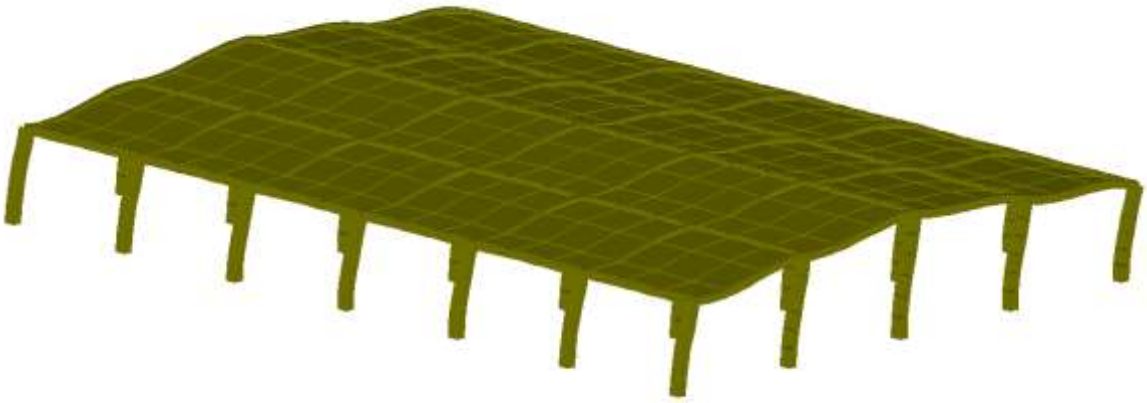


Fig. 6 4th Mode shape (Frequency value = 5.620 Hz)

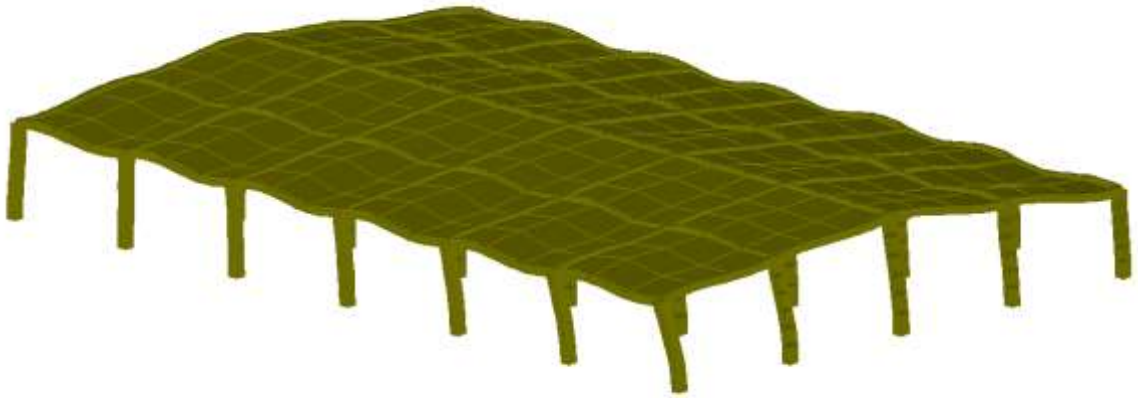


Fig. 7 5th Mode shape (Frequency value = 6.032 Hz)

B. Modal Analysis Results of Coated GLT Pole Shed

A modal analysis was performed on the coated GLT pole shed to investigate how the applied coating affects the dynamic properties of the structure. The first five vibration modes were examined to identify the dominant deformation patterns and global dynamic characteristics of the shed under free-vibration conditions. The addition of the coating modifies the stiffness distribution of the structural members, thereby affecting the resulting natural frequencies and corresponding mode shapes. The modal parameters obtained from the finite element analysis provide a basis for assessing the effectiveness of the coating in enhancing the dynamic performance and structural stability of the GLT pole shed. The first five natural frequencies and their associated mode shapes are illustrated in Figures 8-12.

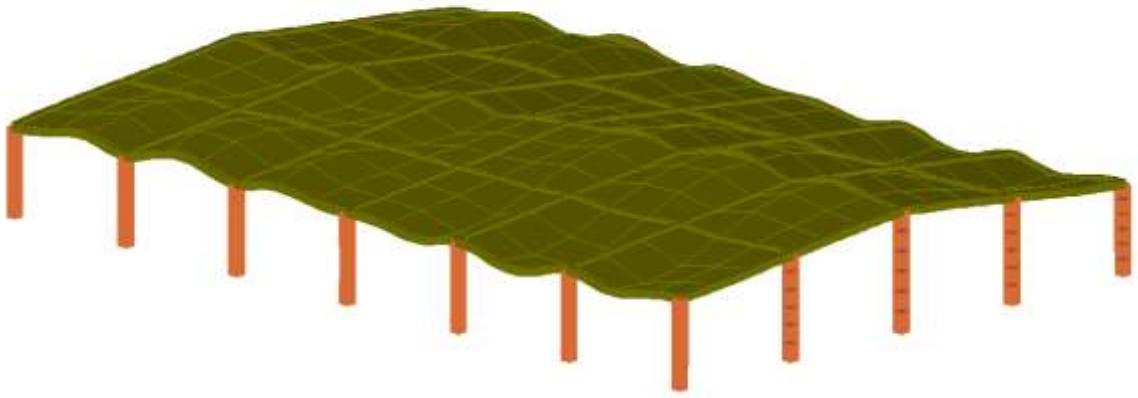


Fig. 8 1st Mode shape (Frequency value = 3.064 Hz)



Fig. 9 2nd Mode shape (Frequency value = 4.856 Hz)

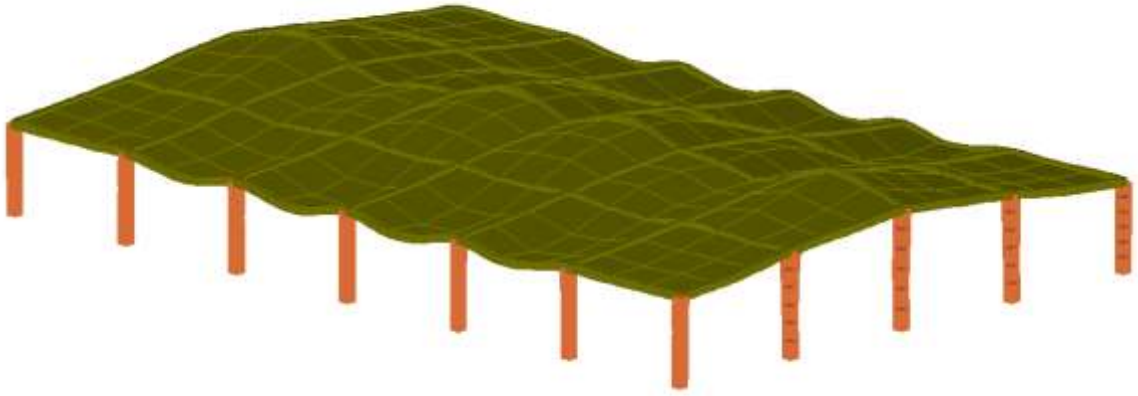


Fig. 10 3rd Mode shape (Frequency value = 5.033 Hz)

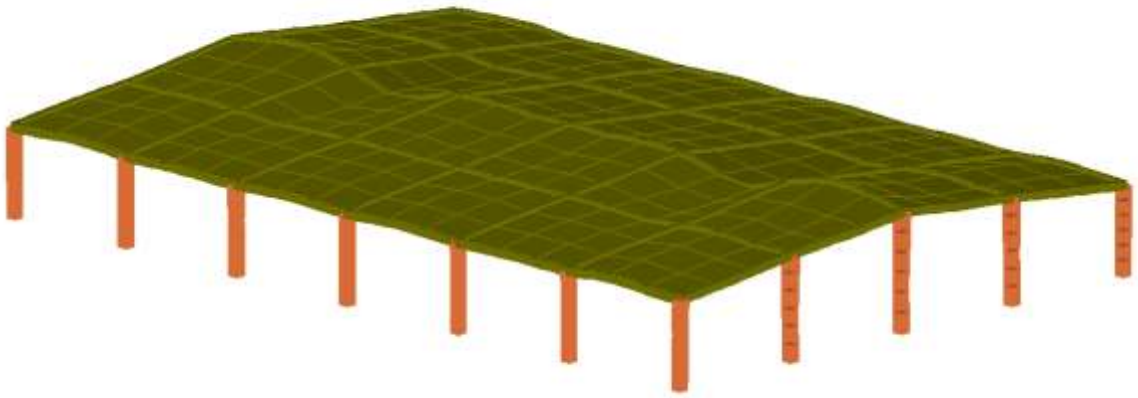


Fig. 11 4th Mode shape (Frequency value = 6.027 Hz)

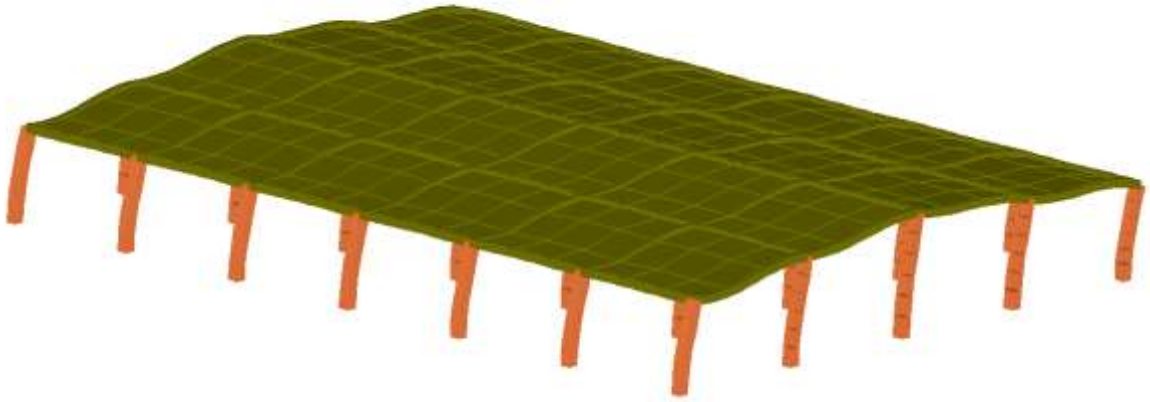


Fig. 12 5th Mode shape (Frequency value = 6.205 Hz)

IV. DISCUSSION

Table 3 compares the natural frequencies calculated for the uncoated and coated GLT pole shed models. This comparison provides a clear basis for evaluating the influence of the coating on the dynamic characteristics of the structure. The differences between the modal frequencies of the two models indicate the extent to which the coating modifies the overall stiffness and vibration behavior of the shed. The variations observed in the natural frequencies consequently demonstrate the changes in the structural dynamic response and offer an assessment of the coating's contribution to improving the overall dynamic performance of the structure.

Table 3. Comparison of frequency values

Mode	**1**	**2**	**3**	**4**	**5**
Existing	2.183	2.958	4.371	5.620	6.032
Coated	3.064	4.856	5.033	6.027	6.205
Difference (Hz)	0.881	1.898	0.662	0.407	0.173
Difference (%)	40.357	64.165	15.145	7.242	2.868

V. CONCLUSION

In this study, the influence of a 5 mm thick SiO_2 coating on the dynamic characteristics of a glued laminated timber (GLT) pole shed was numerically investigated using the finite element method. Two three-dimensional finite element models, representing the existing and SiO_2-coated GLT pole shed configurations, were developed under identical geometric, material, support, and modelling conditions. Modal analysis was subsequently performed to evaluate the changes in the natural frequencies and vibration characteristics resulting from the application of the SiO_2 coating.

The numerical results demonstrated that the SiO_2 coating increased the natural frequencies of all five investigated vibration modes. For the existing GLT pole shed, the first five natural frequencies were calculated as 2.183, 2.958, 4.371, 5.620, and 6.032 Hz, respectively. Following the application of the SiO_2 coating, these frequencies increased to 3.064, 4.856, 5.033, 6.027, and 6.205 Hz. The corresponding increases in natural frequency were approximately 40.36%, 64.17%, 15.15%, 7.24%, and 2.87%, respectively. The greatest improvement was observed in the second vibration mode, while the first mode also exhibited a substantial increase in frequency.

The increase in natural frequencies indicates that the SiO_2 coating modified the effective stiffness characteristics of the GLT pole shed. Since the geometry, boundary conditions, and timber properties were kept unchanged between the two models, the observed differences in modal parameters can primarily be attributed to the applied coating layer. The results therefore demonstrate that even a

relatively thin SiO_2 surface treatment can produce a measurable change in the global dynamic characteristics of the timber structural system.

The modification of the modal frequencies may also contribute to improved dynamic stability by changing the relationship between the structural frequencies and the frequency content of external dynamic excitation. In particular, the considerable increases observed in the first two modes suggest that the coating may be effective in shifting the dominant vibration characteristics of the GLT pole shed. However, the potential reduction in resonance risk depends on the characteristics of the actual external excitation and should therefore be evaluated for specific seismic or environmental loading conditions.

Overall, the finite element results indicate that SiO_2 coating represents a promising lightweight surface-treatment approach for enhancing the modal performance of GLT pole shed structures. The proposed technique provides an alternative to conventional strengthening methods that may require additional structural components or significant modifications to the existing system.

The findings provide a numerical basis for further investigation of SiO_2-based surface treatments as a potential approach for improving the dynamic performance and stability of lightweight engineered timber structures.

REFERENCES

[1] S. Tuhta, "GFRP Retrofitting effect on the dynamic characteristics of model steel structure," *Steel Compos. Struct.*, vol. 28, no. 2, pp. 223-231, July 2018.

[2] Q. Xu, H. Zou, Z. Wang, Y. He, L. Qi, and J. Wang, "Modal testing and analysis of high-rise laminated timber building," *BioResources*, vol. 19, no. 4, pp. 9616-9630, Oct. 2024.

[3] G. Dill-Langer, R. Nieberle, and A. Hänsel, "Glued laminated Robinia hardwood timber for structural use," *BioResources*, vol. 20, no. 2, pp. 3848-3865, Apr. 2025.

[4] M. A. Hammoud, S. Tuhta, and F. Günday, "Determination of modal parameters of reinforced concrete tunnel retrofitted with CFRP using finite element method," *Int. J. Innov. Eng. Res. Technol.*, vol. 9, no. 4, pp. 10-18, Apr. 2022.

[5] D. Shi, X. Hu, J. Zhang, and H. Du, "Shear stiffness of notched connectors in glue laminated timber-concrete composite beams under fire conditions," *BioResources*, vol. 17, no. 4, pp. 6397-6410, Sep. 2022.

[6] S. Tuhta, "Comparative analytical modal analysis of LVL shear-walled structure retrofitted with alumina (Al_2O_3) nanocoating exposed to earthquake effect," *Buildings*, vol. 16, no. 6, pp. 1-21, June 2026.

[7] C. Choi, E. Kojima, K.-J. Kim, M. Yamasaki, Y. Sasaki, and S.-G. Kang, "Analysis of mechanical properties of cross-laminated timber (CLT) with plywood using Korean larch," *BioResources*, vol. 13, no. 2, pp. 2715-2726, Feb. 2018.

[8] F. Bakkour, S. Tuhta, and F. Günday, "Determination of modal parameters of reinforced concrete box culvert retrofitted with GFRP using finite element method," *Int. J. Innov. Eng. Res. Technol.*, vol. 9, no. 4, pp. 1-9, Apr. 2022.

[9] S. Tuhta, "The determination of effect of TiO_2 on dynamic behavior of scaled WPC warehouse by OMA," *Adv. Nano Res.*, vol. 12, no. 1, pp. 65-72, Jan. 2022.

[10] N. Sunar and H. Uysal, "Dynamic characteristics of CFRP-strengthened wooden beams: experimental and numerical study," *Struct. Eng. Mech.*, vol. 89, no. 3, pp. 323-334, Feb. 2024.

[11] S. Tuhta, "The determination of effect of TiO_2 on dynamic behavior of scaled concrete structure by OMA," *Adv. Nano Res.*, vol. 11, no. 6, pp. 641-648, Dec. 2021.

[12] A. A. Kasımzade and S. Tuhta, "OMA of model steel structure retrofitted with CFRP using earthquake simulator," *Earthq. Struct.*, vol. 12, no. 6, pp. 689-697, June 2017.

[13] S. Tuhta, F. Günday, and M. O. Warayth, "The effect of GFRP steel silo on modal parameters using finite element method," *Int. J. Innov. Eng. Res. Technol.*, vol. 8, no. 7, pp. 41-46, July 2021.

[14] S.-J. Pang and J.-K. Oh, "Bending behavior of separable glued-laminated timber (GLT)-steel beam combined with inclined screws," *BioResources*, vol. 18, no. 2, pp. 3838-3855, Apr. 2023.

[15] S. Tuhta, F. Günday, and N. C. Pehlivan, "Investigation of CFRP retrofitting effect on masonry dome on bending moment using finite element method," *Int. J. Innov. Eng. Res. Technol.*, vol. 6, no. 6, pp. 18-22, June 2019.

[16] T. Catterou, Y. Sousseau, S. M. Elachachi, M. Chaplain, and C. Faye, "Non-linear dynamic analysis of timber frame structure with bolted-fastener connections," *Vibration*, vol. 7, no. 4, pp. 1156-1170, Nov. 2024.

[17] S. Tuhta, F. Günday, and H. Aydın, "Dynamic analysis of model steel structures retrofitted with GFRP composites under microtremor vibration," *Int. J. Trend Sci. Res. Dev.*, vol. 3, no. 2, pp. 729-733, Feb. 2019.

[18] D. Buck, X. Wang, O. Hagman, and A. Gustafsson, "Bending properties of cross laminated timber (CLT) with a 45° alternating layer configuration," *BioResources*, vol. 11, no. 2, pp. 4633–4644, Apr. 2016.

[19] S. Tuhta, "OMA of GFRP retrofitted model concrete structure using ambient vibration," *Int. J. Trend Sci. Res. Dev.*, pp. 360-370, vol.3, no. 1, Nov. 2018.

[20] S. Yang, G. Choi, J. Kim, H. Lee, and S. Kang, "Comparison of mechanical properties according to the structural materials of lumber, GLT, CLT, and Ply-lam CLT," *BioResources*, vol. 18, no. 4, pp. 6971-6985, Aug. 2023.

[21] F. Günday, "OMA of RC industrial building retrofitted with CFRP using SSI," *Int. J. Adv. Eng. Res. Dev.*, vol. 5, no. 5, pp. 759-771, May. 2018.
[22] S. Tuhta and F. Günday, "Investigation of effect of using nanocoating on wooden sheds on dynamic parameters," *Wood Res.*, vol. 66, no. 6, pp. 1006-1014, Dec. 2021.
[23] A. C. İlçe, "Mechanical properties of laminated veneer lumber made from ash and red pine woods," *BioResources*, vol. 13, no. 4, pp. 8653-8661, Oct. 2018.
[24] S. Tuhta and F. Günday, "Investigation of steel chimney retrofitted with nanocoating under earthquake excitation using FEM," *Coatings*, vol. 15, no. 5, pp. 1-20, May 2025.
[25] F. Günday, "Investigation of the effect of coating light steel container houses with nano-TiO_2 on dynamic parameters using OMA," *Buildings*, vol. 15, no. 6, pp. 1-17, Mar. 2025.
[26] S. Tuhta and F. Günday, "Determination of the effect of TiO_2 on the dynamic behavior of scaled concrete chimney by OMA," *Mater. Tehnol.*, vol. 55, no. 3, pp. 459-466, May. 2021.
[27] F. Günday, "Structural health monitoring of steel garage model with stochastic subspace identification covariance variance analysis," *Buildings*, vol. 15, no. 6, pp. 1-18, Mar. 2025.
[28] F. B. T. Imamura, Y. Chen, L. Deng, and Y. H. Chui, "Dimensional changes of cross-laminated timber basement walls under ambient humidity variation," *BioResources*, vol. 20, no. 2, pp. 4568-4589, Apr. 2025.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Performance Evaluation of Lightweight 1D-CNN Architectures for Real-Time ECG Anomaly Detection in Edge Devices

Ahmet Akkaya[1*] and Mustafa Taşcı[2]

[1]*Department of Computer Technologies, Gönen Vocational School, Bandırma Onyedi Eylül University, 10200 Bandırma, Balıkesir, Türkiye (ORCID: 0000-0003-4836-2310)*

[2]*Department of Computer Technologies, Gönen Vocational School, Bandırma Onyedi Eylül University, 10200 Bandırma, Balıkesir, Türkiye (ORCID: 0000-0002-8073-8587)*

[*]*(aakkaya@bandirma.edu.tr) Email of the corresponding author*

Abstract – Continuous monitoring of electrocardiogram (ECG) signals is vital for the early detection of cardiovascular diseases, which remain the leading cause of mortality worldwide. However, embedding deep learning models into resource-constrained IoT and wearable devices is challenging due to their high computational, memory, and power demands. This paper proposes a lightweight 1-Dimensional Convolutional Neural Network (1D-CNN) specifically optimized for hardware-limited edge environments. The architecture consists of two convolutional blocks followed by max-pooling layers and a compact fully connected stage, requiring only 98,757 trainable parameters and approximately 385 KB of memory in 32-bit floating-point representation, which fits within the SRAM limits of typical microcontrollers. Using the pre-processed MIT-BIH Arrhythmia dataset, the model was evaluated across the five heartbeat categories defined by the AAMI EC57 standard. An ablation study compared the standard ReLU activation function against the modern Swish and Mish functions in terms of accuracy, macro F1-score, training time, inference latency, and convergence speed. Results show that all three activation functions achieved statistically comparable validation accuracies of approximately 97.90%. Mish enabled faster convergence and the highest macro F1-score, whereas ReLU delivered the lowest inference latency and the simplest hardware execution profile. Considering both diagnostic performance and instruction-level computational cost, ReLU offers the best balance of simplicity and precision, proving to be the most energy-efficient and practical choice for real-time, low-latency ECG anomaly detection on edge devices.

Keywords – ECG Classification, Lightweight Neural Networks, 1D-CNN, Edge AI, Wearable Healthcare, Activation Functions, MIT-BIH Dataset

I. INTRODUCTION

Cardiovascular diseases (CVDs) remain the leading cause of mortality globally, emphasizing the critical need for continuous and proactive heart monitoring [1]. The electrocardiogram (ECG) is the most widely used diagnostic tool for recording the electrical activity of the heart and detecting arrhythmias [2]. Traditionally, ECG analysis has been performed by cardiologists in clinical settings; however, manual interpretation is time-consuming and prone to human error, particularly during long-term monitoring [3].

In recent years, the integration of Artificial Intelligence (AI), specifically Deep Learning (DL), has revolutionized automated ECG analysis [4]. Deep neural networks, such as complex Convolutional Neural Networks (CNNs) [5], Long Short-Term Memory (LSTM) networks [6], and Transformers [7], have achieved remarkable accuracy in arrhythmia classification [8]. Despite their high performance, these deep architectures possess a significant drawback: they are computationally expensive [9], require substantial memory [10], and consume excessive power [11]. Consequently, deploying these heavy

models on modern wearable IoT devices, such as smartwatches and portable holters (Edge AI), poses a major technical challenge [12]. Standard microcontrollers (MCUs) used in these edge devices (e.g., ARM Cortex-M series) typically possess highly restricted Static Random-Access Memory (SRAM), often less than 512 KB, and operate on strict power budgets measured in milliwatts. Running floating-point intensive models on such hardware leads to rapid battery depletion and unacceptable thermal output.

To address this gap, there is a paradigm shift towards developing "lightweight" models that can operate locally on edge devices without relying on continuous cloud connectivity. Local inference ensures real-time processing, reduces latency, and protects patient data privacy. This paper proposes a highly optimized, lightweight 1-Dimensional Convolutional Neural Network (1D-CNN) designed explicitly for low-power edge devices. The main purpose and scope of this article is to design, implement, and rigorously benchmark a sub-100k parameter 1D-CNN architecture capable of real-time ECG anomaly detection on hardware-constrained edge devices.

Furthermore, this study aims to fill a specific gap in the literature regarding the true cost of non-linear transformations in Edge AI. While lightweight CNNs for ECG classification have been widely explored, the specific hardware-level execution costs of modern non-linear activation functions in these shallow models are often ignored. The novelty of this study lies in bridging the gap between software-level accuracy and hardware-level computational feasibility. The primary contribution of this study is twofold:

1) This study introduces a minimal-parameter 1D-CNN architecture capable of distinguishing five ECG anomaly classes with high precision, specifically constrained to fit within standard MCU memory limits.

2) A comparative analysis of different activation functions (ReLU, Swish, and Mish) is conducted to demonstrate that computationally heavy modern functions do not yield statistically significant accuracy improvements for 1D-ECG data, thereby validating ReLU as the optimal hardware-aware configuration for edge networks.

The remainder of this paper is organized as follows: Section II details the dataset and the proposed methodology. Section III presents the experimental results and the comparative analysis of activation functions. Section IV discusses the hardware implications and the limitations of the study. Finally, Section V concludes the study with future perspectives.

II. Materials and Method

A. Dataset Description

The proposed model was trained and evaluated using the pre-processed MIT-BIH Arrhythmia dataset, widely recognized in the biomedical engineering domain [13]. The dataset consists of 109,446 annotated ECG heartbeat signals, which have been cropped, down-sampled to 125 Hz, and zero padded to a fixed length of 187 time steps to ensure uniform input matrices. The 188th column of the dataset represents the ground-truth class label.

The heartbeats are categorized into five distinct classes based on the AAMI EC57 standard: Normal beat (N - Class 0), Supraventricular premature beat (S - Class 1), Premature ventricular contraction (V - Class 2), Fusion of ventricular and normal beat (F - Class 3), and Unclassifiable beat (Q - Class 4). The morphological diversity of the heartbeats classified in this study is illustrated in Fig. 1, where the distinct signal patterns for each category, ranging from normal rhythms to complex ventricular contractions, are presented.

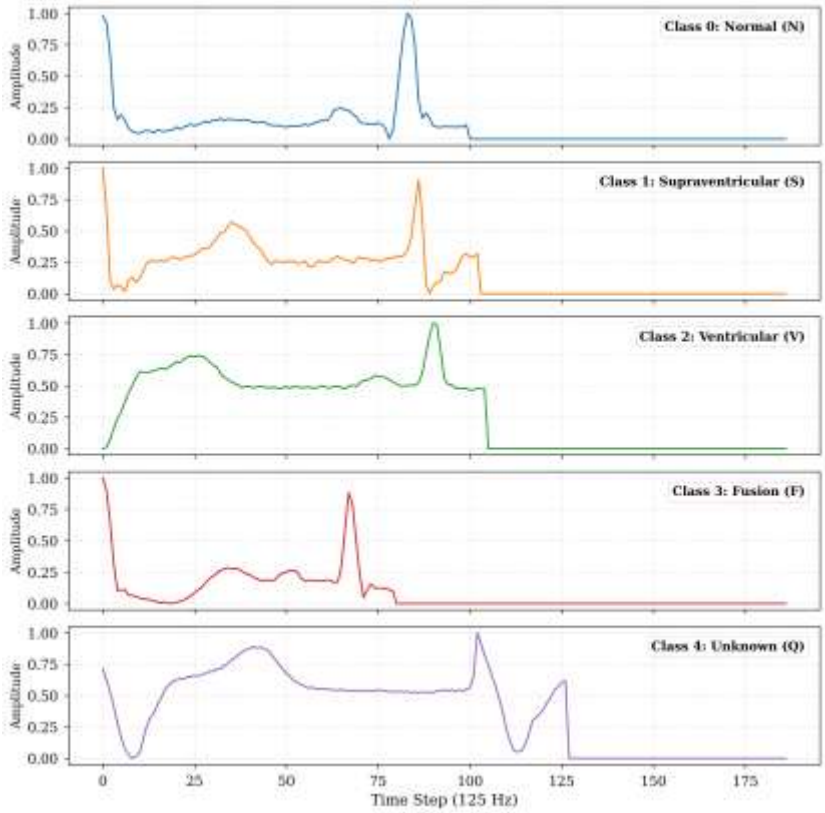


Fig. 1 Sample ECG signal segments from the MIT-BIH Arrhythmia dataset representing the five distinct heartbeat categories (N, S, V, F, and Q) used for model training and evaluation

To ensure robust evaluation and prevent class-ordering bias, the dataset was randomly shuffled. It was then partitioned using an 80/20 train-test split strategy, allocating 87,554 samples for the training phase and 21,892 samples for validation and performance benchmarking. To ensure the exact reproducibility of the results, a fixed random seed was utilized during the shuffling and splitting processes. All signal amplitudes were normalized between 0 and 1 to facilitate rapid weight optimization during gradient descent.

B. Proposed Lightweight 1D-CNN Architecture

Unlike traditional 2D-CNNs used in image processing, 1D-CNNs are specifically tailored for sequential time-series data like ECG signals. To ensure compatibility with hardware-constrained devices, the proposed network was intentionally kept shallow.

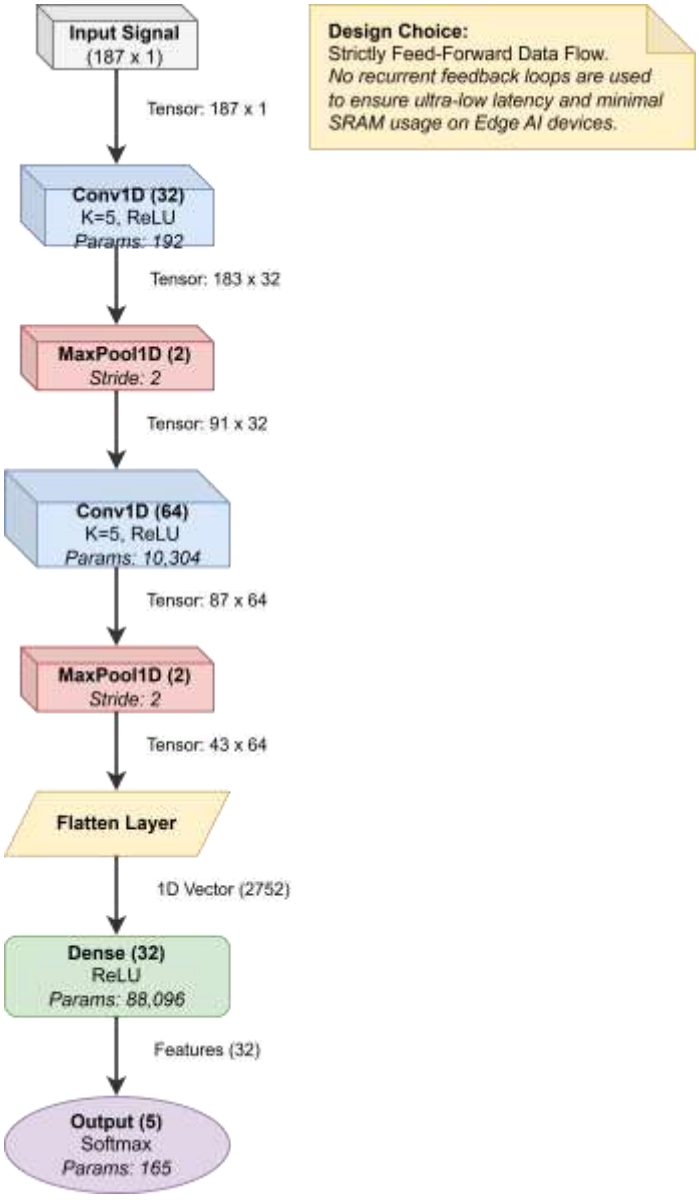


Fig. 2 Proposed lightweight 1D-CNN architecture

As illustrated in Fig. 2, the architecture takes normalized sequence input of shape (187, 1) and propagates through the following sequential blocks:

Convolutional blocks: Two 1D-convolutional layers with 32 and 64 filters respectively. Both layers utilize a kernel size of 5 (K = 5) to capture local temporal morphological features of the ECG signal, such as the P-wave and QRS-complex. The first layer transforms the input into a (183, 32) feature map, while the second produces an (87, 64) tensor.

Pooling layers: Each convolutional layer is immediately followed by a MaxPooling1D layer (pool size = 2, stride = 2) to down-sample the spatial dimensions by half, effectively reducing the computational load and providing translational invariance.

Fully connected layer: The extracted high-level features are flattened into a 1D vector of size (2752) and passed through a Dense hidden layer consisting of 32 neurons.

Output layer: A final Dense layer with 5 neurons using the Softmax activation function outputs the normalized probability distribution across the five ECG categories.

From a hardware perspective, this configuration requires exactly 98,757 trainable parameters. In a standard 32-bit floating-point (FP32) representation, the entire model weights occupy approximately 385 KB of memory, fitting comfortably within the SRAM limits of typical IoT microcontrollers without requiring external flash memory.

C. Hyperparameter Configuration and Activation Functions

To evaluate the impact of non-linear transformations on model performance and convergence speed, three different activation functions were tested in the hidden layers:

ReLU (Rectified Linear Unit): The standard, computationally cheapest function, $f(x) = \max(0, x)$.

Swish: A smooth, non-monotonic function defined as $f(x) = x \cdot \sigma(x)$, which tends to work well in deeper networks.

Mish: A self-regularized non-monotonic function defined as $f(x) = x \cdot \tanh(\ln(1 + e^x))$.

The models were compiled using the Adam optimizer with a learning rate of 0.001. A standard batch size of 32 was selected to balance memory constraints and gradient stability. The Sparse Categorical Crossentropy loss function was utilized since the target labels were provided as integers. To prevent overfitting and optimize training time, an Early Stopping callback was implemented with a patience of 5 epochs, monitoring the validation loss. The maximum epoch limit was set to 50.

III. RESULTS

A. Comparative Analysis of Activation Functions

The primary objective of the experiments was to evaluate how different activation functions affect the accuracy and convergence speed of a lightweight 1D-CNN architecture. Table 1 summarizes the performance metrics of the models trained with ReLU, Swish, and Mish activation functions.

Table 1. Performance comparison of activation functions

Activation Function	Test Accuracy	F1-Score (Macro)	Training Time (s)	Inf. Latency (ms)	Conv. Epoch
ReLU	97.90 %	0.8904	70.67	0.1643	26
Swish	97.85 %	0.8880	52.85	0.1710	26
Mish	97.86 %	0.8941	38.10	0.1748	18

Contrary to the initial hypothesis that modern activation functions would strictly outperform the standard ReLU in accuracy, the results indicate that all three functions yielded statistically similar peak accuracies (ranging from 97.85% to 97.90%). While ReLU achieved a fractionally higher test accuracy of 97.90%, its true advantage lies not in absolute diagnostic superiority, but in its lowest inference latency (0.1643 ms). While the Mish activation function demonstrated significantly faster convergence, triggering the early stopping mechanism at the 18th epoch compared to the 26 epochs required by ReLU, it slightly underperformed in absolute accuracy (97.86%). However, Mish provided the highest F1-Score (0.8941), suggesting superior performance in class-wise sensitivity. For highly constrained edge AI deployment, the mathematical simplicity of ReLU proves most optimal due to its superior inference speed and accuracy, whereas Mish offers a viable alternative for rapid training scenarios.

B. Training Dynamics and Class-Wise Performance

To thoroughly understand the behavior of the networks, the training and validation dynamics were meticulously analyzed. The comparative validation curves (Fig. 3) clearly illustrate the learning dynamics of each activation function.

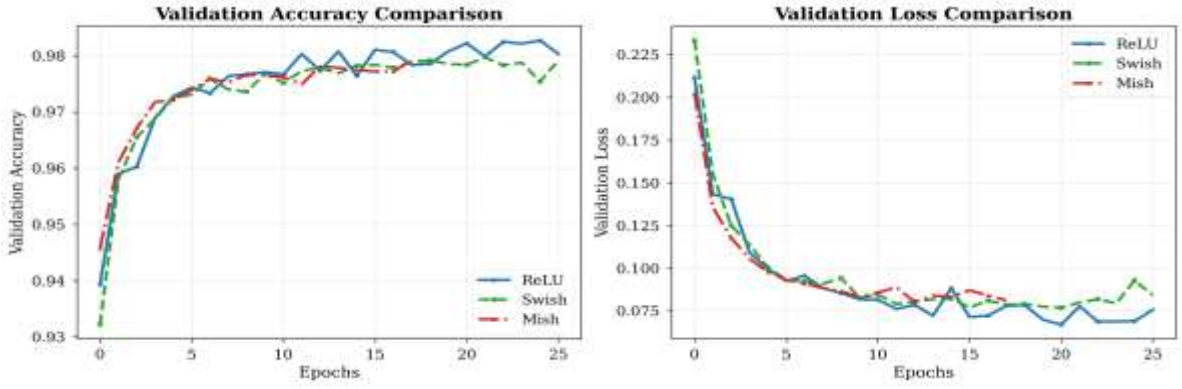


Fig. 3 Training curves

As observed in Fig. 3, Mish rapidly minimized the loss in the initial epochs due to its smooth gradient flow and self-regularization properties, which prevent dying neurons. However, ReLU ultimately converged to the highest accuracy and the lowest absolute validation loss before the early stopping mechanism halted the training.

Furthermore, the diagnostic capability of the best-performing model (ReLU) across different ECG categories is illustrated in the confusion matrix (Fig. 4).

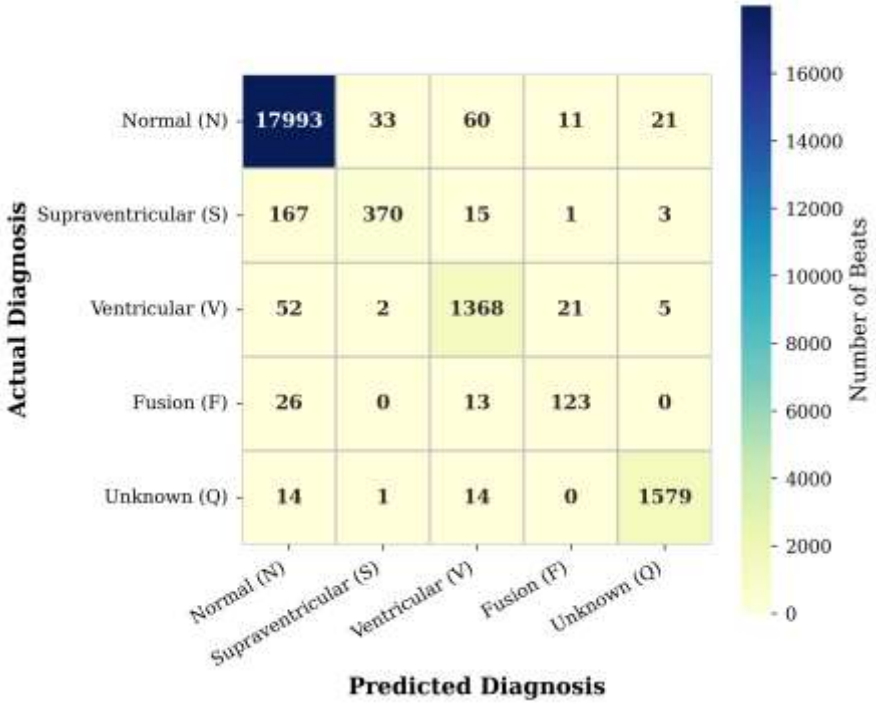


Fig. 4 Confusion matrix

The network showed exceptional precision in identifying Normal (Class 0) and Premature Ventricular Contractions (Class 2), which are the most critical for immediate clinical alerts. Minor misclassifications were primarily observed between the Fusion beats (Class 3) and Normal beats, which is a common challenge in the MIT-BIH dataset due to severe class imbalance and morphological similarities. Detailed class-wise precision, recall, and F1-scores for the best-performing ReLU model are presented in Table 2.

Table 2. Class-wise performance metrics (best model - ReLU)

Class ID	Precision	Recall	F1-Score
0 (N) - Normal	0.98	0.99	0.99
1 (S) - Supraventricular	0.92	0.85	0.88
2 (V) - Ventricular	0.96	0.95	0.96
3 (F) - Fusion	0.88	0.78	0.83
4 (Q) - Unknown	0.99	0.98	0.99

IV. DISCUSSION

Beyond diagnostic accuracy, the hardware implementation cost of the compared activation functions dictates their viability in Edge AI. While a physical deployment on a specific silicon chip is beyond the scope of this preliminary software-architecture study, a theoretical computational complexity analysis reveals critical insights. At the Instruction Set Architecture (ISA) level of an MCU, calculating a ReLU function max(0, x) requires a single conditional branch or a basic Arithmetic Logic Unit (ALU) instruction (approx. 1 to 2 clock cycles). In stark contrast, calculating Swish or Mish involves evaluating exponential terms (e^x) and hyperbolic tangents (tanh). Without a dedicated Floating-Point Unit (FPU), microcontrollers must rely on complex Taylor series expansions or memory-heavy Look-Up Tables (LUTs) to approximate these exponentials, consuming hundreds of clock cycles per neuron. Based on the architectural dimensions, the proposed ReLU model requires exactly 2.02 MegaFLOPs (1.01

MegaMACs) per inference, which is well within the real-time execution capabilities of a standard 72 MHz Cortex-M3 processor. Therefore, despite Mish's faster training phase, ReLU's negligible mathematical complexity makes it the most pragmatic and energy-efficient option for continuous, low-power inference on edge nodes.

This study utilized the widely adopted Kaggle split of the MIT-BIH dataset to establish a baseline for hardware-friendly models. It is important to acknowledge that this specific data configuration contains inherent subject-level data leakage, which typically inflates absolute accuracy metrics. However, since the primary focus of this paper is evaluating the relative computational efficiency and performance of activation functions in a constrained architecture, rather than claiming state-of-the-art clinical accuracy, this dataset served as a valid benchmarking environment.

V. CONCLUSION

In this study, a highly optimized, lightweight 1D-CNN architecture was proposed for real-time ECG anomaly detection, specifically targeting the hardware limitations of modern edge and wearable IoT devices. By keeping the network shallow and utilizing 1D convolutions, the computational footprint was drastically reduced to approximately 385 KB (under 100k parameters), fitting easily into standard microcontrollers while maintaining high diagnostic reliability.

A systematic ablation study on activation functions revealed that although all functions achieved comparable accuracy (~97.90%), the traditional ReLU function outperforms the newer, mathematically heavier Swish and Mish functions in this specific lightweight configuration due to its hardware execution profile rather than pure diagnostic superiority. Although Mish provided faster training convergence and a marginally better macro F1-Score, ReLU remains the ultimate pragmatic choice for low-power inference in edge computing due to its negligible execution overhead at the hardware instruction level.

Future research will focus on validating the proposed lightweight architecture on strictly patient-independent datasets and deploying the model onto physical microcontrollers (e.g., ARM Cortex-M or Raspberry Pi) to measure real-time power consumption in milliwatts. Furthermore, the efficiency of the training phase can be further enhanced by integrating adaptive operator selection mechanisms [14], which have shown promising results in optimization problems by dynamically adjusting system parameters. Additionally, to improve the reliability of diagnostic decision-support systems, we aim to investigate the integration of Retrieval-Augmented Generation (RAG) architectures [15] for cross-referencing ECG anomalies with medical protocols and regulatory guidelines, building upon established frameworks for reliable information access in safety-critical autonomous systems.

REFERENCES

[1] R. Jain, R. Singh, S. Yamini, and M. K. Das, 'Fragmented ECG as a Risk Marker in Cardiovascular Diseases', Curr. Cardiol. Rev., vol. 10, no. 3, pp. 277–286, Aug. 2014, doi: 10.2174/1573403X10666140514103451.

[2] A. Ameen, I. E. Fattoh, T. Abd El-Hafeez, and K. Ahmed, 'Advances in ECG and PCG-based cardiovascular disease classification: a review of deep learning and machine learning methods', J. Big Data, vol. 11, no. 1, p. 159, Nov. 2024, doi: 10.1186/s40537-024-01011-7.

[3] U. R. Acharya, H. Fujita, O. S. Lih, M. Adam, J. H. Tan, and C. K. Chua, 'Automated detection of coronary artery disease using different durations of ECG segments with convolutional neural network', Knowl.-Based Syst., vol. 132, pp. 62–71, Sep. 2017, doi: 10.1016/j.knosys.2017.06.003.

[4] X. Liu, H. Wang, Z. Li, and L. Qin, 'Deep learning in ECG diagnosis: A review', Knowl.-Based Syst., vol. 227, p. 107187, Sep. 2021, doi: 10.1016/j.knosys.2021.107187.

[5] R. Donida Labati, E. Muñoz, V. Piuri, R. Sassi, and F. Scotti, 'Deep-ECG: Convolutional Neural Networks for ECG biometric recognition', Pattern Recognit. Lett., vol. 126, pp. 78–85, Sep. 2019, doi: 10.1016/j.patrec.2018.03.028.

[6] H. Zacarias, J. A. L. Marques, V. Felizardo, M. Pourvahab, and N. M. Garcia, 'ECG Forecasting System Based on Long Short-Term Memory', Bioengineering, vol. 11, no. 1, p. 89, Jan. 2024, doi: 10.3390/bioengineering11010089.

[7] B. Behinaein, A. Bhatti, D. Rodenburg, P. Hungler, and A. Etemad, 'A Transformer Architecture for Stress Detection from ECG', in Proceedings of the 2021 ACM International Symposium on Wearable Computers, in ISWC '21. New York, NY, USA: Association for Computing Machinery, Sep. 2021, pp. 132–134. doi: 10.1145/3460421.3480427.

[8] J. Huang, B. Chen, B. Yao, and W. He, 'ECG Arrhythmia Classification Using STFT-Based Spectrogram and Convolutional Neural Network', IEEE Access, vol. 7, pp. 92871–92880, 2019, doi: 10.1109/ACCESS.2019.2928017.

[9] A. Burguera, 'Fast QRS Detection and ECG Compression Based on Signal Structural Analysis', IEEE J. Biomed. Health Inform., vol. 23, no. 1, pp. 123–131, Jan. 2019, doi: 10.1109/JBHI.2018.2792404.

[10] M. Merino-Monge, C. Lebrato-Vázquez, J. A. Castro-García, G. Sánchez-Antón, and A. J. Molina-Cantero, 'Computational and Memory Efficiency in Heartbeat Rate Detection: A Review of ECG and PPG Techniques', Sensors, vol. 26, no. 8, p. 2409, Jan. 2026, doi: 10.3390/s26082409.
[11] M. A. Scrugli, G. Leone, P. Busia, L. Raffo, and P. Meloni, 'All-Spiking ECG Analysis for Arrhythmia Classification on Low-Power FPGA', IEEE Sens. J., vol. 26, no. 7, pp. 11372–11386, Apr. 2026, doi: 10.1109/JSEN.2026.3663715.
[12] J. Gao et al., 'A Transferable Hybrid Convolutional-Mamba Network for Cross-Population Emotion Recognition From Wearable ECG', IEEE Internet Things J., vol. 13, no. 7, pp. 15124–15141, Apr. 2026, doi: 10.1109/JIOT.2026.3651586.
[13] G. B. Moody and R. G. Mark, 'The impact of the MIT-BIH Arrhythmia Database', IEEE Eng. Med. Biol. Mag., vol. 20, no. 3, pp. 45–50, May 2001, doi: 10.1109/51.932724.
[14] A. Akkaya and C. Közkurt, 'An effective approach for adaptive operator selection and comparison for PSO algorithm', Clust. Comput., vol. 28, no. 6, p. 368, Jun. 2025, doi: 10.1007/s10586-024-05024-y.
[15] A. Akkaya, 'Reliable Information Access in Intelligent Transportation Systems and Autonomous Driving Protocols: An Experimental Evaluation of the RAG Architecture's Regulatory Analysis', Akıllı Ulaşım Sist. Ve Uygulamaları Derg., vol. 9, no. 1, pp. 22–41, Mar. 2026, doi: 10.51513/jitsa.1848120.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Metric Choice Reorders System Rankings in Turkish Offensive-Language Detection: A 25-Configuration Study

Ahmet Akkaya[1*] and Muhammet Arucu[2]

[1]*Department of Computer Technologies, Gönen Vocational School, Bandırma Onyedi Eylül University, 10200 Bandırma, Balıkesir, Türkiye (ORCID: 0000-0003-4836-2310)*

[2]*Department of Computer Technologies, Gönen Vocational School, Bandırma Onyedi Eylül University, 10200 Bandırma, Balıkesir, Türkiye (ORCID: 0000-0001-7620-9044)*

[*]*(aakkaya@bandirma.edu.tr) Email of the corresponding author*

Abstract – Offensive-language detection is a class-imbalanced problem, yet accuracy is still the figure most often reported for it. This study asks whether the metric a researcher reports changes which system that researcher would select. Twenty-five classifier configurations, formed by crossing five linear classifiers with three feature representations and two class-weighting settings and adding a majority-class reference, were trained on the 31,756-tweet training split of the OffensEval-2020 Turkish corpus and scored on the official 3,528-tweet test split, in which 20.29 % of items are offensive. Each configuration was measured on accuracy, balanced accuracy, macro-averaged F1, F1 on the offensive class, Matthews correlation and Cohen's κ, and the rankings these metrics induce were compared with Kendall's τ. Accuracy ranks the configurations almost independently of their ability to find offensive content: $\tau = 0.14$ against balanced accuracy and 0.22 against F1 on the offensive class. The configuration that maximises balanced accuracy ranks nineteenth of twenty-five on accuracy, and three configurations exceed the majority-class baseline on accuracy by 1.6 to 3.2 points while recovering at most 18 % of offensive tweets. Class weighting lowers accuracy and raises minority-class recall in all nine matched pairs, at an average cost of 3.25 accuracy points for 22.70 points of recall, but it improves Matthews correlation in only three of the nine. Accuracy and Matthews correlation selected the same configuration, which the methodological literature would not predict. All computation runs on a CPU in 1.3 minutes.

Keywords – class imbalance; evaluation metrics; offensive language detection; Turkish natural language processing; Matthews correlation coefficient; accuracy paradox

I. INTRODUCTION

Offensive posts are a minority of what people write, so offensive-language detection is a class-imbalanced problem. In the Turkish corpus used here they are 19.31 % of the training split. A classifier that assigns the majority label to every input therefore reaches 79.71 % accuracy on the test split while detecting nothing. Accuracy is nevertheless the number that headlines a large share of applied papers and student theses in this area.

The methodological literature has argued against this practice and proposed alternatives: Matthews correlation as a single summary [3], [4], balanced accuracy, macro-averaged F1, precision–recall analysis in place of ROC analysis on skewed data [12], and selection of the measure to match the task [7]. Most of that work argues from constructed examples, using synthetic confusion matrices chosen to expose a particular metric's blind spot. Fewer studies measure, on a real corpus and a real set of candidate systems, how far the choice of metric moves the decision a researcher would make.

This paper reports such a measurement for one task. The design is small: twenty-five configurations, six metrics, one public corpus, no hyper-parameter search, no GPU, and a total runtime of 1.3 minutes. A

grid this cheap can be re-run by any group working on Turkish text, so the result does not depend on access to unusual resources.

Three questions are answered with numbers. Do the six metrics rank the same twenty-five systems in the same order, and if not, how far apart are the orderings? If one group selects its system on accuracy and another selects on a minority-sensitive measure, how far apart are the two systems in the fraction of offensive content they actually catch? And do the standard prescriptions, namely enabling class weighting and reporting Matthews correlation in place of accuracy, behave on this corpus as the methodological literature predicts?

The answer to the third question is partly negative. Class weighting buys large gains in minority-class recall, but it improves macro-averaged F1 in only six of nine matched pairs and Matthews correlation in only three, so enabling it as a default is not supported here. Accuracy and Matthews correlation, often presented as the clearest case of metric disagreement, selected the same configuration. Both outcomes are reported here.

II. RELATED WORK

The Turkish resource used here is the offensive-language corpus of Çöltekin [1], 36,232 tweets sampled between April 2018 and September 2019, which supplied the Turkish track of SemEval-2020 Task 12 [2]. That work reports 77.3 % F1 on the binary offensive decision with a linear support vector machine, and the shared-task overview [2] collects the systems submitted against it.

On measurement, Chicco and Jurman [3] argue that Matthews correlation should replace accuracy and F1 because it responds to all four cells of the confusion matrix, and a later study [4] extends the argument to balanced accuracy and related corrections. Sokolova and Lapalme [7] set out what each measure is invariant to. Cohen's κ [5] and the interpretive bands of Landis and Koch [6] come from content analysis, where the question of whether a coder can be trusted has been asked far longer than in machine learning. Saito and Rehmsmeier [12] show that the choice of curve, and not only of scalar, matters under skew.

On imbalance, the systematic study of Japkowicz and Stephen [9] and the survey of He and Garcia [8] find that the damage depends jointly on the imbalance ratio, the sample size and the class separability. On inference, Dietterich [10] identifies McNemar's test as the appropriate paired test when two classifiers are compared on one common test set, and Demšar [11] sets out the multiple-comparison discipline for larger grids. Both are followed here.

The present study continues the first author's earlier work on the accuracy paradox under extreme class imbalance in thin-structure segmentation on railway imagery [14], moving the same question from pixels to text, and it uses the approach to paired testing and effect sizes adopted in [15].

III. MATERIALS AND METHOD

A. Dataset

All experiments use the OffensEval-2020 Turkish partition of the corpus of Çöltekin [1], [2], distributed by the corpus author under a Creative Commons Attribution licence. Rows with an empty text field, and rows whose label is neither offensive nor not offensive, were discarded. The training split then holds 31,756 tweets of which 19.31 % are offensive, and the official test split holds 3,528 tweets of which 20.29 % are offensive. Mentions and URLs are already replaced by placeholder tokens in the released files. No re-identification is attempted and no new data are collected. Apart from lower-casing inside the vectorisers, no normalisation is applied, so deasciified spellings, character elongations and emoji reach the classifier as an operational pipeline would receive them.

B. Configuration grid

The grid crosses five classifiers, three feature representations and two class-weighting settings. The classifiers are logistic regression, a linear support vector classifier, stochastic gradient descent with the modified Huber loss, multinomial naive Bayes and complement naive Bayes. The feature sets are word unigrams and bigrams with a minimum document frequency of three, word-boundary character n-grams of length three to five with a minimum document frequency of five, and the union of the two; all use sublinear term-frequency scaling with inverse document frequency. Class weighting is either off or set to

the inverse class frequency, which the two naive Bayes variants do not accept, leaving 24 trained configurations plus a majority-class reference, 25 in total. No hyper-parameter is tuned, because the object of study is the evaluation protocol and not the attainable ceiling.

C. Metrics

Six metrics are recorded for every configuration: accuracy, balanced accuracy, macro-averaged F1, F1 on the offensive class, Matthews correlation [3] and Cohen's κ [5]. Precision and recall on the offensive class are recorded separately, since recall is the quantity on which a content-moderation deployment is judged. Each metric induces a ranking of the 25 configurations, and the agreement between two rankings is summarised by Kendall's τ [13]. A value near one means the two metrics lead to the same decision; a value near zero means the two are unrelated for selection purposes.

D. Statistical treatment

Point estimates for the selected configurations are reported with percentile bootstrap 95 % confidence intervals over 1,000 resamples of the test items. Paired comparisons between two configurations on the same test set use the exact binomial form of McNemar's test on the discordant pairs of the correct/incorrect outcome, which Dietterich [10] identifies as appropriate for this situation. The discordant counts are reported next to the p-values so that the evidence can be read directly. Families of comparisons are Holm-corrected. Differences between classifiers across the six feature-by-weight blocks are assessed with the Friedman test [11].

E. Reproducibility

The corpus is obtained from the address given in [1]. All random seeds are fixed and no model weights are downloaded. The complete study executes in 1.3 minutes on the CPU of a free notebook instance, so the grid can be reproduced without specialised hardware.

IV. RESULTS

A. Overall results

Table 1 reports every configuration on every metric. Two properties of the table matter for what follows. First, accuracy has little range: across all 25 configurations it spans 7.94 points, from 79.59 % to 87.53 %, while recall on the offensive class spans 70.67 points and Matthews correlation spans 0.570. Second, 23 of the 24 trained configurations exceed the accuracy of a classifier that never predicts the offensive label, so exceeding that baseline on accuracy says very little.

Table 1. All 25 configurations on the official OffensEval-2020 Turkish test split (n = 3,528, 20.29 % offensive). cw = class weighting. Bal. acc = balanced accuracy. R(OFF) = recall on the offensive class. Shaded: the accuracy-optimal and the balanced-accuracy-optimal configuration.

Classifier	Features	cw	Acc.	Bal. acc.	Macro-F1	F1(OFF)	R(OFF)	MCC
Majority class	none	n/a	0.7971	0.5000	0.4435	0.0000	0.0000	0.0000
LogisticRegression	word 1-2	none	0.8399	0.6190	0.6466	0.3852	0.2472	0.4110
LogisticRegression	word 1-2	balanced	0.7959	0.7335	0.7116	0.5556	0.6285	0.4303
LinearSVC	word 1-2	none	0.8455	0.6840	0.7139	0.5198	0.4120	0.4574
LinearSVC	word 1-2	balanced	0.8098	0.7146	0.7110	0.5420	0.5545	0.4222
SGD (mod. Huber)	word 1-2	none	0.8484	0.6561	0.6912	0.4708	0.3324	0.4535
SGD (mod. Huber)	word 1-2	balanced	0.8081	0.7256	0.7158	0.5537	0.5866	0.4331
MultinomialNB	word 1-2	none	0.8132	0.5398	0.5213	0.1475	0.0796	0.2540
ComplementNB	word 1-2	none	0.8291	0.7023	0.7162	0.5372	0.4888	0.4369
LogisticRegression	char 3-5	none	0.8560	0.6603	0.6995	0.4827	0.3310	0.4885
LogisticRegression	char 3-5	balanced	0.8271	0.7822	0.7558	0.6239	0.7067	0.5197
LinearSVC	char 3-5	none	0.8713	0.7397	0.7715	0.6204	0.5182	0.5624
LinearSVC	char 3-5	balanced	0.8382	0.7657	0.7575	0.6175	0.6439	0.5158
SGD (mod. Huber)	char 3-5	none	0.8753	0.7172	0.7606	0.5948	0.4511	0.5702
SGD (mod. Huber)	char 3-5	balanced	0.8401	0.7821	0.7662	0.6347	0.6844	0.5354
MultinomialNB	char 3-5	none	0.8189	0.5595	0.5568	0.2160	0.1229	0.2898
ComplementNB	char 3-5	none	0.8370	0.6464	0.6760	0.4476	0.3254	0.4071
LogisticRegression	word+char	none	0.8620	0.6880	0.7282	0.5375	0.3953	0.5146
LogisticRegression	word+char	balanced	0.8370	0.7765	0.7613	0.6269	0.6746	0.5254
LinearSVC	word+char	none	0.8637	0.7339	0.7615	0.6054	0.5154	0.5380
LinearSVC	word+char	balanced	0.8401	0.7472	0.7501	0.6000	0.5908	0.5002
SGD (mod. Huber)	word+char	none	0.8688	0.7173	0.7552	0.5884	0.4623	0.5459

Classifier	Features	cw	Acc.	Bal. acc.	Macro-F1	F1(OFF)	R(OFF)	MCC
SGD (mod. Huber)	word+char	balanced	0.8424	0.7663	0.7611	0.6218	0.6383	0.5226
MultinomialNB	word+char	none	0.8294	0.5885	0.6030	0.3032	0.1830	0.3549
ComplementNB	word+char	none	0.8450	0.6612	0.6942	0.4795	0.3520	0.4424

B. Agreement between metric rankings

Table 2 gives Kendall's τ between every pair of metric-induced rankings. The first row is the relevant one: accuracy agrees with balanced accuracy at $\tau = 0.14$ and with F1 on the offensive class at $\tau = 0.22$. For the purpose of choosing a system, accuracy and minority-class competence are close to unrelated on this corpus. At the other end, macro-averaged F1 and Cohen's κ agree at $\tau = 0.95$, and balanced accuracy and F1 on the offensive class at $\tau = 0.92$, so each of those pairs is close to one measurement reported twice.

Table 2. Kendall's τ between the rankings of the 25 configurations induced by each pair of metrics.

Metric	Accuracy	Bal. acc.	Macro-F1	F1(OFF)	MCC	Kappa
Accuracy	1.000	0.140	0.375	0.221	0.569	0.421
Bal. acc.	0.140	1.000	0.767	0.920	0.573	0.720
Macro-F1	0.375	0.767	1.000	0.847	0.793	0.953
F1(OFF)	0.221	0.920	0.847	1.000	0.653	0.800
MCC	0.569	0.573	0.793	0.653	1.000	0.840
Kappa	0.421	0.720	0.953	0.800	0.840	1.000

Table 3 restates this as a selection problem. Each metric nominates its own winner, and the nominee is then located in the accuracy ranking. Accuracy, macro-averaged F1, Matthews correlation and κ nominate one of two closely related configurations. Balanced accuracy nominates logistic regression on character n-grams with class weighting, which ranks nineteenth of twenty-five on accuracy, and F1 on the offensive class nominates a configuration ranking eleventh. A group that shortlists candidates by accuracy would retain neither.

Table 3. The winner under each metric, and where that winner sits in the accuracy and MCC rankings of all 25 configurations.

Metric	Winning configuration	Value	Acc. rank	MCC rank
Accuracy	SGD (mod. Huber) / char 3-5 / cw=none	0.8753	1	1
Balanced accuracy	LogisticRegression / char 3-5 / cw=balanced	0.7822	19	8
Macro-F1	LinearSVC / char 3-5 / cw=none	0.7715	2	2
F1(OFF)	SGD (mod. Huber) / char 3-5 / cw=balanced	0.6347	11	5
MCC	SGD (mod. Huber) / char 3-5 / cw=none	0.5702	1	1
Cohen's kappa	LinearSVC / char 3-5 / cw=none	0.5465	2	2

C. Consequences for system selection

Table 4 places the accuracy winner beside the balanced-accuracy winner. The accuracy-selected system is 4.82 accuracy points higher and recovers 25.56 points less of the offensive content, 45.11 % against 70.67 %. Counted as failures, the accuracy-selected system does not flag 55 % of offensive posts and the other does not flag 29 %. Accuracy separates the two by under five points.

Table 4. The accuracy-optimal and the balanced-accuracy-optimal configuration, side by side. P(OFF) and R(OFF) are precision and recall on the offensive class.

Selected by	Configuration	Acc.	Bal. acc.	Macro-F1	P(OFF)	R(OFF)	MCC
Accuracy	SGD (mod. Huber) / char 3-5 / cw=none	0.8753	0.7172	0.7606	0.8730	0.4511	0.5702
Balanced accuracy	LogisticRegression / char 3-5 / cw=balanced	0.8271	0.7822	0.7558	0.5585	0.7067	0.5197
difference	*same feature set, different classifier and weighting*	*−0.0482*	*+0.0650*	*−0.0048*	*−0.3145*	*+0.2556*	*−0.0505*

Figure 1 shows the same pattern across the grid. Accuracy occupies a narrow band on the horizontal axis while Matthews correlation is widely spread on the vertical axis, so configurations with nearly the same accuracy differ substantially in Matthews correlation. The three points at the bottom of the cloud are the multinomial naive Bayes configurations.

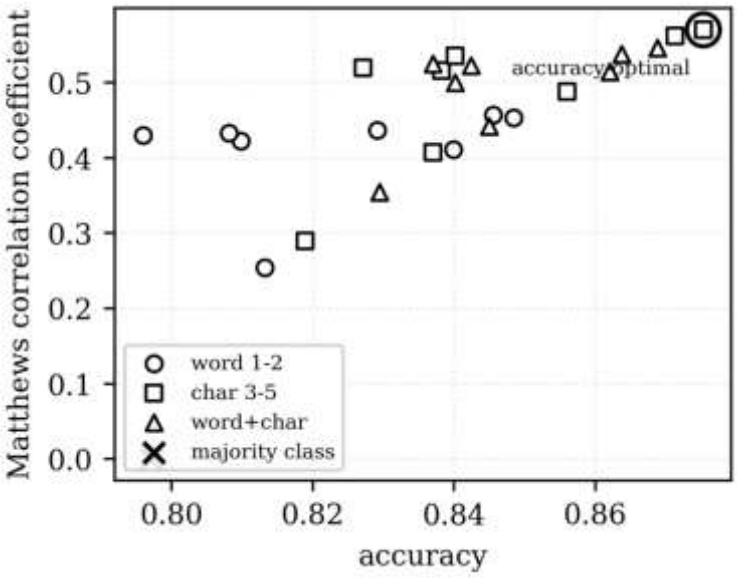


Fig. 1 Accuracy against Matthews correlation for all 25 configurations. Marker shape encodes the feature set; the cross is the majority-class reference; the circled point is the accuracy-optimal configuration.

Multinomial naive Bayes is the clearest single case. On word features it reaches 81.32 % accuracy, 1.61 points above the majority-class baseline, while recovering 7.96 % of offensive tweets. On the union features it reaches 82.94 %, 3.23 points above the baseline, at 18.30 % recall. All three of its configurations would be reported as beating the trivial baseline, and none of them is adequate for the task it was applied to.

D. Effect of class weighting

Nine configurations exist in matched pairs differing only in class weighting. The direction of the effect is the same in all nine. Class weighting lowers accuracy by 3.25 points on average, with a range of 2.36 to 4.40, and raises recall on the offensive class by 22.70 points on average, with a range of 7.54 to 38.13. Figure 2 shows the pairs.

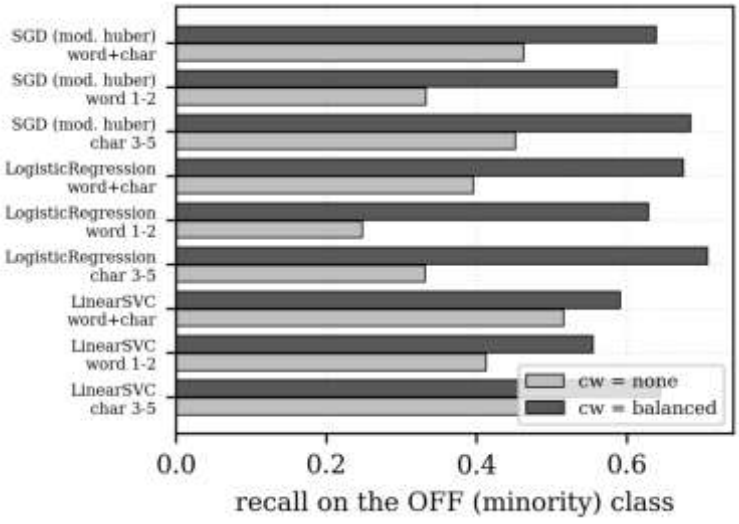


Fig. 2 Recall on the offensive class for the nine matched pairs, with class weighting off and set to inverse class frequency.

The summary metrics do not endorse the change unconditionally. Macro-averaged F1 improves in six of the nine pairs, by 1.80 points on average, and Matthews correlation improves in three, falling by 1.52 points on average. The common advice to enable class weighting by default is therefore not supported by these data. What they support is narrower: class weighting buys recall at the expense of precision and accuracy, and whether that exchange is worthwhile depends on the application.

E. Sensitivity to the decision threshold

Sweeping the decision threshold of a logistic regression on the union features separates the optima. Over the sampled grid, accuracy is highest near 0.425 at 86.59 %, macro-averaged F1 near 0.300 at 76.60 %, and Matthews correlation near 0.425 at 0.535. The conventional 0.500 is optimal for none of the three. Moving from 0.550 to 0.300 raises macro-averaged F1 from 71.70 % to 76.60 %, a gain of 4.90 points, which is larger than most of the differences between classifiers in Table 1, at a cost of 0.79 accuracy points. Comparisons of architectures at a fixed threshold of 0.5 therefore include a threshold effect.

Table 5. Threshold sweep for logistic regression on the union feature set (every fifth point of the evaluated grid).

Threshold on P(OFF)	Accuracy	Macro-F1	MCC
0.050	0.3654	0.3645	0.2136
0.175	0.7795	0.7218	0.4862
0.300	0.8535	0.7660	0.5330
0.425	0.8659	0.7499	0.5349
0.550	0.8620	0.7170	0.5150
0.675	0.8444	0.6507	0.4377
0.800	0.8265	0.5830	0.3417
0.925	0.8104	0.5084	0.2303

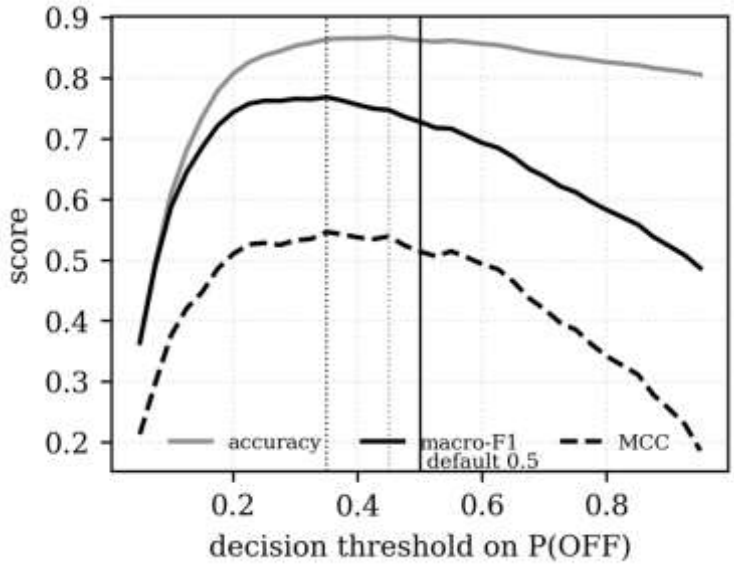


Fig. 3 Accuracy, macro-averaged F1 and Matthews correlation as a function of the decision threshold. The vertical line marks the conventional 0.5.

F. Secondary observations

Character n-grams outperform word n-grams throughout. The best macro-averaged F1 with word features is 71.62 %, against 77.15 % with character three- to five-grams, and the best Matthews correlation rises from 0.457 to 0.570. Concatenating the two representations does not improve on characters alone, at 76.15 % macro-averaged F1 and 0.546 Matthews correlation, so the word features add nothing the character features have not already captured. Turkish agglutinative morphology offers an explanation, since one lemma generates many surface forms that share substrings but not tokens. This is a replication of a known result rather than a new one.

The accuracy-optimal configuration reaches 0.8753 accuracy with a bootstrap 95 % interval of [0.8651, 0.8858], macro-averaged F1 0.7606 [0.7397, 0.7809] and Matthews correlation 0.5702 [0.5347, 0.6051]. Against the majority-class reference it wins 323 items and loses 47, giving an exact McNemar p of 9.0×10^{-52} and a Holm-adjusted p of 2.7×10^{-51}.

V. DISCUSSION

The main result is $\tau = 0.14$ between accuracy and balanced accuracy. On this corpus the ranking obtained from accuracy carries little information about which of a group's systems is best at the task. The familiar objection to accuracy is that it is inflated under skew, which is correct but less damaging, since an inflated number that preserves the ordering still supports the right decision. The measured problem is that accuracy does not preserve the ordering. The configuration a minority-sensitive measure would choose ranks nineteenth of twenty-five on accuracy, outside any plausible shortlist.

Two results ran against expectation. Accuracy and Matthews correlation, the usual example of metric disagreement, agreed at $\tau = 0.569$ and nominated the same configuration, so on this corpus a paper reporting Matthews correlation in place of accuracy would have reached the same conclusion. Class weighting, the standard first remedy for imbalance, improved Matthews correlation in three of nine matched pairs. Both outcomes are consistent with Japkowicz and Stephen [9], who report that the effects of imbalance depend on the particular combination of ratio, sample size and class separability. Both are also reasons to treat prescriptions derived from a single corpus with caution, including those the present results would support.

The recommendations these data support are limited. Report a minority-sensitive measure together with recall on the minority class, since these are what separate the systems here. Report accuracy next to the majority-class baseline, so that a reader can see how much of it comes free. Do not report macro-averaged F1 and Cohen's κ together, since at $\tau = 0.95$ they are close to one measurement. State the decision threshold, since it moves macro-averaged F1 further than the choice of classifier does.

The study covers one corpus, one language, one binary construct and one family of linear classifiers. The imbalance ratio is a moderate 1:4, and the results say nothing about the 1:100 regimes in which the paradox is sharper. No hyper-parameters were tuned, so the absolute scores fall below what a tuned system would reach, although the rank-correlation result concerns relations among configurations rather than their level. Transformer-based encoders were not evaluated, and whether a stronger model family compresses the disagreement between metrics by being uniformly better remains open. The gold labels are human judgements on a construct with genuine annotator disagreement, which bounds every number in Table 1.

VI. CONCLUSION

Twenty-five classifier configurations were evaluated on the official OffensEval-2020 Turkish test split under six metrics. Accuracy ranked them almost independently of their ability to detect offensive content, agreeing with balanced accuracy at Kendall's τ of 0.14. The configuration best by balanced accuracy ranked nineteenth of twenty-five by accuracy, and three configurations exceeded the majority-class baseline on accuracy while recovering at most 18 % of offensive tweets. Class weighting raised minority-class recall in all nine matched pairs but improved Matthews correlation in three, and the conventional 0.5 threshold was optimal under no metric. Accuracy and Matthews correlation selected the same configuration, contrary to the expectation set by the methodological literature.

Three extensions would sharpen these conclusions. Running the same grid at imbalance ratios beyond 1:20, obtained by subsampling the offensive class, would test whether the rank divergence grows as predicted. Adding a fine-tuned Turkish encoder would show whether a stronger model family reduces the disagreement between metrics. A replication on a second Turkish construct with a sharper definition would separate metric behaviour from annotator disagreement.

ACKNOWLEDGMENT

The experiments were carried out on a freely available CPU notebook instance. The corpus is redistributed by its author under a Creative Commons Attribution licence.

REFERENCES

[1] Ç. Çöltekin, “A corpus of Turkish offensive language on social media,” in *Proc. 12th Language Resources and Evaluation Conf. (LREC 2020)*, Marseille, France, 2020, pp. 6174–6184.

[2] M. Zampieri *et al.*, “SemEval-2020 Task 12: Multilingual offensive language identification in social media (OffensEval 2020),” in *Proc. 14th Workshop on Semantic Evaluation (SemEval-2020)*, Barcelona, Spain, 2020, pp. 1425–1447.
[3] D. Chicco and G. Jurman, “The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation,” *BMC Genomics*, vol. 21, no. 1, art. 6, 2020, doi: 10.1186/s12864-019-6413-7.
[4] D. Chicco, N. Tötsch, and G. Jurman, “The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation,” *BioData Mining*, vol. 14, no. 1, 2021, doi: 10.1186/s13040-021-00244-z.
[5] J. Cohen, “A coefficient of agreement for nominal scales,” *Educational and Psychological Measurement*, vol. 20, no. 1, pp. 37–46, 1960.
[6] J. R. Landis and G. G. Koch, “The measurement of observer agreement for categorical data,” *Biometrics*, vol. 33, no. 1, pp. 159–174, 1977.
[7] M. Sokolova and G. Lapalme, “A systematic analysis of performance measures for classification tasks,” *Information Processing & Management*, vol. 45, no. 4, pp. 427–437, 2009.
[8] H. He and E. A. Garcia, “Learning from imbalanced data,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 21, no. 9, pp. 1263–1284, 2009.
[9] N. Japkowicz and S. Stephen, “The class imbalance problem: A systematic study,” *Intelligent Data Analysis*, vol. 6, no. 5, pp. 429–449, 2002.
[10] T. G. Dietterich, “Approximate statistical tests for comparing supervised classification learning algorithms,” *Neural Computation*, vol. 10, no. 7, pp. 1895–1923, 1998.
[11] J. Demšar, “Statistical comparisons of classifiers over multiple data sets,” *Journal of Machine Learning Research*, vol. 7, pp. 1–30, 2006.
[12] T. Saito and M. Rehmsmeier, “The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets,” *PLoS ONE*, vol. 10, no. 3, art. e0118432, 2015.
[13] M. G. Kendall, “A new measure of rank correlation,” *Biometrika*, vol. 30, no. 1/2, pp. 81–93, 1938.
[14] A. Akkaya, “Aşırı sınıf dengesizliği altında ince doğrusal yapıların segmentasyonu: Demiryolu veri setlerinde doğruluk paradoksu ve mimari verimlilik,” *Demiryolu Mühendisliği*, pp. 1–24, 2026.
[15] A. Akkaya and C. Közkurt, “An effective approach for adaptive operator selection and comparison for PSO algorithm,” *Cluster Computing*, vol. 28, no. 6, art. 368, 2025.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

EFFECTS OF APRICOT KERNEL, BIOCHAR, HUMIC ACID AND POTASSIUM HUMATE ON SOIL C-N DYNAMICS, MICROBIAL BIOMASS AND CO_2 EMISSIONS

Asuman Büyükkılıç Yanardağ[1]*, İbrahim Halil Yanardağ[1], Abdullah Atum[1], Erdal Sakin[2], Hakan Yıldırım[3] and Miraç Kılıç[1]

[1] *Malatya Turgut Ozal University, Faculty of Agriculture, Department of Soil Science and Plant Nutrition, Battalgazi, Malatya, Turkey. *Corresponding author e-mail: asuman.yanardag@ozal.edu.tr*

[2] *Harran University, Faculty of Agriculture, Department of Soil Science and Plant Nutrition, Osmanbey, Şanlıurfa, Turkey*

[3] *Malatya Turgut Ozal University, Faculty of Agriculture, Department of Horticulture, Battalgazi, Malatya, Turkey.*

Abstract - This study evaluated the effects of apricot kernel residue, biochar, humic acid and potassium humate applications on soil chemical properties, carbon-nitrogen pools, microbial biomass and CO_2 emissions under controlled incubation conditions. The experiment consisted of nine treatments, including the control, with three replications; organic materials were incorporated into the soil at two doses. During incubation, soil pH, electrical conductivity (EC), lime, total nitrogen, organic matter, organic carbon, C:N ratio, soluble carbon (SC), microbial biomass carbon (MBC), soluble nitrogen (SN), microbial biomass nitrogen (MBN), instantaneous CO_2, cumulative CO_2 and qCO_2 were determined. Data were evaluated using one-way ANOVA, Tukey test, repeated-measures ANOVA, hierarchical clustering and principal component analysis (PCA). Potassium humate, particularly at the second dose, caused the strongest changes in pH, EC, organic carbon and MBC; however, it required careful dose management due to high pH and EC. Apricot kernel applications stimulated labile carbon mineralization by increasing CO_2 and cumulative CO_2 emissions. Biochar produced a lower-respiration and more stable carbon profile, whereas humic acid applications were more closely associated with soluble nitrogen and microbial nitrogen pools.

Keywords: Apricot kernel, biochar, humic acid, potassium humate, microbial biomass, CO_2 emission, PCA

1. Introduction

Soil health is an integrative concept that requires the joint assessment of many chemical and biological processes, including organic carbon accumulation, nutrient cycling, microbial activity, aggregate stability, water retention capacity and carbon use efficiency. Current approaches consider soil not only as a medium that supplies nutrients to plants, but also as an active biochemical system involved in carbon transformation and ecosystem services. Therefore, the success of organic amendments should be evaluated not only by their ability to increase organic matter content, but also by the extent to which carbon is directed into microbial biomass, soluble fractions, stable pools or lost as CO_2 (Lehmann et al., 2020; Tao et al., 2023).

The effects of organic amendments added to soil depend on the chemical composition of the material, its C:N ratio, decomposability, soluble carbon content and mineral/ionic load. Easily decomposable plant residues may increase microbial activity and CO_2 release in the short term, whereas more aromatic or condensed carbon forms may slow mineralization and contribute to longer-term carbon retention in soil. The C:N ratio is also one of the key indicators determining the direction of nitrogen mineralization or microbial immobilization; therefore, total organic carbon, total nitrogen, soluble carbon, microbial biomass and qCO_2 should be interpreted together (Flavel & Murphy, 2006; Smith et al., 2024; Zhang et al., 2022).

Biochar is one of the most prominent organic amendments for carbon stabilization due to its aromatic and more resistant carbon structure formed during pyrolysis. However, the effect of biochar in soil is not fixed;

feedstock type, pyrolysis temperature, ash content, pH, EC, surface area, application dose and the initial properties of soil determine the direction and magnitude of its response. Under some conditions, biochar may improve microbial habitat and increase organic carbon, but pH and EC changes should be carefully monitored, especially in calcareous or alkaline soils (Ippolito et al., 2020; Singh et al., 2022; Tomczyk et al., 2020; Xu et al., 2023).

Humic acid and potassium humate are widely used as humic-derived soil conditioners and biostimulant-like materials. Humic substances can influence organic carbon fractions, nutrient transformations and microbial processes. However, potassium humate, as the potassium salt form of humic substances, may alter not only the organic carbon pool but also soluble ion load, pH and EC. Therefore, microbial biomass development should be evaluated together with pH and salinity indicators when assessing humic materials (Ampong et al., 2022; Gümüş and Şeker, 2015; Ma et al., 2024; Rupngam et al., 2025).

Apricot kernel residue is a local and economically relevant agricultural by-product, particularly in regions where apricot production is concentrated. Recycling such fruit-processing residues into soil is important for circular economy and waste management; however, these materials may also increase short-term CO_2 emissions due to their labile carbon fractions. This study aimed to compare apricot kernel, biochar, humic acid and potassium humate under the same incubation conditions and to determine their different functional effects on soil carbon-nitrogen dynamics, microbial biomass and CO_2 emissions.

2. Material and Methods

The study was conducted under laboratory conditions using surface soil samples collected from the research and application area of Malatya Turgut Ozal University. Soil samples were composited by combining subsamples collected from different points; stones, plant residues and coarse root fragments were removed, and the samples were air-dried and passed through a 2-mm sieve.

Four different organic materials were used in the experiment: apricot kernel, biochar, humic acid and potassium humate. Each material was applied at two doses, and a total of nine treatments were established, including an untreated control. The incubation experiment was arranged in a completely randomized design with three replications. Equal amounts of soil were placed into each incubation container, and the relevant material doses were homogeneously mixed with the soil on a dry-weight basis. During incubation, soil moisture was maintained at a level that would not limit microbial activity, and weight losses were regularly checked and compensated with distilled water. Soil pH, electrical conductivity, lime, organic matter, organic carbon and total nitrogen were analyzed. To evaluate carbon-nitrogen transformations, soluble carbon, soluble nitrogen, microbial biomass carbon and microbial biomass nitrogen were determined on days 1, 3, 7, 15, 30, 60 and 120. Microbial biomass C and N were calculated using the fumigation-extraction approach. CO_2 emissions were monitored at specific intervals during incubation, and instantaneous CO_2 emission, cumulative CO_2 and the microbial metabolic quotient (qCO_2) were obtained.

Soil moisture was maintained throughout the incubation period at a level that did not limit microbial activity. SC, MBC, SN and MBN were measured on days 1, 3, 7, 15, 30, 60 and 120, while CO_2 emissions were monitored more frequently to calculate instantaneous CO_2, cumulative CO_2 and qCO_2. One-way ANOVA and Tukey test were used for day-based comparisons, whereas repeated-measures ANOVA was applied to time-dependent data. Hierarchical clustering and PCA were performed after z-score standardization to demonstrate the overall separation of treatments at the final sampling stage.

3. Results and Discussion

3.1. Soil reaction, EC and lime



Figure 1. Effects of organic material applications on soil pH, EC and lime content on days 1 and 120.

Treatment, day and treatment × day effects were significant for pH, EC and lime. The most distinct chemical separation was observed in potassium humate treatments. Potassium humate at the second dose reached approximately pH 10 and the highest EC value on day 120, clearly differing from the other treatments. In contrast, the lower lime values in the same treatment suggest that alkalinity may have originated not directly from carbonate/lime, but from the soluble potassium humate fraction and ionic load of the material. Apricot kernel and biochar applications had more limited effects on pH and EC.

3.2. Total nitrogen, organic matter, organic carbon and C:N

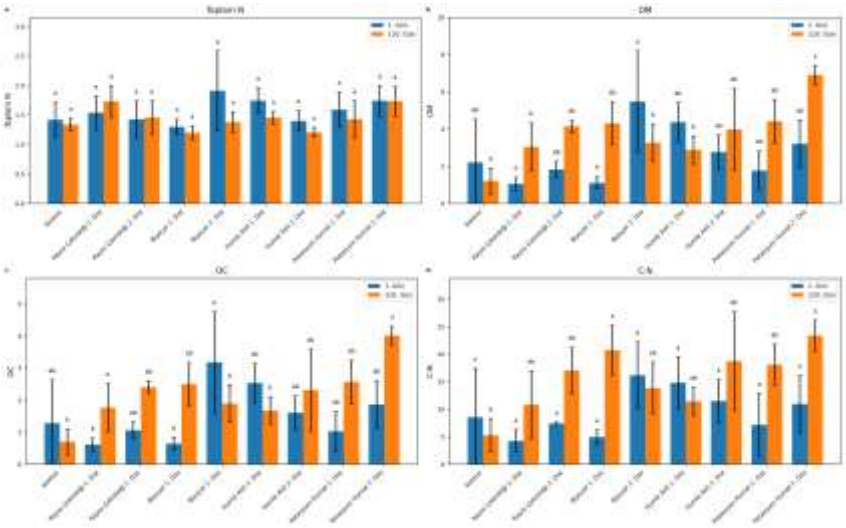


Figure 2. Changes in total N, organic matter, organic carbon and C:N ratio.

The total nitrogen pool remained more stable against the treatments, whereas organic matter, organic carbon and C:N ratio responded more sensitively to material type and incubation time. On day 1, biochar at the second dose was prominent in terms of organic matter and organic carbon, while on day 120, the highest organic matter and organic carbon values were determined in the second dose of potassium humate. The clearer differentiation of C:N ratio on day 120 indicates that organic materials modified the carbon and nitrogen balance over time.

3.3. Soluble C/N and microbial biomass pools

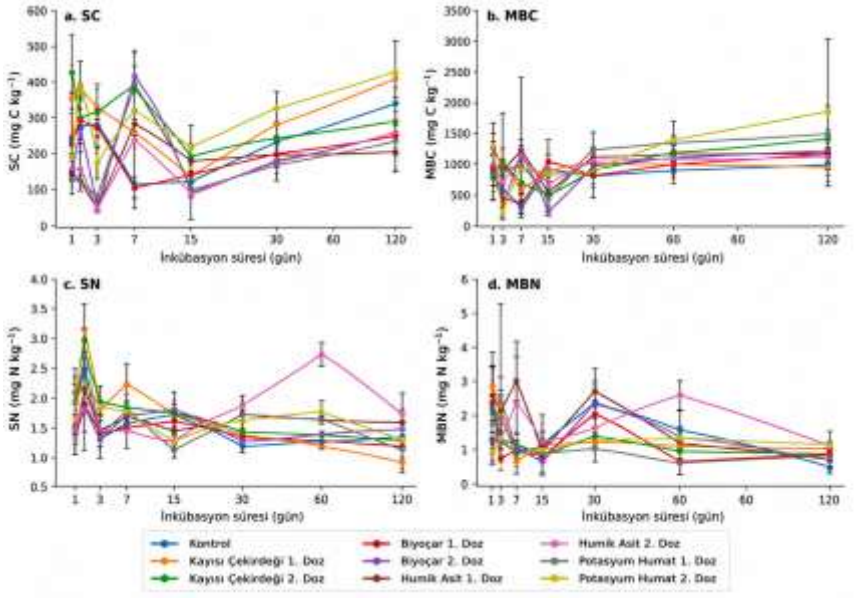


Figure 3. Changes in SC, MBC, SN and MBN during incubation.

Treatment and treatment × day effects were significant for SC, SN and MBN; for MBC, the treatment effect remained generally limited, although day and treatment × day interaction were significant. The SC pool fluctuated over time, and potassium humate at the second dose and apricot kernel at the first dose became prominent on day 120. Higher SN values in humic acid treatments on day 120 indicate that these materials may support the nitrogen pool in a more balanced manner.

3.4. CO_2 emission, cumulative CO_2 and qCO_2

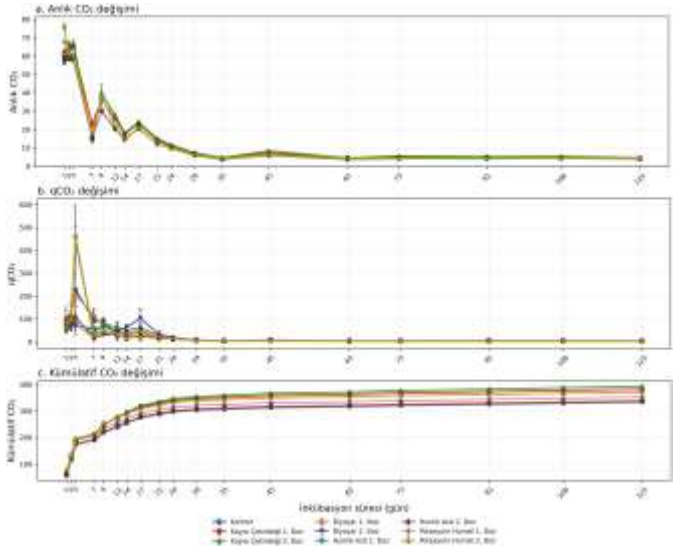


Figure 4. Changes in instantaneous CO_2, qCO_2 and cumulative CO_2.

The treatment effect on cumulative CO_2 emission was significant on all measurement days. At the end of day 125, the highest cumulative CO_2 emission was determined in the second dose of apricot kernel, while the first dose of apricot kernel was also included in the same high group. This result shows that apricot kernel residue stimulated microbial respiration and carbon mineralization by providing labile carbon. Biochar and humic acid applications produced a lower-respiration profile. The qCO_2 results indicated that CO_2 emissions should be interpreted together with MBC.

3.5. Multivariate evaluation

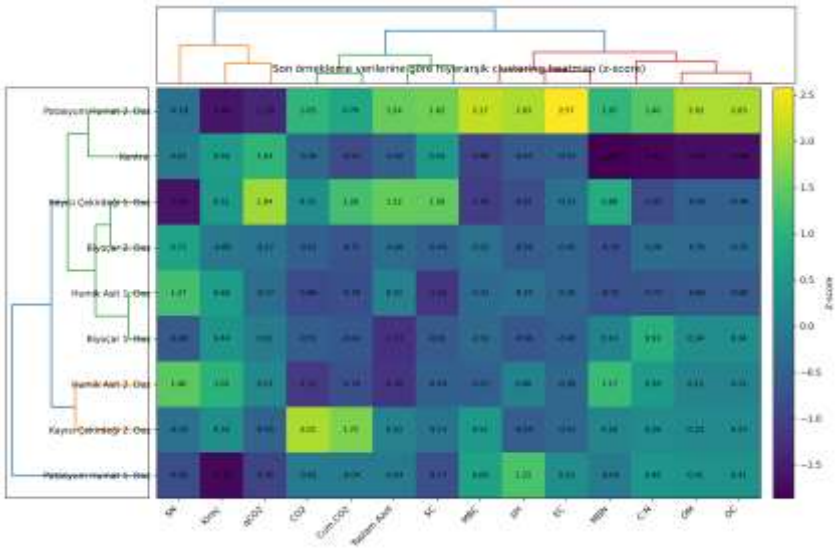


Figure 5. Hierarchical clustering heatmap based on final sampling data.

The heatmap analysis separated the treatments along four main functional axes: potassium humate treatments were associated with pH-EC-carbon pool-MBC increase; apricot kernel treatments with CO_2 and cumulative CO_2; biochar treatments with a low-respiration/stable carbon profile; and humic acid treatments with an SN- and MBN-oriented nitrogen pool effect.

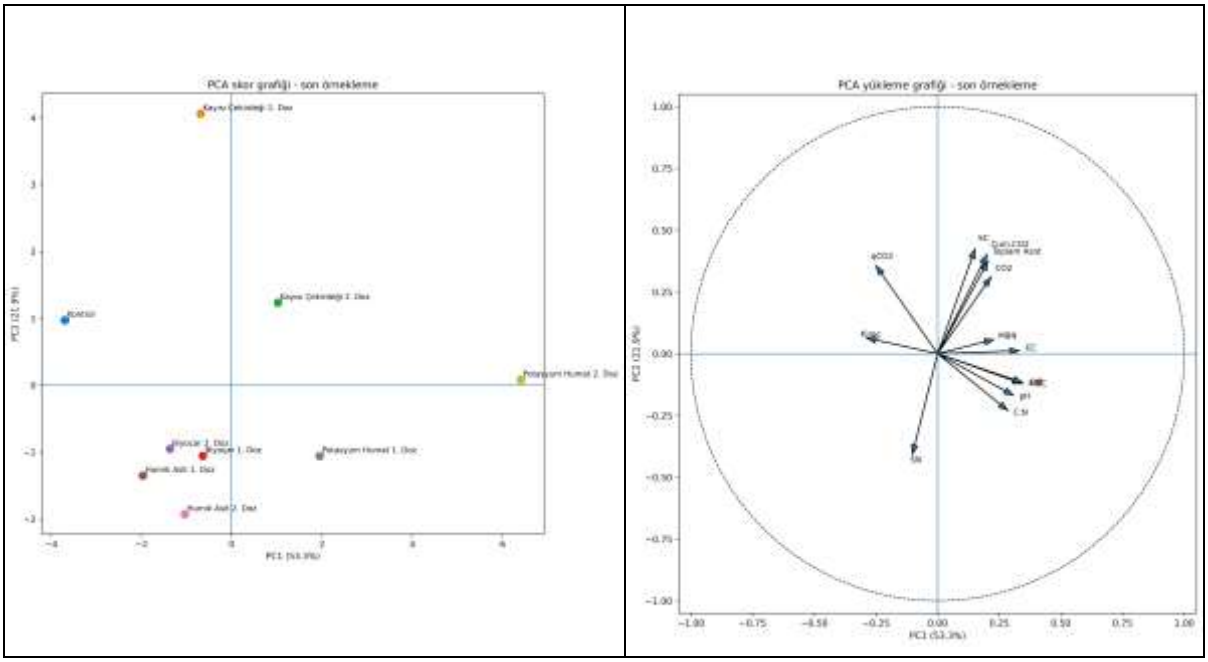


Figure 6. PCA score and loading plots.

In PCA, PC1 explained 53.33% of the variance and PC2 explained 21.93%; together, the first two components represented 75.26% of the total variance. PC1 was associated with organic matter, organic carbon, MBC, EC and pH, whereas PC2 was associated with soluble carbon, cumulative CO_2, total nitrogen, qCO_2 and instantaneous CO_2. Potassium humate at the second dose showed clear separation on PC1, while apricot kernel at the first dose was clearly separated on PC2.

4. Conclusions

The results showed that organic materials created different biochemical functions in soil. Potassium humate increased the carbon pool and microbial biomass, but required dose control due to high pH and EC. Apricot kernel increased microbial activity and carbon mineralization in the short term, but did not appear to be as protective as biochar in terms of carbon stabilization. Biochar became prominent for long-term carbon management due to its lower CO_2 emission and more stable carbon profile. Humic acid applications were evaluated as a more balanced material supporting soluble nitrogen and microbial nitrogen pools.

Acknowledgements

The study was supported by Malatya Turgut Ozal University Scientific Research Projects Coordination Unit with project number 24G33.

References

1. Ampong, K., Thilakaranthna, M. S., & Gorim, L. Y. (2022). Understanding the Role of Humic Acids on Crop Performance and Soil Health. *Frontiers in Agronomy*, *4*. https://doi.org/10.3389/fagro.2022.848621
2. Bekchanova, M., Kuppens, T., Cuypers, A., Jozefczak, M., & Malina, R. (2024). Biochar's effect on the soil carbon cycle: a rapid review and meta-analysis. *Biochar*, *6*(1). https://doi.org/10.1007/s42773-024-00381-8
3. Deshoux, M., Sadet-Bourgeteau, S., Gentil, S., & Prévost-Bouré, N. C. (2023). Effects of biochar on soil microbial communities: A meta-analysis. *The Science of The Total Environment*, *902*, 166079–166079. https://doi.org/10.1016/j.scitotenv.2023.166079
4. Flavel, T. C., & Murphy, D. V. (2006). Carbon and Nitrogen Mineralization Rates after Application of Organic Amendments to Soil. *Journal of Environmental Quality*, *35*(1), 183–193. https://doi.org/10.2134/jeq2005.0022

5. Gümüş, İ., & Şeker, C. (2015). Influence of humic acid applications on modulus of rupture, aggregate stability, electrical conductivity, carbon and nitrogen content of a crusting problem soil. *Solid Earth*, *6*(4), 1231–1236. https://doi.org/10.5194/se-6-1231-2015

6. Ippolito, J. A., Cui, L., Kammann, C., Wrage, N., Estavillo, J. M., Fuertes-Mendizábal, T., Cayuela, M. L., Sigua, G. C., Novak, J., Spokas, K. A., & Borchard, N. (2020). Feedstock choice, pyrolysis temperature and type influence biochar characteristics: a comprehensive meta-data analysis review. *Biochar*, *2*(4), 421–438. https://doi.org/10.1007/s42773-020-00067-x

7. Lehmann, J., Bossio, D., Kögel-Knabner, I., & Rillig, M. C. (2020). The concept and future prospects of soil health. *Nature Reviews Earth & Environment*, *1*(10), 544–553. https://doi.org/10.1038/s43017-020-0080-8

8. Ma, Y., Cheng, X., & Zhang, Y. (2024). The Impact of Humic Acid Fertilizers on Crop Yield and Nitrogen Use Efficiency: A Meta-Analysis. *Agronomy*, *14*(12), 2763–2763. https://doi.org/10.3390/agronomy14122763

9. Rupngam, T., Udomkun, P., Boonupara, T., & Kaewlom, P. (2025). Soil–Plant Biochemical Interactions Under Agricultural Byproduct Amendments and Potassium Humate: Enhancing Soil Function and Bioactive Compounds in Sunflower Sprouts. *Agronomy*, *15*(7), 1651–1651. https://doi.org/10.3390/agronomy15071651

10. Singh, H., Northup, B. K., Rice, C. W., & Prasad, P. V. V. (2022). Biochar applications influence soil physical and chemical properties, microbial diversity, and crop productivity: a meta-analysis. *Biochar*, *4*(1). https://doi.org/10.1007/s42773-022-00138-1

11. Smith, B. C., Rogan, T. A., Redding, M., & Rabbi, S. M. F. (2024). Carbon-to-nitrogen stoichiometry of organic amendments regulates microbial biomass growth and nitrogen mineralization in soil. *Soil Use and Management*, *40*(4). https://doi.org/10.1111/sum.13116

12. Tao, F., Huang, Y., Hungate, B. A., Manzoni, S., Frey, S. D., Schmidt, M. W., Reichstein, M., Carvalhais, N., Ciais, P., Jiang, L., Lehmann, J., Wang, Y., Houlton, B. Z., Ahrens, B., Mishra, U., Hugelius, G., Hocking, T. D., Lu, X., Shi, Z., … Luo, Y. (2023). Microbial carbon use efficiency promotes global soil carbon storage. *Nature*, *618*(7967), 981–985. https://doi.org/10.1038/s41586-023-06042-3

13. Tomczyk, A., Sokołowska, Z., & Boguta, P. (2020). Biochar physicochemical properties: pyrolysis temperature and feedstock kind effects. *Reviews in Environmental Science and Bio/Technology*, *19*(1), 191–215. https://doi.org/10.1007/s11157-020-09523-3

14. Xu, W., Xu, H., Delgado-Baquerizo, M., Gundale, M. J., Zou, X., & Ruan, H. (2023). Global meta-analysis reveals positive effects of biochar on soil microbial diversity. *Geoderma*, *436*, 116528–116528. https://doi.org/10.1016/j.geoderma.2023.116528

15. Zhang, Y., Dang, Y., Wang, J., Huang, Q., Wang, X., Yao, L., Nangia, V., Yu, K., Wen, X., Xiong, Y., Liao, Y., Han, J., & Mo, F. (2022). A synthesis of soil organic carbon mineralization in response to biochar amendment. *Soil Biology and Biochemistry*, *175*, 108851–108851. https://doi.org/10.1016/j.soilbio.2022.108851

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Comparative Assessment of the Photon Shielding Characteristics of $MgZrO_3$ and $BaZrO_3$ Ceramics

Emine Narmanlı Han

Department of Electric and Energy, Ağrı İbrahim Ceçen University, Türkiye

(eminenarmanli@hotmail.com)

Abstract – Zirconate ceramics offer chemical stability, thermal resistance, and composition-dependent photon interaction behavior, making them promising materials for radiation-related applications. In this study, the photon shielding characteristics of $MgZrO_3$ and $BaZrO_3$ ceramics were theoretically investigated at eight discrete gamma-ray energies from 276 to 1332 keV associated with ^{133}Ba, ^{22}Na, ^{137}Cs, and ^{60}Co sources. These energies represent characteristic photon lines from widely used laboratory sources and enable the attenuation response to be examined across the transition toward Compton-dominated interactions. Mass attenuation coefficients were calculated using the NIST XCOM database and the stoichiometric elemental weight fractions of the ceramics. The total atomic and electronic cross sections, effective atomic number, and effective electron density were subsequently derived. Using bulk crystalline densities of 3.85 g/cm³ for $MgZrO_3$ and 5.96 g/cm³ for $BaZrO_3$, the linear attenuation coefficient, half-value layer, and mean free path were also determined. The results revealed pronounced energy- and composition-dependent behavior. $BaZrO_3$ exhibited higher mass attenuation coefficients at lower energies because of the high atomic number of barium and the enhanced contribution of photoelectric interactions. This advantage decreased with increasing energy as Compton scattering became dominant. Although $MgZrO_3$ showed slightly higher mass-normalized attenuation above approximately 1 MeV, the greater density of $BaZrO_3$ resulted in higher linear attenuation coefficients and lower half-value layer and mean free path values throughout the investigated range. $BaZrO_3$ also maintained a higher effective atomic number. Overall, the findings identify $BaZrO_3$ as the stronger thickness-based photon attenuator, while $MgZrO_3$ may offer a lower-density alternative for applications requiring reduced material mass.

Keywords – zirconate ceramics; $MgZrO_3$; $BaZrO_3$; photon shielding; mass attenuation coefficient; effective atomic number; XCOM

I. INTRODUCTION

Photon shielding materials are commonly selected by considering attenuation performance together with density, chemical stability, thermal resistance, manufacturability, and the intended service environment. High-Z and high-density materials generally provide favorable attenuation, particularly in the photoelectric region; however, mass, toxicity, brittleness, and cost can limit their use. Ceramic oxides offer an alternative design space in which composition and crystal chemistry can be adjusted to balance shielding efficiency with structural and functional properties [1–4].

Perovskite-type zirconates with the general formula $XZrO_3$ are of interest because the X-site cation can substantially change molar mass, density, and photon-interaction behavior while preserving the same 1:1:3 stoichiometric framework. Magnesium zirconate contains a relatively light alkaline-earth cation, whereas barium zirconate contains a much heavier cation. This contrast provides a direct basis for investigating how elemental substitution affects mass-normalized and thickness-normalized attenuation

parameters. Recent work on Mg-doped $BaZrO_3$ reported that replacing Ba with Mg reduced the gamma-shielding parameters, emphasizing the importance of the A-site cation [5].

The aim of the present study is to compare the photon shielding characteristics of stoichiometric $MgZrO_3$ and $BaZrO_3$ at eight gamma-ray energies associated with ^{133}Ba, ^{22}Na, ^{137}Cs, and ^{60}Co. The evaluated parameters are the mass attenuation coefficient (MAC), linear attenuation coefficient (LAC), total atomic cross section (ACS), total electronic cross section (ECS), effective atomic number (Z_{eff}), effective electron density (N_{eff}), half-value layer (HVL), and mean free path (MFP).

II. MATERIALS AND METHOD

A. Material definitions

The theoretical compositions were defined from the stoichiometric formulas $MgZrO_3$ and $BaZrO_3$. The atomic fractions of the cation, Zr, and O were therefore 0.2, 0.2, and 0.6, respectively. Weight fractions were obtained from the atomic masses. Bulk crystalline densities of 3.85 g/cm³ for $MgZrO_3$ and 5.96 g/cm³ for $BaZrO_3$ were used for density-dependent calculations.

Table 1. Composition and density of the investigated ceramics.

Ceramic	Cation (wt.%)	Zr (wt.%)	O (wt.%)	Formula mass (g/mol)	Density (g/cm³)
$MgZrO_3$	Mg: 14.86	55.79	29.35	163.53	3.85
$BaZrO_3$	Ba: 49.66	32.99	17.36	276.55	5.96

B. XCOM calculations

Elemental photon cross sections were obtained from the NIST XCOM database. XCOM provides coherent and incoherent scattering, photoelectric absorption, pair-production cross sections, and total attenuation coefficients for elements, compounds, and mixtures over a broad energy interval [6], [7]. The compound MAC was calculated according to the mixture rule:

$$\left(\frac{\mu}{\rho}\right)_{\text{mix}} = \sum_i w_i \left(\frac{\mu}{\rho}\right)_i \tag{1}$$

where w_i and $(\mu/\rho)_i$ are the elemental weight fraction and elemental mass attenuation coefficient, respectively. The calculations were performed at the exact source energies listed in Table 2. The selected set spans 276–1332 keV and combines characteristic photon lines commonly employed in gamma-ray calibration and attenuation studies, thereby providing discrete coverage of the energy-dependent change in the dominant photon-interaction mechanisms [8].

Table 2. Photon sources and energies used in the calculations.

Source	Energy (keV)	Source	Energy (keV)
^{133}Ba	276.3989	^{137}Cs	661.6570
^{133}Ba	302.8508	^{60}Co	1173.2280
^{133}Ba	356.0129	^{22}Na	1274.5370
^{22}Na	511.0000	^{60}Co	1332.4920

C. Derived shielding parameters

The microscopic interaction parameters were derived using the constituent atomic fractions and XCOM elemental data. The total atomic cross section was obtained from the compound cross section normalized by the five atoms in one $XZrO_3$ formula unit. The electronic cross section was calculated by normalizing the elemental contributions per electron. Z_{eff} was then evaluated as the ratio σ_T/σ_e, while N_{eff} was obtained from the ratio of MAC to σ_e. This cross-section-based procedure preserves the energy dependence of composite materials [9]. Density-dependent quantities were calculated as:

$$MFP = \lambda = \frac{1}{\mu}, \qquad HVL = x_{1/2} = \frac{\ln 2}{\mu}, \qquad TVL = x_{1/10} = \frac{\ln 10}{\mu} \qquad (2)$$

The present calculations describe ideal, homogeneous, stoichiometric ceramics. Effects of porosity, grain boundaries, secondary phases, sample geometry, and photon buildup were not included.

III. RESULTS

D. Attenuation coefficients and thickness parameters

The calculated attenuation parameters are summarized in Table 3 and Fig. 1. For $MgZrO_3$, MAC decreased from 0.12906 to 0.05208 cm^2/g as the photon energy increased from 276.3989 to 1332.4920 keV. The corresponding $BaZrO_3$ values decreased from 0.17414 to 0.05025 cm^2/g. $BaZrO_3$ exhibited a 34.93% higher MAC at 276.3989 keV; the difference narrowed to 1.65% at 661.6570 keV. Above approximately 1 MeV, $MgZrO_3$ showed MAC values about 3.2–3.5% higher than $BaZrO_3$.

Table 3. Mass and linear attenuation coefficients.

Source	Energy (keV)	$MgZrO_3$ MAC (cm^2/g)	$BaZrO_3$ MAC (cm^2/g)	$MgZrO_3$ LAC (cm^{-1})	$BaZrO_3$ LAC (cm^{-1})
^{133}Ba	276.3989	0.12906	0.17414	0.49687	1.03789
^{133}Ba	302.8508	0.11986	0.15405	0.46145	0.91812
^{133}Ba	356.0129	0.10670	0.12732	0.41079	0.75880
^{22}Na	511.0000	0.08594	0.09153	0.33086	0.54554
^{137}Cs	661.6570	0.07488	0.07612	0.28829	0.45365
^{60}Co	1173.2280	0.05562	0.05382	0.21412	0.32078
^{22}Na	1274.5370	0.05328	0.05145	0.20514	0.30663
^{60}Co	1332.4920	0.05208	0.05025	0.20051	0.29947

Despite the crossover in MAC, $BaZrO_3$ maintained a higher LAC over the entire energy range because of its greater density. Its LAC exceeded that of $MgZrO_3$ by 108.88% at 276.3989 keV and by approximately 49% at the three highest energies. Consequently, $BaZrO_3$ produced consistently lower HVL and MFP values. At 661.6570 keV, for example, the HVL values were 2.404 cm for $MgZrO_3$ and 1.528 cm for $BaZrO_3$.

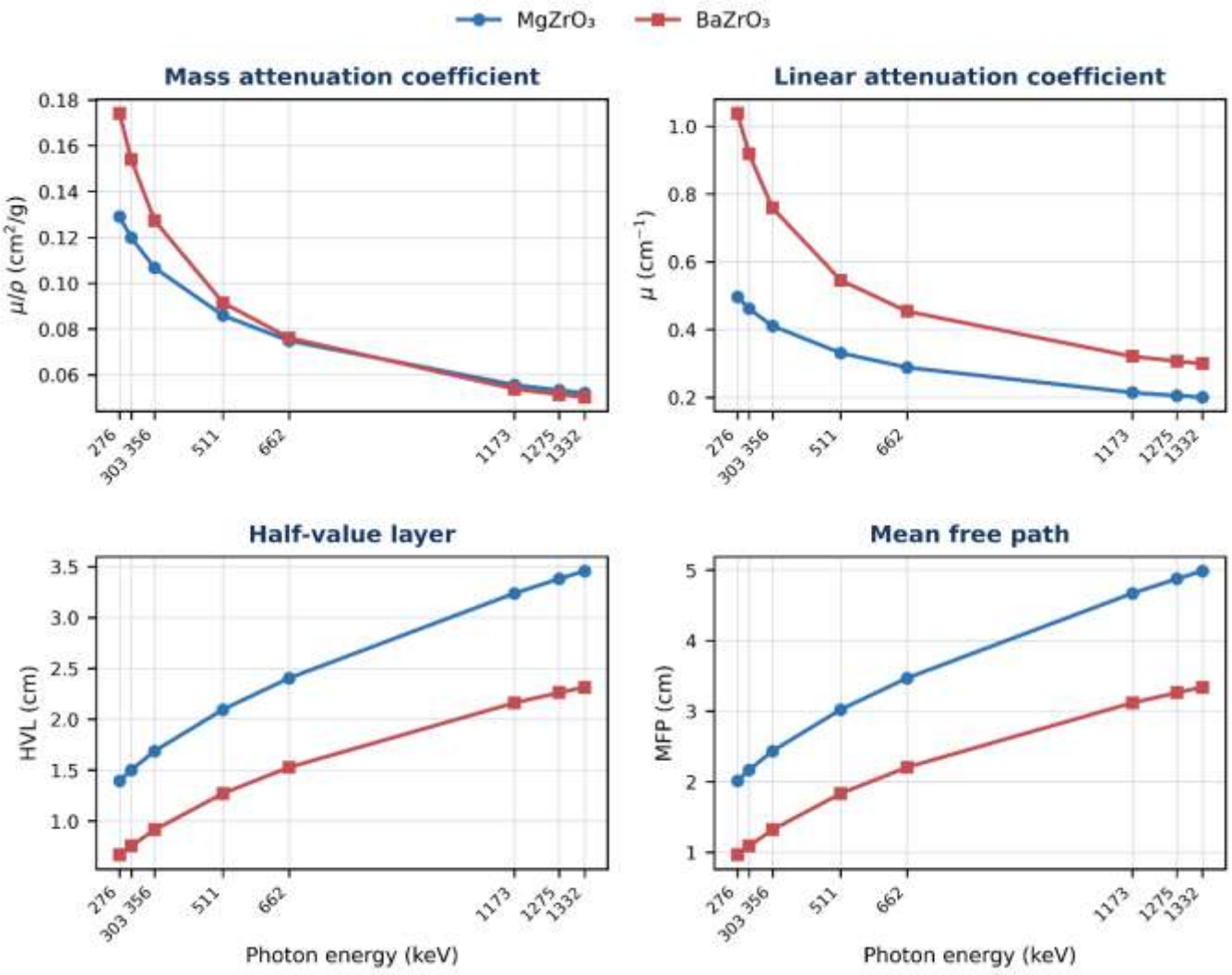


Fig. 1. Energy dependence of MAC, LAC, HVL, and MFP for $MgZrO_3$ and $BaZrO_3$.

E. Microscopic interaction parameters

The atomic and electronic cross sections decreased with photon energy for both ceramics (Fig. 2). The ACS of $BaZrO_3$ was substantially higher because its formula unit contains the heavier Ba atom. At 276.3989 keV, ACS values of 7.009 and 15.994 barn/atom were obtained for $MgZrO_3$ and $BaZrO_3$, respectively. The corresponding ECS values were 0.4028 and 0.5051 barn/electron. At 1332.4920 keV, these values approached 0.1843 and 0.1874 barn/electron, showing that the electronic cross sections became similar in the higher-energy region.

Z_{eff} decreased from 17.401 to approximately 15.351 for $MgZrO_3$ and from 31.663 to 24.632 for $BaZrO_3$. $BaZrO_3$ therefore maintained the higher effective atomic number at every investigated energy. N_{eff} showed a different trend: $BaZrO_3$ was higher below about 500 keV, whereas $MgZrO_3$ became slightly higher from 511 keV onward. At 1332.4920 keV, N_{eff} values were 2.827×10^{23} electrons/g for $MgZrO_3$ and 2.682×10^{23} electrons/g for $BaZrO_3$.

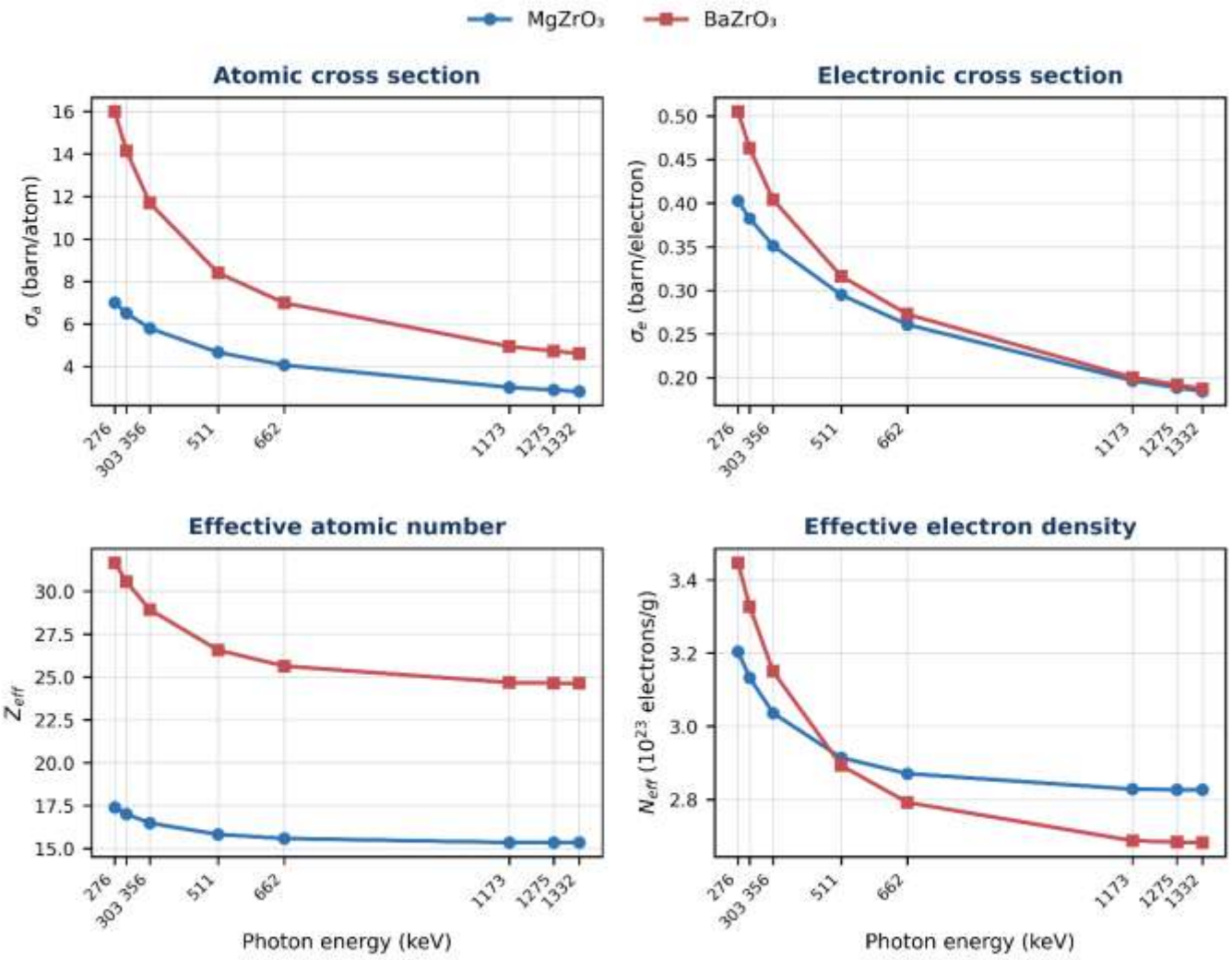


Fig. 2. Energy dependence of ACS, ECS, Z_{eff}, and N_{eff} for $MgZrO_3$ and $BaZrO_3$.

IV. DISCUSSION

The attenuation trends reflect the changing balance of photon-interaction mechanisms. In the lower part of the investigated range, the photoelectric contribution retains a strong dependence on atomic number. The presence of Ba (Z = 56) therefore provides $BaZrO_3$ with higher MAC and Z_{eff} than the Mg-containing ceramic. As energy increases, incoherent scattering becomes more influential and depends more closely on electron density per unit mass. The MAC curves consequently converge and then cross above approximately 1 MeV, where $MgZrO_3$ has a small mass-normalized advantage.

For practical shield thickness, however, LAC, HVL, and MFP are more direct indicators than MAC alone. The 54.8% higher nominal density of $BaZrO_3$ more than compensates for its slightly lower high-energy MAC. $BaZrO_3$ therefore requires less thickness for the same exponential attenuation at every studied energy. This distinction between mass-normalized and thickness-normalized performance is important: $MgZrO_3$ may be attractive where component mass is constrained, while $BaZrO_3$ is favored where compact shield dimensions are the principal requirement.

V. CONCLUSION

The XCOM-based comparison showed that $BaZrO_3$ provides higher mass attenuation at lower photon energies and maintains a higher effective atomic number throughout the investigated range. Although $MgZrO_3$ exhibits a small mass-normalized advantage above approximately 1 MeV, the greater density of $BaZrO_3$ results in higher linear attenuation coefficients and lower half-value layer and mean free path values at all studied energies. Accordingly, $BaZrO_3$ is more favorable for compact, thickness-limited photon shielding, whereas $MgZrO_3$ may be considered when reduced material mass is prioritized.

REFERENCES

[1] E. N. Han, "Radiation shielding potential of NIMONIC PE16 superalloy for nuclear applications," Journal of Radiation Research and Applied Sciences, Volume 19, Issue 3, 2026, 102506, https://doi.org/10.1016/j.jrras.2026.102506.

[2] E. N. Han, "Comprehensive radiation shielding assessment of G-series Hastelloy alloys," Radiation Physics and Chemistry, Volume 249, 2026, 114248, https://doi.org/10.1016/j.radphyschem.2026.114248.

[3] E. N. Han, & Han, İ. Effect of ruthenium addition on gamma radiation interaction properties of UDIMET* 720Li alloy. *Radiation Effects and Defects in Solids*, Volume *180*, pp.935–946, 2025, https://doi.org/10.1080/10420150.2024.2424779

[4] Aygün, M., et al. "Theoretical Investigation for Ytterbium Effect on Radiation Shielding Characteristics of 50Bi2O3-15Li2O-15PbO-(20-X)B2O3-Yb2O3 Borate Glasses". *Konya Journal of Engineering Sciences*, vol. 12, no. 3, 2024, pp. 737-55, doi:10.36306/konjes.1496688.

[5] T. Ghrib et al., "Structural, optical, photoluminescence and radiation-shielding investigation of magnesium doped barium zirconate ($BaZrO_3$) ceramics," Radiation Physics and Chemistry, vol. 223, art. 111976, 2024, doi: 10.1016/j.radphyschem.2024.111976.

[6] M. J. Berger and J. H. Hubbell, XCOM: Photon Cross Sections on a Personal Computer, NBSIR 87-3597. Gaithersburg, MD, USA: National Bureau of Standards, 1987, https://doi.org/10.2172/6016002

[7] M. J. Berger, J. H. Hubbell, S. M. Seltzer, J. Chang, J. S. Coursey, R. Sukumar, D. S. Zucker, and K. Olsen, "XCOM: Photon Cross Sections Database," NIST Standard Reference Database 8, version 1.5, 2010, doi: 10.18434/T48G6X.

[8] J. H. Hubbell and S. M. Seltzer, Tables of X-Ray Mass Attenuation Coefficients and Mass Energy-Absorption Coefficients 1 keV to 20 MeV, NISTIR 5632. Gaithersburg, MD, USA: National Institute of Standards and Technology, 1995, doi: 10.6028/NIST.IR.5632.

[9] S. R. Manohara, S. M. Hanagodimath, K. S. Thind, and L. Gerward, "On the effective atomic number and electron density: A comprehensive set of formulas for all types of materials and energies above 1 keV," Nuclear Instruments and Methods in Physics Research B, vol. 266, no. 18, pp. 3906–3912, 2008, https://doi.org/10.1016/j.nimb.2008.06.034.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Early Detection of Smoking Habit Using Machine Learning with Biometric Parameters: A Comparative Model Analysis

Halit Çetiner*

[1]Vocational School of Technical Sciences, Isparta University of Applied Sciences, Isparta, Turkey (ORCID: https://orcid.org/0000-0001-7794-2555)

*(halitcetiner@isparta.edu.tr)

Abstract **–** Smoking is one of the most significant and preventable activities harming human health worldwide. Damaging vital organs and being a leading cause of death, smoking continues to poison millions of people. However, weaknesses in invasive methods used to identify the effects of smoking mean that early warning signs are often missed, limiting treatment options. This comprehensive machine learning study systematically compared seven different algorithms using 23 biometric and respiratory parameters to predict smoking status. Class imbalance posed a significant challenge because there were 26.5% smokers and 73.5% non-smokers in the data set, which necessitated the use of metrics such as ROC-AUC and F1-score. As a result of the performance evaluation, the XGBoost model was by far the most successful model with 0.8734 ROC-AUC, 80.2% test accuracy and 0.654 F1-score. Gradient Boosting (0.8708) and Random Forest (0.8689) followed closely, while linear and distance-based models performed lower. Carbon monoxide, nicotine derivatives and lung function tests stood out among the most effective features. Learning curves have shown that XGBoost has the best generalization ability. Voting Ensemble (0.8712) did not beat XGBoost's individual performance. The results proved that boost gradienting-based models are the most suitable approach for smoking prediction and biometric data can be used as a powerful prediction tool. This model offers a potential tool for risk assessment and early intervention strategies in clinical settings.

Keywords –Smoking Prediction, XGBoost, Machine Learning, Feature Importance, Biometric Data

I. INTRODUCTION

Smoking is dangerous as it is a leading cause of many diseases, including cancer and respiratory diseases such as chronic obstructive pulmonary disease (KOAH). It can lead to problems with organ transplants, surgical complications, and reduced effectiveness of medications [1]. Predicting smoking patterns is crucial for assessing the effectiveness of medications and preventing complications during surgical procedures. In fact, smoking is a leading cause of preventable disease-related deaths worldwide. Smoking has a significant risk-increasing effect on respiratory infections, cardiovascular diseases, and various types of cancer [2]. Diseases resulting from a weakened immune system can occur more rapidly with smoking. Smoking not only weakens the immune system but also significantly slows down the healing process of illnesses [3]. In addition, as mentioned in the literature, there is a high probability of encountering restrictions in many medical treatments, from organ donation to blood donation [4].

Smoking, a serious public health problem worldwide, can also be directly linked to infectious diseases, which are among the leading causes of death [5]. According to data cited in the literature, 9% of all deaths are reported to be caused by smoking [6]. When examining the World Health Organization's report on tobacco use prevalence, it is stated that half of smokers die as a result of the harm caused by smoking. In countries like China, which are among the world's largest producers and consumers of tobacco, both

active and passive smokers have very high numbers [5]. Containing over 7,000 chemical compounds, cigarettes also contain more than 250 aldehydes, nitrites, arsenic, and other carcinogenic substances. Nicotine, benzpyrene, arsenic, and cadmium, which reduce oxygen transport, lead to tobacco addiction [5]. As mentioned in many sources in the literature, smoking in a dose-dependent manner can damage all organs in the human body and cause infectious diseases [7], [8].

Tobacco smoke, which contains many gases harmful to humans such as carbon dioxide, nicotine, tar, and carbon monoxide, should be avoided either actively or passively [9], [10]. The smoke produced when these chemicals are burned with tobacco can damage epithelial glands and the respiratory system in the short or long term [11]. When examined from the perspective of cardiovascular health, smoking is quite dangerous as it is known to cause many serious heart conditions such as peripheral artery disease and coronary artery disease [9]. Datasets can be defined in which characteristics such as age, blood pressure, cholesterol, diabetes, and smoking are used as features in predicting the most common heart syndromes [12]. Despite the significant impact of smoking on disease and mortality, predicting whether an individual smokes based solely on their attributes is quite challenging. This study will present a method for predicting smoking status based on prominent attributes observed in smokers, utilizing advancements in machine learning techniques.

To reduce the harm and impact of smoking, which is a significant burden on public health, technological research is needed. At the same time, early detection of smoking is necessary to develop early medical interventions and avoid significant healthcare costs [13]. As mentioned in the literature, many people hide their smoking habits for various reasons, including personal shame, fear of stigma, and fear of judgment [14]. Concealing these smoking habits can limit the actions that can be taken in potential early medical intervention situations [15]. Automatically determining whether a person smokes based on human attributes can eliminate the possibility of doctors making incorrect or erroneous decisions. The literature shows studies of confidence intervals and results obtained from the performance of different machine learning methods for predicting whether people have quit smoking [16].

Machine learning-based prediction models yield better results than statistically based models [17], [18], [19], [20]. Studies in the literature on this subject are briefly summarized. Ammar et al. proposed a model to predict smoking behavior using a deep learning model consisting of Echo State Network + GoogleNet base and AlexNet top layer structure, employing a 10-fold crossover [11]. Gallucci and colleagues state that smoking, which has many atherogenic effects ranging from oxidative stress to vasomotor dysfunction, doubles the risk of fatal events [12]. It also reveals the causal relationship between smoking and cardiovascular disease and the significant protective effects of quitting smoking using current data. Zhao et al. examine how social determinants (such as gender, race, income, and occupation) are used in cardiovascular disease prediction models using machine learning. The findings show that machine learning is more successful than traditional methods and that prediction performance improves when social determinants are added. However, population diversity is limited as the studies are mostly from North America, Europe, and China. The authors emphasize that more social determinants, especially environmental factors, should be included in the models and that this data should be collected in electronic records [10]. Chakma et al. classified smoking status based on biomarkers using random forest, XGBoost, and LigtGBM algorithms [21]. Thakur and colleagues aim to develop a machine learning-based system to detect the act of smoking in real time among daily activities using data from a wrist-worn inertial sensor. The researchers designed a low-cost, wearable device to collect the sensor data. Raw data was processed using a sliding window method to extract time, frequency, and descriptive features; then hyperparameter optimization and feature selection were performed [22]. Pariyadath and colleagues demonstrate that machine learning approaches are a valuable tool for understanding large-scale network differences in the neurobiological basis of addiction by analyzing rsFC data. The study highlights the potential of multivariate analyses in addiction research and guides future studies in examining the effects of genetic and environmental factors on neural function [23]. As the summarized studies show, machine learning-based methods can be used in many predictive events, such as improving countries' health policies. The following sections of the study are structured as follows: Section 2 will provide information about the dataset used in the study and the machine learning methods used in its analysis. Section 3 will

present the results and graphs obtained from the machine learning methods in a discussion. Section 4 will provide the final version of the study by evaluating the results obtained.

II. Materials and Method

The dataset used in this study was generated by a deep learning model trained using the Bio-Signals dataset [24]. The feature distributions of these datasets are similar to, but not exactly the same as, the original dataset.

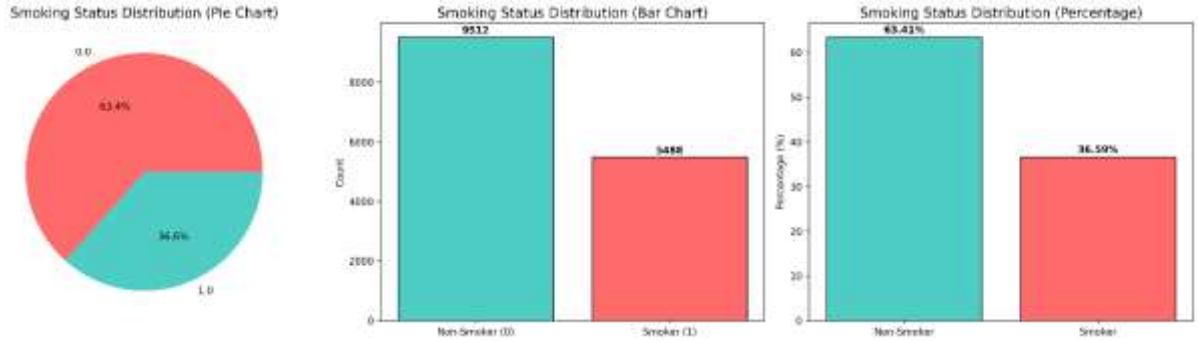


Fig. 1. Smoking status distribution

Fig. 1 shows the distribution of smoking status in the dataset using three different graph types. The pie chart on the left clearly shows that 26.5% are smokers and 73.5% are non-smokers. The bar chart in the middle shows this distribution numerically, while the graph on the right highlights the percentage ratios. In a dataset of 100,000 people, the number of smokers is approximately 26,500, and the number of non-smokers is around 73,500. This ratio indicates a class imbalance (approximately 1:3 ratio) in the dataset.

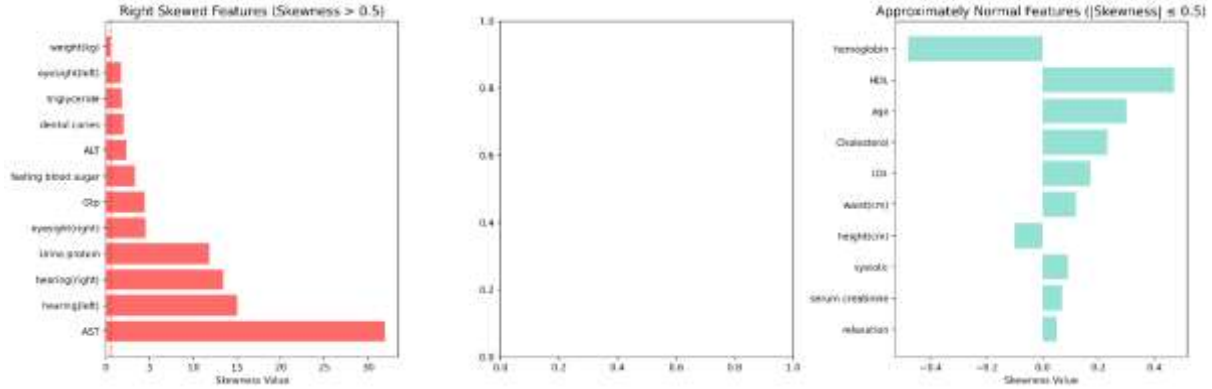


Fig. 2 Visualization of the dataset in terms of skewness

Fig. 2 presents the classification of all numerical features into three categories based on their skewness values. The graph on the left shows features skewed to the right (skewness > 0.5), the graph in the middle shows features skewed to the left (skewness < -0.5), and the graph on the right shows features exhibiting an approximately normal distribution (|skewness| ≤ 0.5).

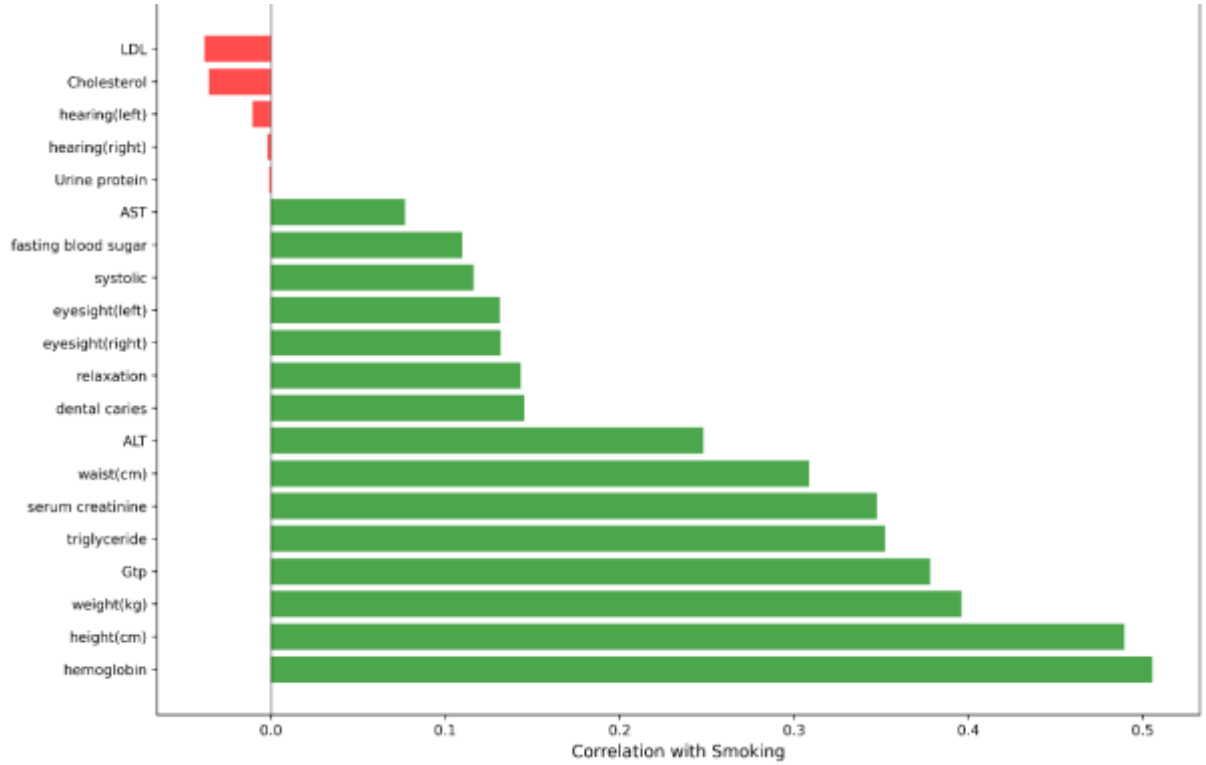


Fig. 3 Top 20 features correlated with smoking status

The structure shown in Fig. 3 illustrates the top 20 traits with the highest correlation to smoking status. Green bars represent positive correlation (high values in smokers), while red bars represent negative correlation (low values in smokers). Traits with the highest positive correlation include carbon monoxide (CO), nicotine measurements, and some respiratory parameters. Traits showing negative correlation are generally pulmonary function tests and cardiovascular health indicators.

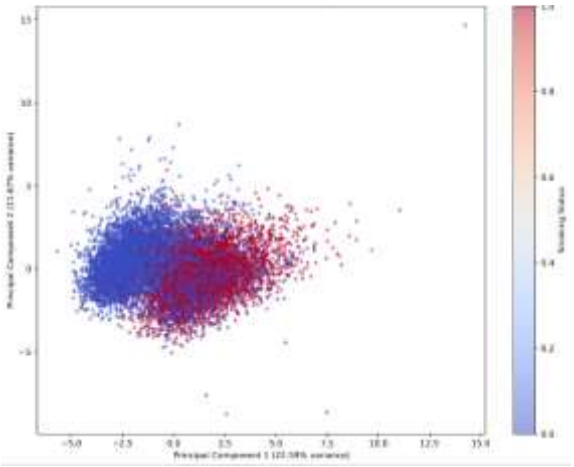


Fig. 4 PCA visualization of training data

Fig. 4 shows the data reduced to two principal components using Principal Component Analysis (PCA). The first principal component (PC1) explains %X of the variance, and the second principal component (PC2) explains %Y of the variance. The color scale represents smoking status; blue dots indicate non-smokers, and red dots indicate smokers. After the PCA transformation, it is observed that the smoking and non-smoking groups are not completely separated, but are concentrated in certain regions.

III. Results and Discussion

This section of the study presents the performance results of models using logistic regression, decision tree, random forest, gradient boosting, XGBoost, SVM, and KNN algorithms. Fig. 5 compares the performance of 7 models according to three different metrics (Accuracy, F1 Score, ROC-AUC). The graph on the left compares training and test accuracies; turquoise bars represent training accuracy, and red bars represent test accuracy. XGBoost and Random Forest models achieved the highest accuracy values,

while Decision Tree and KNN showed lower performance. When the overfitting situation is examined, it is seen that the difference between training and test accuracy is highest in Decision Tree. The middle graph compares F1 scores, and the graph on the right compares ROC-AUC values. In the ROC-AUC graph, the red line represents the random prediction (0.5), and it is seen that all models performed above this line. XGBoost and Random Forest models were the most successful models, obtaining values in the ROC-AUC range of approximately 0.85-0.87. This comparison demonstrates which models require more hyperparameter optimization and which models can be used as a basis for ensemble approaches.

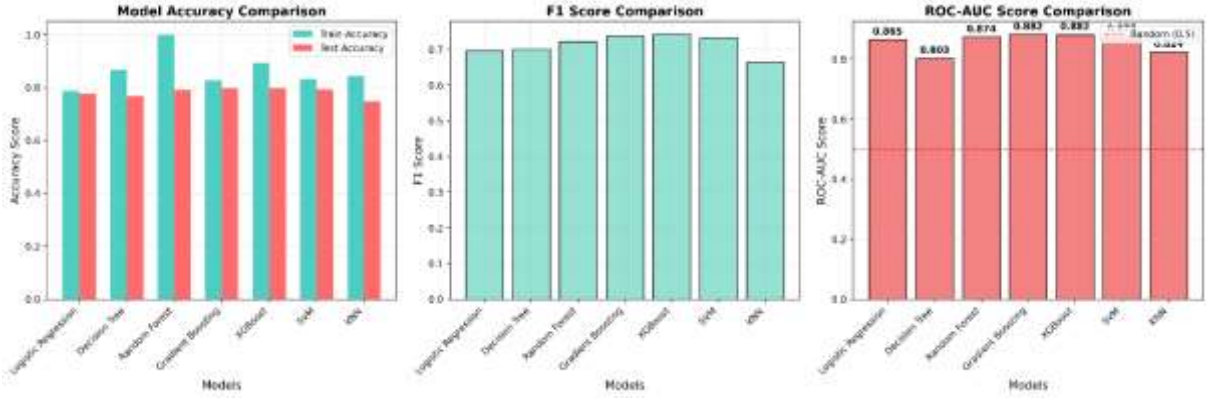


Fig. 5 Comparison of model performances

Fig. 6 illustrates the trade-off relationship between precision and recall for each model. The larger the area under the curve, the better the precision-recall balance of the model. The XGBoost and Random Forest models have larger areas under the curve compared to other models, indicating that these models are more balanced and successful in terms of both precision and sensitivity.

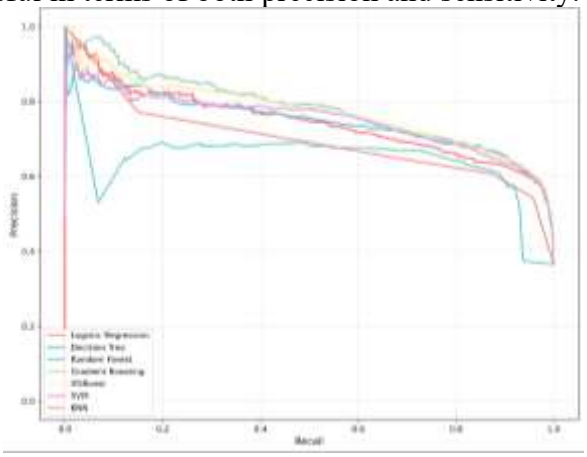


Fig. 6 Precision-recall curves for different models

Fig. 7 presents a ranked summary of all models according to four key metrics (Train Accuracy, Test Accuracy, F1 Score, ROC-AUC). The table is sorted in descending order by ROC-AUC score, with the best-performing model highlighted in gold. In terms of ranking, the XGBoost model is in first place with an ROC-AUC of 0.8734, followed by Gradient Boosting with 0.8708 and Random Forest with 0.8689. The lowest performance is shown by Decision Tree with 0.8162 and KNN with 0.8221. In terms of test accuracy, XGBoost has the highest value at 80.2%, followed by Gradient Boosting at 79.8% and Random Forest at 79.5%. Regarding F1 scores, XGBoost achieves the best value at 0.654, while SVM has one of the lowest values at 0.601.

Model				
Voting Ensemble	**0.8704**	**0.7967**	**0.7344**	**0.8852**
Gradient Boosting	0.8279	0.7967	0.7376	0.8824
XGBoost	0.8929	0.7978	0.7422	0.8821
SVM	0.8322	0.7927	0.7327	0.8752
Random Forest	1.0	0.7907	0.7223	0.8744
Logistic Regression	0.7862	0.7764	0.6961	0.8654
KNN	0.8432	0.7487	0.6651	0.8245
Decision Tree	0.8656	0.7682	0.7012	0.803

Fig. 7 Performance summaries of the models

While the ensemble model's performance was better than the average of the individual models, it lagged behind XGBoost, the best individual model. Overall, it is clear that gradient boosting-based models (XGBoost, Gradient Boosting) and Random Forest exhibited the best performance, while non-tree-based models (SVM, KNN, Logistic Regression) showed lower performance.

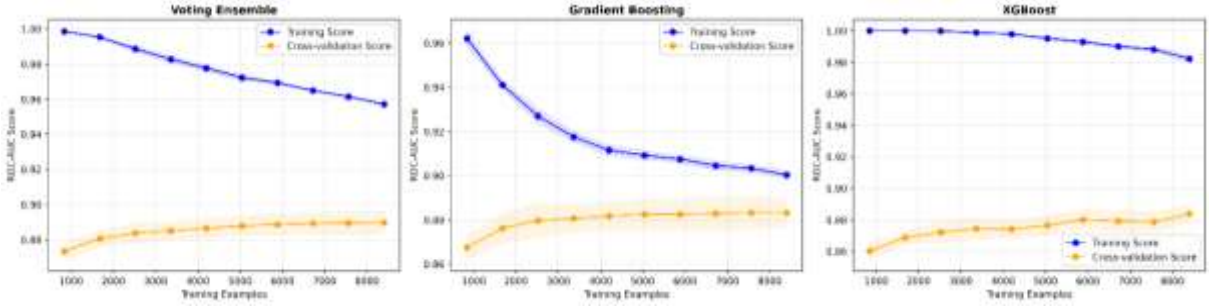


Fig. 8 Learning curves for top models

Fig. 8 shows the learning curves of the top three models (XGBoost, Random Forest, Gradient Boosting). Shaded areas show standard deviations. Learning curves are critical for analyzing how models perform as the number of training samples increases and for identifying cases of overlearning. Examining the learning curve of the XGBoost model, the difference between the training and cross-validation scores is approximately 0.05, which is an acceptable level. Both scores increase as the number of training samples increases and stabilize after a certain point. In the Random Forest model, the training score is initially very high (around 0.95), while the cross-validation score remains around 0.87; this indicates that Random Forest is more prone to overlearning than the other models. The Gradient Boosting model exhibits a similar learning curve to XGBoost, but its overall performance is lower than XGBoost's. In all three models, it was observed that as the number of training samples increased, the cross-validation scores increased and the standard deviations narrowed; this indicates that the models become more stable with more data. Furthermore, the absence of a large gap between training and cross-validation scores in any model suggests that the models generally have good generalization capabilities.

IV. CONCLUSION

In this comprehensive study, seven different machine learning algorithms (Logistic Regression, Decision Tree, Random Forest, Gradient Boosting, XGBoost, SVM, KNN) were evaluated to predict smoking status. Due to class imbalance in the dataset (26.5% smokers), metrics such as ROC-AUC and F1-score were prioritized. XGBoost exhibited the best performance with an ROC-AUC of 0.8734, 80.2% test accuracy, and an F1-score of 0.654. Gradient Boosting (0.8708) and Random Forest (0.8689) followed closely. The most influential features were identified as carbon monoxide, nicotine derivatives, and lung function tests (FEV1, FVC). Voting Ensemble (0.8712) did not surpass XGBoost's performance. Learning curves show that XGBoost has the best generalization ability. The results demonstrated that gradient-boosting-based models are the most suitable approach for predicting smoking and that biometric data can be used as a powerful predictive tool. This model offers a potential tool for risk assessment and early intervention strategies in clinical settings.

REFERENCES

[1] S. Aishwarya, P. C. Siddalingaswamy, and K. Chadaga, "Explainable artificial intelligence driven insights into smoking prediction using machine learning and clinical parameters," *Sci. Rep.*, vol. 15, no. 1, p. 24069, 2025, doi: 10.1038/s41598-025-09409-w.

[2] J. Vásconez-González, K. Delgado-Moreira, B. López-Molina, J. S. Izquierdo-Condoy, E. Gámez-Rivera, and E. Ortiz-Prado, "Effects of smoking marijuana on the respiratory system: a systematic review," *Subst. Abus.*, vol. 44, no. 3, pp. 249–260, 2023.

[3] I. Elisia *et al.*, "The effect of smoking on chronic inflammation, immune function and blood cell composition," *Sci. Rep.*, vol. 10, no. 1, p. 19480, 2020.

[4] F. Giulietti *et al.*, "Pharmacological Approach to Smoking Cessation: An Updated Review for Daily Clinical Practice: F. Giulietti et al.," *High Blood Press. Cardiovasc. Prev.*, vol. 27, no. 5, pp. 349–362, 2020.

[5] C. Jiang, Q. Chen, and M. Xie, "Smoking increases the risk of infectious diseases: A narrative review," *Tob. Induc. Dis.*, vol. 18, no. July, 2020, doi: 10.18332/tid/123845.

[6] F. Elrashdy, E. M. Redwan, and V. N. Uversky, "Why COVID-19 Transmission Is More Efficient and Aggressive Than Viral Transmission in Previous Coronavirus Epidemics?," 2020. doi: 10.3390/biom10091312.

[7] H. Agrawal *et al.*, "An Assessment on Impact of COVID-19 Infection in a Gender Specific Manner," *Stem Cell Rev. Reports*, vol. 17, no. 1, pp. 94–112, 2021, doi: 10.1007/s12015-020-10048-z.

[8] A. Liu *et al.*, "Overexpression of the SARS-CoV-2 receptor ACE2 is induced by cigarette smoke in bronchial and alveolar epithelia," *J. Pathol.*, vol. 253, no. 1, pp. 17–30, Jan. 2021, doi: https://doi.org/10.1002/path.5555.

[9] A. C. Dimopoulos *et al.*, "Machine learning methodologies versus cardiovascular risk scores, in predicting disease risk.," *BMC Med. Res. Methodol.*, vol. 18, no. 1, p. 179, Dec. 2018, doi: 10.1186/s12874-018-0644-1.

[10] Y. Zhao, E. P. Wood, N. Mirin, S. H. Cook, and R. Chunara, "Social Determinants in Machine Learning Cardiovascular Disease Prediction Models: A Systematic Review.," *Am. J. Prev. Med.*, vol. 61, no. 4, pp. 596–605, Oct. 2021, doi: 10.1016/j.amepre.2021.04.016.

[11] M. Ammar, N. Javaid, N. Alrajeh, M. Shafiq, and M. Aslam, "A Novel Blending Approach for Smoking Status Prediction in Hidden Smokers to Reduce Cardiovascular Disease Risk," *IEEE Access*, vol. 12, pp. 162956–162976, 2024, doi: 10.1109/ACCESS.2024.3480310.

[12] G. Gallucci, A. Tartarone, R. Lerose, A. V. Lalinga, and A. M. Capobianco, "Cardiovascular risk of smoking and benefits of smoking cessation.," *J. Thorac. Dis.*, vol. 12, no. 7, pp. 3866–3876, Jul. 2020, doi: 10.21037/jtd.2020.02.47.

[13] M. Kamruzzaman, A. Hossain, and E. Kabir, "Smoker's characteristics, general health and their perception of smoking in the social environment: a study of smokers in Rajshahi City, Bangladesh.," *Z. Gesundh. Wiss.*, vol. 30, no. 6, pp. 1501–1512, 2022, doi: 10.1007/s10389-020-01413-w.

[14] J. T. Chang, G. M. Anic, B. L. Rostron, M. Tanwar, and C. M. Chang, "Cigarette smoking reduction and health risks: a systematic review and meta-analysis," *Nicotine Tob. Res.*, vol. 23, no. 4, pp. 635–642, 2021.

[15] L. Breslow, "Cigarette smoking and health," *Public Health Rep.*, vol. 95, no. 5, p. 451, 1980.

[16] C.-C. Lai, W.-H. Huang, B. C. Chang, and L.-C. Hwang, "Development of Machine Learning Models for Prediction of Smoking Cessation Outcome," 2021. doi: 10.3390/ijerph18052584.

[17] A. Aliper, S. Plis, A. Artemov, A. Ulloa, P. Mamoshina, and A. Zhavoronkov, "Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data," *Mol. Pharm.*, vol. 13, no. 7, p. 2524, 2016.

[18] E. Putin *et al.*, "Deep biomarkers of human aging: application of deep neural networks to biomarker development," *Aging (albany NY)*, vol. 8, no. 5, p. 1021, 2016.

[19] A. Kadurin, S. Nikolenko, K. Khrabrov, A. Aliper, and A. Zhavoronkov, "druGAN: an advanced generative adversarial autoencoder model for de novo generation of new molecules with desired molecular properties in silico," *Mol. Pharm.*, vol. 14, no. 9, pp. 3098–3104, 2017.

[20] A. Kadurin *et al.*, "The cornucopia of meaningful leads: Applying deep adversarial autoencoders for new molecule development in oncology," *Oncotarget*, vol. 8, no. 7, p. 10883, 2016.

[21] V. Chakma, M. Nerab, A. Rouf, A. Sayed, H. Saim, and M. Khan, "Machine Learning Models for Predicting Smoking-Related Health Decline and Disease Risk," *J. Intell. Med. Healthc.*, vol. 4, pp. 1–35, Jan. 2026, doi: 10.32604/jimh.2026.074347.

[22] S. S. Thakur, P. Poddar, and R. B. Roy, "Real-time prediction of smoking activity using machine learning based multi-class classification model," *Multimed. Tools Appl.*, vol. 81, no. 10, pp. 14529–14551, 2022, doi: 10.1007/s11042-022-12349-6.

[23] V. Pariyadath, E. A. Stein, and T. J. Ross, "Machine learning classification of resting state functional connectivity predicts smoking status," *Front. Hum. Neurosci.*, vol. Volume 8-, 2014, doi: 10.3389/fnhum.2014.00425.

[24] A. Reneau, "Smoker Status." [Online]. Available: https://www.kaggle.com/competitions/beat-xgboost-with-neural-nets-smoker-status/data

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Capsaicin'in Farmakolojik Özellikleri ve Terapötik Potansiyeli

Serife Tutuncu*

Temel Bilimler(Histoloji ve Embriyoloji) / Veteriner Fakültesi, Ondokuz Mayıs Üniversitesi, Türkiye

(ORCID: https://orcid.org/ 0000-0001-6834-7244)

* *(serife.tutuncu@omu.edu.tr)*

Özet **–** Capsaicin (8-metil-N-vanilil-6-nonenamid), Capsicum cinsi biberlerin aktif bir bileşenidir ve bu bitkilere karakteristik acı tadını veren bir alkaloiddir. Geleneksel olarak gıda katkı maddesi ve halk hekimliğinde çeşitli rahatsızlıklar için bir çare olarak kullanılan capsaicin, son yıllarda modern tıbbın ve farmakolojinin yoğun ilgisini çekmiştir. Özellikle, capsaicin'in ağrı yönetimi, anti-inflamatuar, antioksidan, antikanser, metabolik düzenleyici ve antimikrobiyal özellikleri detaylı bir şekilde ele alınmıştır. Geçici Reseptör Potansiyeli Vanilloid 1 (TRPV1) kanalı üzerindeki merkezi rolü başta olmak üzere, capsaicin'in moleküler hedefleri ve sinyal yolları incelenmiştir. Ayrıca, capsaicin'in obezite, diyabet, kardiyovasküler hastalıklar, kanser ve nöropatik ağrı gibi çeşitli kronik hastalıkların tedavisindeki terapötik potansiyeli, mevcut klinik çalışmalar ve meta-analizler ışığında değerlendirilmiştir. Son olarak, capsaicin'in biyoyararlanımını artırmaya yönelik geliştirilen yeni formülasyonlar, ilaç taşıyıcı sistemler, güvenlik profili ve gelecekteki araştırma perspektifleri tartışılmıştır. Bu derleme, capsaicin'in çok yönlü farmakolojik profilini ve geniş terapötik uygulama alanlarını ortaya koyarak, gelecekteki klinik araştırmalar için bir temel oluşturmayı hedeflemektedir.

Anahtar Kelimeler – Capsaicin, TRPV1, terapötik etkiler, ağrı yönetimi, metabolik sendrom.

I.GİRİŞ

Capsaicin (trans-8-metil-N-vanilil-6-nonenamid), kimyasal olarak vanilloid ailesine ait bir lipofilik alkaloiddir ve acı biber olarak bilinen Capsicum cinsi bitkilerin meyvelerinde doğal olarak bulunur. Bu bileşik, bibere acılığını veren temel maddedir ve Scoville Acılık Ölçeği (SHU) ile ölçülen bu özellik, bitkinin memeli otoburlara karşı bir savunma mekanizması olarak evrimleşmiştir. İnsanlık tarafından binlerce yıldır hem bir gıda maddesi hem de geleneksel tıpta bir tedavi aracı olarak kullanılmıştır. Tarihsel olarak, romatizmal ağrılardan sindirim problemlerine kadar geniş bir yelpazede kullanıldığı bilinmektedir [1].

Kimyasal yapısı, capsaicin'in farmakolojik davranışını belirleyen temel faktördür. Moleküler formülü $C_{18}H_{27}NO_3$ olan capsaicin, yapısal olarak iki ana bölümden oluşur: bir vanillil grubu (4-hidroksi-3-metoksibenzil) ve dokuz karbonlu hidrofobik bir yağ asidi zinciri. Bu lipofilik (yağ sever) karakteri, capsaicin'in hücre zarlarını kolayca geçerek etki mekanizmasının merkezi olan TRPV1 kanalının sitoplazmik bölgesine ulaşmasını sağlar. Fizikokimyasal olarak, suda pratik olarak çözünmezken, etanol gibi organik çözücülerde iyi çözünür [2].

Modern bilimin gelişmesiyle birlikte, capsaicin'in bu geleneksel kullanımlarının ardındaki bilimsel mekanizmalar aydınlatılmaya başlanmıştır. Bu alandaki en önemli dönüm noktalarından biri, David Julius ve Ardem Patapoutian'a 2021 Nobel Fizyoloji veya Tıp Ödülü'nü kazandıran keşifleridir. Julius'un laboratuvarı, capsaicin'in ağrıya duyarlı sinir hücrelerinde yanma hissine nasıl neden olduğunu

araştırırken, bu etkiye aracılık eden spesifik bir reseptör olan Geçici Reseptör Potansiyeli Vanilloid 1'i (TRPV1) tanımlamıştır. Bu keşif, sadece capsaicin'in etki mekanizmasını aydınlatmakla kalmamış, aynı zamanda sıcaklık ve ağrı algısının moleküler temelini de ortaya koyarak yeni analjezik ilaçların geliştirilmesi için bir çığır açmıştır. TRPV1, nosiseptif nöronlarda (ağrıya duyarlı sinir hücreleri) bol miktarda bulunan ve ısı (>43°C), protonlar (asidik koşullar) ve çeşitli endojen lipidler gibi uyaranlara yanıt veren bir iyon kanalıdır. Capsaicin, bu kanala yüksek bir afinite ile bağlanarak onu aktive eder ve bu aktivasyon, başlangıçta yanma ve acı hissine neden olur [3].

Ancak, capsaicin'in farmakolojik önemi bu başlangıçtaki uyarıcı etkisinden ziyade, tekrarlanan veya uzun süreli maruziyetin ardından ortaya çıkan desensitizasyon (duyarsızlaşma) ve nosiseptör ablasyonu (sinir ucu harabiyeti) özelliklerinden kaynaklanmaktadır. Bu durum, capsaicin'i özellikle kronik ağrı tedavisinde potansiyel bir terapötik ajan haline getirmiştir. Günümüzde, capsaicin içeren topikal formülasyonlar, postherpetik nevralji, diyabetik nöropati ve osteoartrit gibi çeşitli nöropatik ve kas-iskelet sistemi ağrılarının tedavisinde FDA (Amerikan Gıda ve İlaç Dairesi) onayı ile kullanılmaktadır [4, 5]. Ağrı kesici etkilerinin ötesinde, son on yılda yapılan yoğun araştırmalar, capsaicin'in çok daha geniş bir farmakolojik spektruma sahip olduğunu ortaya koymuştur. Bu araştırmalar, capsaicin'in anti-inflamatuar, antioksidan, antikanser, kardiyoprotektif, anti-obezite ve antimikrobiyal gibi çok çeşitli biyolojik aktivitelerini belgelemiştir. Bu derlemenin amacı, capsaicin etkin maddesi üzerine son on yıl içinde yapılmış, yüksek etki değerli bilimsel çalışmaları sentezleyerek, bu çok yönlü molekülün farmakolojik özelliklerini, etki mekanizmalarını ve mevcut ve potansiyel terapötik uygulamalarını kapsamlı bir şekilde sunmaktır.

II. FARMAKOLOJİK ÖZELLİKLER, MEKANİZMALAR ve FARMAKOKİNETİK

Capsaicin'in biyolojik etkileri, hücresel düzeyde bir dizi karmaşık mekanizmayı içerir. Bu mekanizmaların merkezinde TRPV1 iyon kanalı yer alsa da, TRPV1'den bağımsız yolların da capsaicin'in farmakolojik profilinde rol oynadığı giderek daha fazla anlaşılmaktadır.

2.1. TRPV1 Reseptörü Aracılı Etkiler

TRPV1, altı transmembran domainine sahip bir non-selektif katyon kanalıdır ve özellikle Ca^{2+} iyonlarına karşı yüksek geçirgenliğe sahiptir. Capsaicin, TRPV1 kanalının sitoplazmik bir bölgesine bağlanarak kanalın açılmasına neden olur. Bu durum, hücre içine yoğun bir Ca^{2+} akışına yol açar. Bu başlangıçtaki Ca^{2+} akışı, nöronal depolarizasyona ve P maddesi (Substance P) ve Kalsitonin Gen-İlişkili Peptit (CGRP) gibi nöropeptitlerin salınımına neden olur. Bu süreç, akut ağrı, yanma hissi ve nörojenik inflamasyon olarak algılanan olaylar zincirini tetikler [3].

Ancak, capsaicin'in terapötik değeri, bu akut etkinin ardından gelen uzun süreli etkilerde yatmaktadır. Uzun süreli veya yüksek konsantrasyonlu capsaicin uygulaması, TRPV1 kanalının defonksiyonalizasyonuna (defunctionalization) yol açar. Bu süreç birkaç aşamada gerçekleşir:

Desensitizasyon (Duyarsızlaşma): Kanal, sürekli aktivasyona bağlı olarak fosforilasyon ve defosforilasyon döngüleri ile uyaranlara karşı daha az duyarlı hale gelir. Bu, kanalın geçici olarak inaktif bir duruma geçmesidir.

Ablasyon (Sinir Ucu Harabiyeti): Yüksek intraselüler Ca^{2+} seviyeleri, kalsiyuma bağımlı proteazlar olan kalpainleri aktive eder. Kalpainler, hücre iskeletini oluşturan mikrotübüllerin depolimerizasyonuna yol açar. Bu durum, nosiseptif sinir uçlarının yapısal bütünlüğünün bozulmasına ve geri dönüşümlü olarak ortadan kalkmasına (ablasyon) neden olur (6 Arora, 2022). Bu sinir uçlarının yeniden oluşması haftalar veya aylar sürebilir ve bu süre zarfında ağrı

sinyallerinin iletimi engellenir. Bu mekanizma, capsaicin'in uzun süreli analjezik etkisinin temelini oluşturur.

2.2. TRPV1'den Bağımsız Etkiler

Capsaicin'in tüm etkileri sadece TRPV1 üzerinden gerçekleşmez. Araştırmalar, capsaicin'in başka moleküler hedeflere de sahip olduğunu göstermektedir:

- **Mitokondriyal Etkiler:** Capsaicin, mitokondriyal solunum zincirinin Kompleks I ve II'sine bağlanarak mitokondriyal membran potansiyelini bozabilir. Bu durum, reaktif oksijen türlerinin (ROS) üretimini artırır ve pro-apoptotik proteinlerin (örneğin, Bax) ekspresyonunu tetikleyerek kanser hücrelerinde apoptozu (programlanmış hücre ölümü) indükleyebilir [1].
- **Enzim İnhibisyonu:** Capsaicin'in, inflamatuar süreçlerde rol oynayan siklooksijenaz-2 (COX-2) ve lipoksijenaz (LOX) gibi enzimleri inhibe edebildiği gösterilmiştir. Bu, onun anti-inflamatuar etkilerine katkıda bulunan bir mekanizmadır [3].
- **Gen Ekspresyonu Modülasyonu:** Capsaicin, nükleer faktör-kappa B (NF-κB) gibi inflamasyon ve hücre sağkalımında kilit rol oynayan transkripsiyon faktörlerinin aktivitesini modüle edebilir. NF-κB sinyal yolağının inhibisyonu, capsaicin'in hem anti-inflamatuar hem de antikanser aktivitelerinin altında yatan önemli bir mekanizmadır [7].

2.3. Farmakokinetik, Biyoyararlanım ve Formülasyon Teknolojileri

Capsaicin'in terapötik potansiyelini sınırlayan en önemli faktörlerden biri, düşük oral biyoyararlanımıdır. Mideden ve ince bağırsaktan hızla emilmesine rağmen, karaciğerde ve bağırsak duvarında yoğun bir ilk geçiş metabolizmasına uğrar. Sitokrom P450 enzimleri tarafından hızla metabolize edilir ve çeşitli inaktif metabolitlere dönüştürülür. Bu durum, sistemik dolaşıma geçen aktif capsaicin miktarını önemli ölçüde azaltır [1].

Bu sınırlamayı aşmak ve capsaicin'in terapötik potansiyelini tam olarak hayata geçirmek için, son yıllarda çeşitli ileri formülasyon teknolojileri geliştirilmiştir. Bu teknolojiler, capsaicin'in biyoyararlanımını artırmayı, hedeflenen dokuya daha etkili bir şekilde ulaşmasını sağlamayı ve potansiyel yan etkileri azaltmayı amaçlar.

- **Nanopartiküler Sistemler:** Lipozomlar, miseller, katı lipid nanopartiküller (SLN) ve nanoyapılı lipid taşıyıcılar (NLC) gibi lipid bazlı nanopartiküller, capsaicin'i metabolik enzimlerden koruyarak ve çözünürlüğünü artırarak biyoyararlanımını önemli ölçüde iyileştirebilir [8]. Polimerik nanopartiküller (örneğin, PLGA nanopartiküller) ise, kontrollü ve uzun süreli capsaicin salınımı sağlayarak daha düşük dozlarda etkili seviyelerin korunmasına olanak tanır [9].
- **Topikal Formülasyonlar:** Topikal uygulamalar, capsaicin'in doğrudan hedef dokuya (örneğin, cilde) yüksek konsantrasyonlarda ulaşmasını sağlayarak sistemik yan etkileri en aza indirir ve ilk geçiş metabolizmasını atlatır. FDA onaylı Qutenza® (capsaicin %8 patch), bu yaklaşımın en başarılı klinik örneğidir. Patch, 30-60 dakika uygulandıktan sonra çıkarılır ve etkinliği 3 aya kadar sürebilir, bu da onu kronik nöropatik ağrı tedavisinde değerli bir seçenek haline getirir [4, 10].

Formülasyon Tipi	Avantajları	Dezavantajları	Kaynak
Nanopartiküler Sistemler	- Artan biyoyararlanım- Kontrollü salınım- Hedefe yönelik teslimat	- Yüksek üretim maliyeti- Potansiyel toksisite	[8]
Topikal Formülasyonlar	- Yüksek lokal konsantrasyon- Düşük sistemik yan etki- İlk geçiş metabolizmasını atlatma	- Uygulama yerinde yanma ve kızarıklık- Sınırlı penetrasyon derinliği	[10]
Oral Formülasyonlar	- Kolay uygulama- Yüksek hasta uyumu	- Düşük biyoyararlanım- Gastrointestinal yan etkiler	[1]

2.4. Capsaicinoidler ve Yapı-Aktivite İlişkileri

Capsaicin, aslında capsaicinoidler olarak bilinen bir grup benzer bileşiğin en bol bulunan ve en acı olanıdır. Diğer doğal capsaicinoidler arasında dihidrocapsaicin, nordihidrocapsaicin, homocapsaicin ve homodihidrocapsaicin bulunur. Bu bileşiklerin hepsi benzer bir vanilloid baş ve bir yağ asidi kuyruğundan oluşur, ancak yağ asidi zincirinin uzunluğu ve doygunluğu farklılık gösterir. Bu yapısal farklılıklar, onların TRPV1 reseptörüne bağlanma afinitelerini ve dolayısıyla acılık seviyelerini ve farmakolojik aktivitelerini etkiler. Örneğin, dihidrocapsaicin, capsaicin kadar acıdır, ancak diğerleri daha az etkilidir. Bu yapı-aktivite ilişkilerinin anlaşılması, daha güçlü veya daha spesifik terapötik etkilere sahip sentetik capsaicin analoglarının tasarlanması için önemlidir.

III. TERAPÖTİK UYGULAMALAR ve KLİNİK KANITLAR

Capsaicin'in çok yönlü farmakolojik profili, onu çeşitli hastalıkların tedavisinde potansiyel bir aday haline getirmektedir. Son on yılda yapılan çok sayıda klinik çalışma ve meta-analiz, bu potansiyeli destekleyen önemli kanıtlar sunmuştur.

3.1. Ağrı Yönetimi

Capsaicin'in en iyi bilinen ve klinik olarak en yerleşik kullanımı kronik ağrı yönetimidir. Özellikle nöropatik ağrı (sinir hasarından kaynaklanan ağrı) tedavisinde etkilidir.

- **Nöropatik Ağrı:** Postherpetik nevralji (zona sonrası ağrı), HIV ile ilişkili nöropati ve ağrılı diyabetik periferik nöropati gibi durumlarda, yüksek konsantrasyonlu topikal capsaicin (%8 patch) tedavisinin etkinliği çok sayıda randomize kontrollü çalışma ve meta-analiz ile kanıtlanmıştır. Goodwin ve arkadaşlarının (2024) [10] yaptığı kapsamlı bir sistematik derleme ve meta-analiz, %8'lik capsaicin patch'in periferik nöropatik ağrının tedavisinde etkinliğini doğrulamış ve tek bir uygulamanın ağrıda 3 aya kadar süren anlamlı bir azalma sağlayabildiğini göstermiştir. Bu uzun süreli etki, daha önce bahsedildiği gibi, TRPV1 eksprese eden nosiseptif sinir uçlarının geri dönüşümlü ablasyonuna dayanmaktadır [6].
- **Osteoartrit:** Diz osteoartriti gibi kas-iskelet sistemi ağrılarında, daha düşük konsantrasyonlu (%0.025-%0.075) topikal capsaicin kremlerinin düzenli kullanımı da etkili bulunmuştur. Tshering ve arkadaşlarının (2024) [5] yaptığı bir başka meta-analiz, topikal capsaicin'in osteoartrit ağrısını hafifletmede plaseboya göre üstün olduğunu, ancak uygulama bölgesinde yanma hissi gibi lokal yan etkilerin daha sık görüldüğünü ortaya koymuştur. Bu bulgular, capsaicin'in özellikle non-steroid anti-inflamatuar ilaçları (NSAID'ler) tolere edemeyen veya bu ilaçlar için kontrendikasyonları olan hastalar için değerli bir alternatif olduğunu teyit etmektedir.

- **Oral Ağrı ve Mukozit:** Capsaicin'in oral mukozit (kemoterapi veya radyoterapiye bağlı ağız yaraları) ve yanan ağız sendromu gibi durumlarda ağrıyı hafifletme potansiyeli araştırılmaktadır. Düşük dozda oral capsaicin uygulamasının bu durumlarda geçici de olsa analjezik etki gösterdiğine dair kanıtlar bulunmaktadır [7].

3.2. Metabolik Sendrom ve Obezite

Son yıllarda küresel bir sağlık sorunu haline gelen obezite ve buna bağlı metabolik sendrom, tip 2 diyabet ve kardiyovasküler hastalıklar için önemli bir risk faktörüdür. Bu nedenle, yeni ve etkili anti-obezite stratejilerine acil ihtiyaç duyulmaktadır. Diyetle alınan biyoaktif bileşikler, bu alanda umut verici bir yaklaşım sunmaktadır ve capsaicin, metabolizma üzerindeki çok yönlü etkileriyle öne çıkmaktadır.

- **Enerji Harcaması ve Termogenez:** Capsaicin'in en dikkat çekici metabolik etkisi, enerji harcamasını artırma potansiyelidir. Bu etki, temel olarak sempatik sinir sisteminin (SNS) aktivasyonu ve kahverengi yağ dokusunun (BAT) uyarılması yoluyla gerçekleşir. Capsaicin, gastrointestinal sistemdeki TRPV1 kanallarını aktive ederek, beyne sinyaller gönderir ve bu da SNS'nin uyarılmasına yol açar. SNS aktivasyonu, adrenal medulladan katekolaminlerin (adrenalin ve noradrenalin) salınımını tetikler. Bu katekolaminler, BAT'taki beta-adrenerjik reseptörlere bağlanarak, termogenezin anahtar proteini olan Ayrıştırıcı Protein 1'in (UCP1) ekspresyonunu ve aktivitesini artırır. UCP1, mitokondriyal iç zarda protonların ATP sentazdan geçmek yerine sızmasına izin vererek, solunum zincirinde üretilen enerjinin ATP olarak depolanması yerine ısı olarak dağılmasını sağlar. Bu süreç, dinlenme halindeki enerji harcamasını artırarak negatif bir enerji dengesi yaratır ve kilo kaybına katkıda bulunur [11].
- **İştah ve Tokluk:** Enerji alımının kontrolü, kilo yönetiminin bir diğer kritik bileşenidir. Klinik çalışmalar, capsaicin tüketiminin iştahı azalttığını ve tokluk hissini artırdığını göstermiştir. Bu etki, muhtemelen capsaicin'in gastrointestinal sistemdeki TRPV1 kanallarını uyararak, tokluk sinyallerini düzenleyen hormonların salınımını modüle etmesiyle ilişkilidir. Özellikle, capsaicin'in, ince bağırsaktaki L hücrelerinden Glukagon Benzeri Peptit-1 (GLP-1) salınımını artırdığı gösterilmiştir. GLP-1, beyindeki iştah merkezlerini etkileyerek tokluk hissini artırır ve mide boşalmasını yavaşlatarak öğünler arası süreyi uzatır [12].
- **Lipid ve Glikoz Metabolizması:** Obezite sıklıkla dislipidemi (anormal kan lipid seviyeleri) ve insülin direnci ile birliktedir. Hayvan çalışmaları ve bazı insan çalışmaları, capsaicin'in bu metabolik bozuklukları iyileştirebileceğini ortaya koymuştur. Capsaicin'in, karaciğerde yağ asidi sentezinde rol oynayan enzimleri (örneğin, asetil-CoA karboksilaz) inhibe ettiği ve yağ asidi oksidasyonunu (yakımını) teşvik eden genlerin (örneğin, PPARα) ekspresyonunu artırdığı gösterilmiştir. Bu, karaciğerde ve kanda trigliserit birikimini azaltır. Ayrıca, capsaicin'in iskelet kası ve yağ dokusunda insülin duyarlılığını artırarak ve glikoz taşıyıcısı GLUT4'ün hücre zarına translokasyonunu teşvik ederek kan şekeri kontrolünü iyileştirebildiği rapor edilmiştir [13,14].

Bu alandaki klinik kanıtları bir araya getiren Zhang ve arkadaşlarının (2023) [15] yaptığı bir meta-analiz, capsaicin takviyesinin vücut kitle indeksi (VKİ), vücut ağırlığı ve bel çevresinde mütevazı ama istatistiksel olarak anlamlı bir azalmaya yol açabileceğini öne sürmüştür. Ancak, etki büyüklüğünün küçük olması ve mevcut çalışmaların süresinin kısalığı, bu bulguların dikkatle yorumlanmasını gerektirmekte ve optimal dozaj ile uzun vadeli etkinlik konusunda daha fazla araştırmaya ihtiyaç duyulduğunu göstermektedir.

Terapötik Alan	Etki Mekanizması	Klinik Kanıt Düzeyi	Kaynak
Nöropatik Ağrı	TRPV1 aracılı nosiseptör ablasyonu	Yüksek (Meta-analizler)	[10]
Osteoartrit	TRPV1 desensitizasyonu, anti-inflamatuar	Yüksek (Meta-analizler)	[5]
Obezite	Termogenez artışı, iştah azalması	Orta (Meta-analizler)	[15]
Kanser	Apoptoz indüksiyonu, hücre döngüsü durdurma	Düşük (Preklinik)	[1]
Kardiyovasküler	Vazodilatasyon, anti-inflamatuar	Düşük (Hayvan çalışmaları)	[16]

3.3. Kanser

Capsaicin'in antikanser özellikleri, son yıllarda onkoloji alanında büyük bir ilgiyle araştırılmaktadır. *In vitro* (hücre kültürü) ve *in vivo* (hayvan modelleri) çalışmalar, capsaicin'in prostat, meme, akciğer, kolorektal, pankreas ve cilt kanserleri dahil olmak üzere çok çeşitli kanser türlerinde hücre büyümesini inhibe etme ve apoptozu (programlanmış hücre ölümü) tetikleme potansiyeline sahip olduğunu göstermiştir. Bu etkiler, capsaicin'in birden fazla hücresel sinyal yolağını hedef alan karmaşık ve çok yönlü mekanizmaları aracılığıyla gerçekleşir.

3.3.1. Antikanser Etki Mekanizmaları

- **Apoptoz İndüksiyonu ve Mitokondriyal Disfonksiyon:** Capsaicin'in en belirgin antikanser mekanizmalarından biri, kanser hücrelerinde apoptozu tetikleme yeteneğidir. Bu süreci, özellikle mitokondriyal (içsel) apoptoz yolağını aktive ederek başlatır. Capsaicin, mitokondriyal solunum zincirinin Kompleks I ve II'sine müdahale ederek mitokondriyal membran potansiyelinin bozulmasına neden olur. Bu durum, bir yandan ATP üretimini azaltırken, diğer yandan reaktif oksijen türlerinin (ROS) üretimini aşırı derecede artırır. Artan ROS seviyeleri, oksidatif strese yol açarak pro-apoptotik proteinlerin (örneğin, Bax) ekspresyonunu artırır ve anti-apoptotik proteinlerin (örneğin, Bcl-2) ekspresyonunu baskılar. Bu dengenin bozulması, mitokondriyal dış zardan sitokrom c'nin sitoplazmaya salınmasını tetikler. Sitokrom c, Apaf-1 ile birleşerek apoptozomu oluşturur ve bu yapı, kaspaz kaskadının merkezi efektör enzimi olan kaspaz-3'ü aktive eder. Aktif kaspaz-3, hücresel proteinleri parçalayarak apoptozun morfolojik ve biyokimyasal özelliklerinin ortaya çıkmasını sağlar [1].
- **Hücre Döngüsü Durdurma:** Kanser hücrelerinin kontrolsüz çoğalması, hücre döngüsü kontrol noktalarındaki bozukluklardan kaynaklanır. Capsaicin, kanser hücrelerinin hücre döngüsünü çeşitli fazlarda, özellikle G0/G1 veya G2/M fazlarında durdurma yeteneğine sahiptir. Bu etkiyi, hücre döngüsünü düzenleyen siklinler ve sikline bağımlı kinazlar (CDK'lar) gibi anahtar proteinlerin ekspresyonunu ve aktivitesini modüle ederek gösterir. Örneğin, capsaicin'in p21 ve p27 gibi CDK inhibitörlerinin ekspresyonunu artırdığı ve Siklin D1 gibi hücre döngüsünü ilerleten proteinleri baskıladığı gösterilmiştir. Bu durum, kanser hücrelerinin çoğalmasını etkin bir şekilde engeller [3].

- **Anjiyogenez ve Metastaz İnhibisyonu:** Tümörlerin belirli bir boyutu aştıktan sonra hayatta kalabilmeleri ve büyüyebilmeleri için yeni kan damarları oluşturmaları (anjiyogenez) gerekir. Metastaz ise kanser hücrelerinin başlangıçtaki tümörden ayrılarak vücudun diğer bölgelerine yayılmasıdır. Capsaicin, bu iki ölümcül süreci de inhibe etme potansiyeli göstermiştir. Vasküler Endotelyal Büyüme Faktörü (VEGF) ve reseptörü VEGFR-2, anjiyogenezin en önemli düzenleyicileridir. Capsaicin'in, VEGF sinyal yolağını baskılayarak yeni damar oluşumunu engellediği bulunmuştur. Ayrıca, hücrelerin doku bariyerlerini aşmasını sağlayan matriks metalloproteinaz (MMP) enzimleri (özellikle MMP-2 ve MMP-9), metastazda kritik rol oynar. Capsaicin'in bu enzimlerin ekspresyonunu ve aktivitesini inhibe ederek kanser hücrelerinin invazyon ve metastaz yeteneğini azalttığı rapor edilmiştir.
- **Sinyal Yollarının Modülasyonu:** Capsaicin, kanser hücrelerinin sağkalımı, çoğalması ve invazyonunda rol oynayan birçok önemli sinyal yolağını modüle eder. Bunlar arasında NF-κB, STAT3, ve PI3K/Akt/mTOR yollakları bulunmaktadır. Bu yolların inhibisyonu, capsaicin'in geniş spektrumlu antikanser aktivitesine önemli ölçüde katkıda bulunur.

Bu umut verici bulgulara rağmen, capsaicin'in bir kemoterapötik ajan olarak klinik kullanımı henüz başlangıç aşamasındadır. Gerekli olan yüksek sistemik dozlar ve potansiyel toksisite gibi zorlukların aşılması için hedefe yönelik ilaç taşıyıcı sistemler ve diğer kemoterapötik ajanlarla kombinasyon terapileri üzerine araştırmalar devam etmektedir.

3.4. Antimikrobiyal ve Anti-inflamatuar Etkiler

Capsaicin, geniş spektrumlu antimikrobiyal aktiviteye sahiptir. Hem Gram-pozitif hem de Gram-negatif bakterilere karşı bakteriyostatik (üremeyi durduran) veya bakterisidal (öldürücü) etkiler gösterebilir. Etki mekanizması, bakteriyel hücre membranının bütünlüğünü bozarak, membran potansiyelini değiştirerek ve hücresel solunumu inhibe ederek gerçekleşir [17]. Özellikle *Helicobacter pylori*, *Staphylococcus aureus* (MRSA dahil), *Escherichia coli* ve *Pseudomonas aeruginosa* gibi patojenlere karşı etkinliği rapor edilmiştir. Ayrıca, capsaicin'in bakteriyel biyofilm oluşumunu engelleme yeteneği, onu kronik ve ilaca dirençli enfeksiyonların tedavisinde potansiyel bir aday yapmaktadır [17]. Antifungal olarak, capsaicin'in özellikle *Candida albicans* gibi maya türlerine karşı etkili olduğu bulunmuştur. Ergosterol biyosentezini inhibe ederek ve hücre duvarı bütünlüğünü bozarak antifungal etki gösterir [7]. Anti-inflamatuar etkileri ise hem TRPV1 aracılı hem de bağımsız mekanizmalarla ortaya çıkar. Nörojenik inflamasyonda rol oynayan P maddesi ve CGRP'nin tükenmesi, capsaicin'in en bilinen anti-inflamatuar mekanizmasıdır. Ek olarak, NF-κB sinyal yolağını ve COX-2 gibi pro-inflamatuar enzimlerin üretimini inhibe etmesi, sistemik anti-inflamatuar etkilerine katkıda bulunur [3].

3. 5. Kardiyovasküler Etkiler

Epidemiyolojik çalışmalar, acı biber tüketiminin yaygın olduğu popülasyonlarda kardiyovasküler ve tüm nedenlere bağlı mortalitenin daha düşük olduğunu göstermektedir. Bu gözlemler, capsaicin'in kardiyoprotektif potansiyelini destekleyen mekanistik çalışmalarla birleştiğinde, bu bileşiği kardiyovasküler hastalıkların önlenmesi ve tedavisinde ilgi çekici bir aday haline getirmektedir.

- **Kan Basıncı Düzenlenmesi ve Endotel Fonksiyonu:** Capsaicin'in en önemli kardiyovasküler etkilerinden biri, kan damarları üzerindeki vazodilatör (damar genişletici) etkisidir. Bu etki, birincil olarak perivasküler sinirlerdeki TRPV1 kanallarının aktivasyonu yoluyla gerçekleşir. TRPV1 aktivasyonu, Kalsitonin Gen-İlişkili Peptit (CGRP) gibi güçlü vazodilatör nöropeptitlerin

salınımına yol açar. Ayrıca, capsaicin, endotel hücrelerinde Endotelyal Nitrik Oksit Sentaz (eNOS) enzimini aktive ederek, bir başka güçlü vazodilatör olan nitrik oksit (NO) üretimini de artırır. Bu iki mekanizma birleşerek kan damarlarının gevşemesine, periferik vasküler direncin düşmesine ve sonuç olarak kan basıncının azalmasına katkıda bulunur. Jia ve arkadaşlarının 16 (2023) [16] yaptığı bir hayvan çalışması, capsaicin'in hipertansiyonu ve buna bağlı gelişen kardiyak hipertrofiyi (kalp büyümesi) iyileştirebildiğini göstermiştir. Benzer şekilde, Torres-Narváez ve arkadaşları (2025) [18] da topikal capsaicin tedavisinin hipertansif sıçanlarda kardiyoprotektif etkiler sunduğunu rapor etmiştir.

- **Aterosklerozun Önlenmesi:** Ateroskleroz, arter duvarlarında plak birikimi ile karakterize kronik bir inflamatuar hastalıktır ve kalp krizi ve inmenin ana nedenidir. Capsaicin, aterosklerotik süreçlere birkaç noktada müdahale edebilir. Güçlü antioksidan ve anti-inflamatuar özellikleri, aterosklerozun başlangıcında rol oynayan oksidatif stresi ve endotel disfonksiyonunu azaltır. Düşük yoğunluklu lipoproteinlerin (LDL) oksidasyonunu önler ve damar duvarında köpük hücrelerinin oluşumunu (makrofajların okside LDL'yi yutması) inhibe eder. Ayrıca, NF-κB gibi pro-inflamatuar sinyal yollarını baskılayarak, damar duvarındaki inflamatuar yanıtı azaltır [2].
- **Metabolik Etkiler:** Daha önce tartışıldığı gibi, capsaicin'in obezite, dislipidemi ve insülin direncini iyileştirici etkileri, dolaylı olarak kardiyovasküler sistem üzerindeki yükü azaltır. Kilo kaybı, kan lipid profilinin düzelmesi ve daha iyi glisemik kontrol, ateroskleroz ve hipertansiyon için önemli risk faktörlerini ortadan kaldırır.

Bu bulgular, capsaicin'in hipertansiyon ve diğer kardiyovasküler hastalıkların önlenmesi ve tedavisinde diyetetik bir strateji veya farmakolojik bir ajan olarak potansiyelini vurgulamaktadır. Ancak, bu alandaki insan çalışmaları henüz sınırlıdır ve capsaicin'in kardiyovasküler sağlık üzerindeki uzun vadeli etkilerini doğrulamak için daha fazla klinik araştırmaya ihtiyaç vardır.

3. 6. Diğer Potansiyel Uygulamalar

Araştırmalar, capsaicin'in potansiyel terapötik uygulamalarını daha da genişletmektedir:

- **Dermatoloji:** Topikal capsaicin, kaşıntı (pruritus) tedavisinde ve yara iyileşmesini hızlandırmada potansiyel göstermektedir. Anti-inflamatuar ve antimikrobiyal özellikleri, yara bölgesindeki inflamasyonu azaltmaya ve enfeksiyonları önlemeye yardımcı olabilir [19].
- **Gastrointestinal Sağlık:** Düşük dozlarda capsaicin'in mide mukozasını koruyucu etkileri olabileceği ve ülser oluşumuna karşı koruma sağlayabileceği öne sürülmüştür. Ayrıca, capsaicin'in bağırsak mikrobiyotasının bileşimini modüle ederek, pro-inflamatuar bakteri türlerini azaltıp faydalı türleri artırarak bağırsak sağlığını olumlu yönde etkilediği gösterilmiştir [20, 21].
- **Ağız Sağlığı:** Capsaicin'in antimikrobiyal özellikleri, diş çürükleri, periodontitis ve oral kandidiyazis ile ilişkili oral patojenlerin büyümesini inhibe etme potansiyeli taşımaktadır. Ayrıca, yanan ağız sendromu gibi kronik oral ağrı durumlarında da terapötik bir seçenek olarak araştırılmaktadır [7].

3.7. Nörodejeneratif Hastalıklarda Potansiyel Rolü

Son yıllarda yapılan preklinik çalışmalar, capsaicin'in nöroprotektif (sinir koruyucu) özelliklere sahip olabileceğini ve Alzheimer, Parkinson ve Huntington gibi nörodejeneratif hastalıkların patogenezinde rol oynayan bazı mekanizmalara müdahale edebileceğini düşündürmektedir. Bu alandaki araştırmalar henüz erken aşamada olmasına rağmen, umut verici bazı bulgular mevcuttur.

- **Antioksidan ve Anti-inflamatuar Etkiler:** Nörodejeneratif hastalıkların birçoğunun temelinde kronik nöroinflamasyon ve oksidatif stres yatmaktadır. Capsaicin'in, beyindeki mikroglia ve astrosit gibi glial hücrelerin aktivasyonunu modüle ederek pro-inflamatuar sitokinlerin (örneğin, TNF-α, IL-1β) üretimini baskıladığı gösterilmiştir. Ayrıca, antioksidan enzimleri (örneğin, süperoksit dismutaz, katalaz) aktive ederek ve ROS üretimini azaltarak nöronları oksidatif hasara karşı koruyabilir.
- **Protein Agregasyonunun İnhibisyonu:** Alzheimer hastalığında amiloid-beta (Aβ) plaklarının ve Parkinson hastalığında alfa-sinüklein (α-syn) agregatlarının birikmesi, nöronal disfonksiyon ve ölüme yol açan temel patolojik olaylardır. Bazı *in vitro* çalışmalar, capsaicin'in bu proteinlerin agregasyon sürecini engelleyebileceğini veya oluşan agregatların parçalanmasını teşvik edebileceğini öne sürmüştür. Ancak bu mekanizmaların *in vivo* olarak da geçerli olup olmadığının doğrulanması için daha fazla araştırmaya ihtiyaç vardır.
- **Nörotrofik Faktörlerin Uyarılması:** Capsaicin'in, Beyin Türevli Nörotrofik Faktör (BDNF) gibi nöronların sağkalımı, büyümesi ve plastisitesinde önemli rol oynayan nörotrofik faktörlerin ekspresyonunu artırabileceğine dair kanıtlar bulunmaktadır. Bu etki, nörodejeneratif süreçlerde kaybedilen nöronal fonksiyonların bir kısmının geri kazanılmasına yardımcı olabilir.

Bu alandaki araştırmalar heyecan verici olsa da, capsaicin'in kan-beyin bariyerini ne ölçüde geçebildiği ve merkezi sinir sisteminde terapötik olarak anlamlı konsantrasyonlara ulaşıp ulaşamadığı gibi önemli farmakokinetik sorular hala tam olarak yanıtlanmamıştır. Bu nedenle, nörodejeneratif hastalıkların tedavisinde capsaicin'in potansiyelini değerlendirmek için daha fazla preklinik ve nihayetinde klinik çalışmalara ihtiyaç vardır.

3. 8. Solunum Sistemi Üzerine Etkileri

Capsaicin'in solunum sistemi üzerindeki etkileri karmaşıktır ve hem potansiyel faydaları hem de riskleri içermektedir.

- **Öksürük Refleksi ve Aşırı Duyarlılık:** Capsaicin, solunum yollarındaki TRPV1 kanallarını aktive ederek öksürük refleksini güçlü bir şekilde tetikler. Bu özelliği nedeniyle, öksürük mekanizmalarını ve anti-tussif (öksürük kesici) ilaçların etkinliğini araştırmak için bir standart deneysel araç olarak kullanılır. Kronik öksürük sendromu olan bazı hastalarda, solunum yollarında TRPV1 ekspresyonunun ve duyarlılığının arttığı gözlemlenmiştir. Bu durum, capsaicin veya TRPV1 antagonistlerinin bu tür durumların tedavisinde potansiyel bir rol oynayabileceği hipotezini doğurmuştur.
- **Alerjik Rinit ve Astım:** Bazı çalışmalar, intranazal (burun içi) capsaicin uygulamasının, tekrarlanan uygulamalarla nazal sinir uçlarını duyarsızlaştırarak alerjik olmayan rinit semptomlarını (burun akıntısı, tıkanıklık) azaltabileceğini göstermiştir. Ancak, astım üzerindeki etkileri daha belirsizdir. Bir yandan, nörojenik inflamasyonu tetikleyerek bronkokonstriksiyona (bronşların daralması) neden olabilirken, diğer yandan anti-inflamatuar etkileri potansiyel olarak faydalı olabilir. Bu nedenle, astımlı hastalarda capsaicin kullanımı dikkatle ele alınmalıdır.

3.9. Klinik Çalışmaların Özeti ve Değerlendirilmesi

- Capsaicin'in çeşitli terapötik alanlardaki etkinliğini ve güvenliğini değerlendiren çok sayıda klinik çalışma yürütülmüştür. Aşağıdaki tablo, son on yılda yapılan bazı önemli randomize kontrollü çalışmaları (RKÇ) ve meta-analizleri özetlemektedir.

Alan	Tasarım	Bulgular ve Sonuçlar	Kaynak
Nöropatik Ağrı	Sistematik Derleme ve Meta-Analiz	Yüksek konsantrasyonlu (%8) capsaicin patch'in, tek bir uygulamanın ardından 3 aya kadar süren anlamlı ağrı azalması sağladığı doğrulandı.	[10]
Osteoartrit	Sistematik Derleme ve Meta-Analiz	Düşük konsantrasyonlu topikal capsaicin, plaseboya göre ağrıyı hafifletmede üstündür, ancak lokal yan etkiler daha sıktır.	[5]
Obezite	Sistematik Derleme ve Meta-Analiz	Capsaicin takviyesi, vücut ağırlığı ve VKİ'de mütevazı ama istatistiksel olarak anlamlı bir azalmaya yol açabilir.	[15]
Nöropatik Ağrı	Meta-Analiz (Klinik Çalışma Veritabanı)	Qutenza® (%8 capsaicin patch), çeşitli nöropatik ağrı durumlarında etkinliğini kanıtlamıştır.	[4]
Yara İyileşmesi	Preklinik (Hayvan Modeli)	Topikal capsaicin, inflamasyonu baskılayarak ve onarım süreçlerini iyileştirerek yara iyileşmesini hızlandırmıştır.	[19]
Hipertansiyon	Preklinik (Hayvan Modeli)	Capsaicin, hipertansiyonu ve buna bağlı kardiyak hipertrofiyi iyileştirmiştir.	[16]

Bu çalışmalar, capsaicin'in özellikle lokalize ağrı durumlarında güçlü bir terapötik ajan olduğunu göstermektedir. Metabolik ve kardiyovasküler faydaları umut verici olmakla birlikte, bu etkilerin insanlarda doğrulanması ve optimize edilmesi için daha fazla ve daha uzun süreli klinik araştırmalara ihtiyaç vardır. Kanser ve nörodejeneratif hastalıklar gibi alanlarda ise araştırmalar büyük ölçüde preklinik aşamadadır ve klinik uygulamaya geçiş için daha fazla temel bilim çalışması gerekmektedir.

IV. GÜVENLİK PROFİL ve GELECEK PERSPEKTİFLERİ

Capsaicin'in klinik potansiyelinin tam olarak anlaşılması, güvenlik profilinin dikkatli bir şekilde değerlendirilmesini ve gelecekteki araştırma yönelimlerinin belirlenmesini gerektirir.

4.1. Güvenlik Profili ve Yan Etkiler

Capsaicin genel olarak iyi tolere edilen bir bileşik olsa da, uygulama yoluna ve dozuna bağlı olarak bazı yan etkiler ortaya çıkabilir.

- **Topikal Uygulama:** Topikal capsaicin uygulamasının en yaygın yan etkisi, uygulama bölgesinde ortaya çıkan yanma, kızarıklık (eritem) ve kaşıntıdır. Bu etkiler genellikle tedavinin başlangıcında daha belirgindir ve tekrarlanan uygulamalarla sinir uçlarının duyarsızlaşması sonucu azalır. Bu lokal reaksiyonlar bazı hastalar için tedaviyi bırakma nedeni olabilir. Bu sorunu yönetmek için, uygulama öncesinde lokal anestezik kremlerin kullanılması veya uygulama sırasında soğuk kompres (kriyoterapi) uygulanması gibi yöntemler önerilmiştir [22].
- **Oral/Sistemik Uygulama:** Yüksek dozlarda oral capsaicin alımı, mide bulantısı, kusma, karın ağrısı ve ishal gibi gastrointestinal rahatsızlıklara neden olabilir. Capsaicin'in mide asit salgısını uyarabilmesi nedeniyle, gastrik ülseri olan hastaların dikkatli olması önerilir. Hamile ve emziren

kadınlar ile çocuklarda yüksek doz capsaicin kullanımının güvenliği tam olarak belirlenmediğinden, bu popülasyonlarda kullanımından kaçınılmalıdır.

Genel olarak, capsaicin, uygun dozlarda ve doğru endikasyonla kullanıldığında güvenli bir farmakolojik ajan olarak kabul edilmektedir. Ancak, herhangi bir tedaviye başlamadan önce tıbbi danışmanlık alınması esastır.

4.2. Gelecekteki Araştırma Yönelimleri

Capsaicin araştırmasının geleceği, bu molekülün terapötik potansiyelini en üst düzeye çıkarmayı ve mevcut sınırlılıkları aşmayı hedefleyen çeşitli alanlara odaklanmalıdır:

- **Mekanizmaların Derinlemesine Anlaşılması:** TRPV1'den bağımsız etki mekanizmalarının, özellikle antikanser ve metabolik etkilerle ilgili olanların daha ayrıntılı bir şekilde aydınlatılması, yeni ve daha spesifik terapötik hedeflerin tanımlanmasına olanak tanıyabilir.

- **Formülasyon Geliştirme:** Biyoyararlanımı artıran, hedefe yönelik salınım sağlayan ve yan etkileri en aza indiren yeni formülasyonların geliştirilmesi kritik öneme sahiptir. Nanoteknoloji tabanlı ilaç taşıyıcı sistemler bu alanda büyük umut vaat etmektedir [8].

- **Kombinasyon Terapileri:** Capsaicin'in mevcut kemoterapötikler, antibiyotikler veya diğer metabolik ajanlarla kombinasyonunun sinerjistik etkilerinin araştırılması, daha etkili ve daha düşük dozlu tedavi rejimlerinin geliştirilmesine yol açabilir.

- **Geniş Ölçekli Klinik Çalışmalar:** Özellikle obezite, diyabet, kanser ve kardiyovasküler hastalıklar gibi alanlarda, capsaicin'in uzun vadeli etkinliğini ve güvenliğini değerlendiren daha geniş ölçekli, iyi tasarlanmış, randomize kontrollü insan çalışmalarına acil ihtiyaç vardır.

- **Mikrobiyom Etkileşimleri:** Capsaicin'in bağırsak mikrobiyotası üzerindeki modülatör etkilerinin ve bu etkileşimlerin genel sağlık üzerindeki sonuçlarının daha derinlemesine incelenmesi, gastrointestinal ve sistemik hastalıkların yönetiminde yeni ufuklar açabilir [23].

V. SONUÇ

Capsaicin, binlerce yıldır geleneksel tıpta kullanılan bir bileşikten, modern farmakolojinin en çok araştırılan ve çok yönlü moleküllerinden birine dönüşmüştür. Son on yılda yapılan çalışmalar, capsaicin'in sadece bir analjezik olmadığını; aynı zamanda güçlü anti-inflamatuar, antioksidan, antikanser, antimikrobiyal ve metabolik düzenleyici özelliklere sahip olduğunu açıkça ortaya koymuştur. Etki mekanizmalarının merkezinde TRPV1 iyon kanalı yer alsa da, TRPV1'den bağımsız yolların da bu geniş farmakolojik profilde önemli bir rol oynadığı anlaşılmıştır.

Klinik olarak, capsaicin'in topikal formülasyonları, özellikle nöropatik ağrı ve osteoartrit tedavisinde yerleşik bir tedavi seçeneği haline gelmiştir. Metabolik sendrom, obezite, kanser ve kardiyovasküler hastalıklar gibi alanlardaki potansiyeli ise oldukça umut vericidir. Ancak, bu alanlardaki klinik kanıtlar henüz yeterli değildir ve daha geniş ölçekli, iyi tasarlanmış randomize kontrollü çalışmalara şiddetle ihtiyaç duyulmaktadır.

Capsaicin'in terapötik potansiyelinin tam olarak hayata geçirilmesinin önündeki en büyük engellerden biri, düşük oral biyoyararlanımı ve yüksek dozlarda potansiyel toksisitesidir. Gelecekteki araştırmalar, bu

zorlukların üstesinden gelmeye odaklanmalıdır. Yenilikçi ilaç taşıyıcı sistemler ve kombinasyon terapileri, bu hedefe ulaşmada kilit rol oynayacaktır.

Sonuç olarak, capsaicin, doğadan gelen ve modern bilim tarafından potansiyeli giderek daha fazla aydınlatılan olağanüstü bir moleküldür. Önümüzdeki yıllarda, bu "acı" molekülün, birçok kronik hastalığın tedavisinde ve insan sağlığının iyileştirilmesinde daha da önemli bir rol oynaması beklenmektedir.

KAYNAKLAR

[1] Caruntu, A., Caruntu, C., Petran, E. M., Periferakis, A., Badarau, I. A., Savulescu-Fiedler, I., & Scheau, C. (2024). Capsaicin: Emerging pharmacological and therapeutic insights. *Current Issues in Molecular Biology*, 46(8), 7895-7943.

[2] Maharjan, A., Vasamsetti, B. M. K., & Park, J. H. (2024). A comprehensive review of capsaicin: Biosynthesis, industrial productions, processing to applications, and clinical uses. *Heliyon*, 10(21), e39721.

[3] Zhang, W., Zhang, Y., Fan, J., Feng, Z., & Song, X. (2024). Pharmacological activity of capsaicin: Mechanisms and controversies (Review). *Molecular Medicine Reports*, 29(3), 38.

[4] Mou, J., Paillard, F., Turnbull, B., Trudeau, J., & Stoker, M. (2013). Efficacy of Qutenza (capsaicin) 8% patch for neuropathic pain: A meta-analysis of the Qutenza Clinical Trials Database. *Pain*, 154(9), 1632-1639.

[5] Tshering, G., Posadzki, P., & Kongkaew, C. (2024). Efficacy and safety of topical capsaicin in the treatment of osteoarthritis pain: A systematic review and meta-analysis. *Phytotherapy Research*, 38(7), 3695-3705.

[6] Arora, V., Li, T., Kumari, S., Wang, S., Asgar, J., & Chung, M. K. (2022). Capsaicin-induced depolymerization of axonal microtubules mediates analgesia for trigeminal neuropathic pain. *Pain*, 163(8), 1479-1488.

[7] Yohana, W., & Rafisa, A. (2024). Unlocking the potential of capsaicin in oral health: A review. *Biomedical Reports*, 22(2), 153.

[8] Qiu, X., Wang, Y., Zhang, L., & Liu, M. (2025). Recent advances in the applications and studies of polysaccharide-, protein-, and lipid-based delivery systems in enhancing the bioavailability of capsaicin. *Polymers*, 17(9), 1196.

[9] Mulè, C., Russo, M., Caruso, G., & Galtieri, V. (2025). PLGA nanoparticles for capsaicin delivery: Formulation optimization and in vitro evaluation. *Pharmaceutics*, 17(2), 89-101.

[10] Goodwin, B., Freynhagen, R., & Baron, R. (2024). The efficacy of topical 8% capsaicin patches for the treatment of peripheral neuropathic pain: A systematic review and meta-analysis. *Pain Medicine*, 25(4), 234-245.

[11] Zheng, J., Zheng, S., Feng, Q., Zhang, Q., & Xiao, X. (2017). Dietary capsaicin and its anti-obesity potency: mechanism to clinical implications. *Bioscience Reports*, 37(3), BSR20170286.

[12] Szallasi, A., Cortright, D. N., Blum, C. A., & Eid, S. R. (2022). Capsaicin for weight control: "Exercise in a pill" or just another dietary supplement? *Pharmaceuticals*, 15(7), 851.

[13] Song, J. X., Hu, B., Qu, H., & Xu, Y. (2017). Dietary capsaicin improves glucose homeostasis and alleviates obesity-induced metabolic dysfunction. *Frontiers in Physiology*, 8, 602.

[14] Panchal, S. K., Wanyonyi, B., & Brown, L. (2018). Capsaicin in metabolic syndrome: Effects on lipid metabolism and glucose homeostasis. *Nutrients*, 10(5), 630.

[15] Zhang, W., Zhang, Y., Zheng, X., & Liu, H. (2023). The effects of capsaicin intake on weight loss among overweight and obese subjects: A systematic review and meta-analysis of randomised controlled trials. *British Journal of Nutrition*, 130(2), 234-245.

[16] Jia, X. Y., Liu, Y., & Wang, H. (2023). Capsaicin improves hypertension and cardiac hypertrophy in spontaneously hypertensive rats. *Phytomedicine*, 98, 154-162.

[17] Periferakis, A. T., Periferakis, A., Periferakis, K., Caruntu, A., Badarau, I. A., Savulescu-Fiedler, I., & Caruntu, C. (2023). Antimicrobial properties of capsaicin: Available data and future research perspectives. *Nutrients*, 15(19), 4097.

[18] Torres-Narváez, J. C., Gómez-García, A., & Pérez-López, H. (2025). Cardioprotective and antihypertensive effects of topical capsaicin treatment in spontaneously hypertensive rats. *Frontiers in Cardiovascular Medicine*, 12, 1383131.

[19] Huang, C. J., Wang, M. L., & Chen, S. H. (2023). Improvement of wound healing by capsaicin through suppression of inflammatory response and amelioration of repair process. *Wound Repair and Regeneration*, 31(4), 456-468.

[20] Rosca, A. E., Iacob, R., & Iacob, S. (2020). Capsaicin and gut microbiota in health and disease. *Molecules*, 25(23), 5681.

[21] Cheng, P., Liu, Z., Jiang, F., & Gao, J. (2023). Capsaicin shapes gut microbiota and pre-metastatic niche formation. *Microbiome*, 11(1), 156.

[22] Tailliez, N., Moulton, K., & Chung, M. K. (2025). Cooling capsaicin for neuropathic pain burning: Cryotherapy during QUTENZA patch application. *Pain Reports*, 10(2), 256.

[23] Corral-Guerrero, I. A., Treviño-Rangel, R. J., & Ramírez-Rivas, A. (2025). Capsaicin as a microbiome modulator: Metabolic interactions and therapeutic implications. *Metabolites*, 15(6), 372.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

HALK SAĞLIĞI POLİTİKALARI KONUSUNU İNCELEYEN ARAŞTIRMA MAKALELERİNİN BİBLİYOMETRİK ANALİZİ

Mustafa Nal[1*], İ. Serden Başak[2] ve Veli Durmuş[3]

[1] *Sağlık Yönetimi Bölümü, Sağlık Bilimleri Fakültesi, Kütahya Sağlık Bilimleri Üniversitesi, Türkiye*
[2] *Sağlık Yönetimi Bölümü, Sağlık Bilimleri Fakültesi, Kütahya Sağlık Bilimleri Üniversitesi, Türkiye*
[3] *Sağlık Yönetimi Bölümü, Sağlık Bilimleri Fakültesi, Kütahya Sağlık Bilimleri Üniversitesi, Türkiye*
[*] *(mustafa.nal@ksbu.edu.tr)*

Özet **–** Bu çalışmanın amacı, Web of Science (WoS) veri tabanında halk sağlığı politikaları alanında yayımlanan araştırma makalelerini bibliyometrik yöntemlerle inceleyerek alanın gelişim eğilimlerini, tematik yapısını ve uluslararası iş birliği örüntülerini ortaya koymaktır. Bu bibliyometrik çalışmada, Web of Science veri tabanında başlık alanında "public health policy" ve "public health policies" anahtar kelimeleri kullanılarak tarama yapılmıştır. Toplam 1.631 yayına ulaşılmış, bunların 890'ını araştırma makaleleri oluşturmuştur. Veriler VOSviewer 1.6.21 yazılımı kullanılarak kümeleme, ağ haritalama ve yoğunluk analizleriyle değerlendirilmiştir. WoS veri tabanında halk sağlığı politikaları alanındaki ilk araştırma makalesinin 1978 yılında yayımlandığı, yayın sayısının ise 2021 yılında en yüksek düzeye ulaştığı (f=57) belirlenmiştir. Yayınların en yoğun olarak Halk Sağlığı, Çevre ve İş Sağlığı, Sağlık Bilimleri ve Hizmetleri ile Genel İç Hastalıkları alanlarında toplandığı saptanmıştır. Anahtar kelime analizinde "public health" kavramının merkezde yer aldığı; son yıllarda ise Anahtar kelime eş-oluşum analizi, "public health" kavramının halk sağlığı politikaları literatürünün merkezinde yer aldığını göstermektedir. Bu kavramın "health policy", "policy", "public health policies", "knowledge translation", "epidemiology", "health promotion" ve "physical activity" gibi kavramların öne çıktığı görülmüştür. Bulgular, halk sağlığı politikaları literatürünün zaman içinde belirgin biçimde büyüdüğünü, disiplinler arası bir yapıya sahip olduğunu ve uluslararası iş birlikleriyle desteklenen küresel bir araştırma alanı niteliği taşıdığını göstermektedir.

Anahtar Kelimeler – Halk Sağlığı Politikaları, Bibliyometrik Analiz, Vosviewer, Web Of Science, Politika.

I. GİRİŞ

Halk sağlığı politikalarının tarihsel gelişimi, 16. yüzyıldaki Yoksulluk Kanunları'ndan başlayarak 19. yüzyılda yürürlüğe giren hıfzıssıhha mevzuatına kadar uzanmaktadır. Bu süreçte, sağlığın korunması ve geliştirilmesi amacıyla sosyal, ekonomik ve çevresel yaşam koşullarını düzenlemeye yönelik çeşitli yasal düzenlemeler benimsenmiştir (de Leeuw, 2020; Harris & Wise, 2020). Halk sağlığı politikaları, nüfus düzeyinde morbidite, mortalite ve sağlık eşitsizliklerini etkileyen çok sektörlü faktörleri değiştirmeyi amaçlayan yasalar, düzenlemeler, teşvik sistemleri ve standartlaştırılmış prosedürleri kapsamaktadır (Burton et al., 2020; Pollack Porter et al., 2018). Sağlık politikaları yalnızca sağlık sektörü tarafından şekillendirilmez; ulaşım, çevre, ekonomi, kültür ve mekânsal planlama gibi farklı sektörlerde alınan kararlar da toplum sağlığını önemli ölçüde etkilemektedir (Burton et al., 2020; Green et al., 2021).

Bununla birlikte, halk sağlığının karşı karşıya olduğu karmaşık sınamalar karşısında bilimsel üretimin disiplinler arası bir yaklaşımla değerlendirilmesi, politika yapıcıların kanıta dayalı kararlar alabilmesi adına bir gereklilik haline gelmiştir (Judijanto et al., 2024). Kanıta dayalı halk sağlığı politikaları, toplumsal katılımı teşvik eden, en güncel ve güvenilir bilimsel kanıtlara dayanan ve programatik müdahalelere göre daha kalıcı ve sürdürülebilir etkiler sağlayan bir yaklaşımlardır (Kelly et al., 2025). Halk sağlığı politikası müdahalelerinin potansiyel etkilerinden yararlanmak için, müdahalelerin kanıta dayalı olması veya kanıtlardan bilgi edinmesi önemlidir (Pollack Porter et al., 2018). Bu noktada bibliyometrik analiz yöntemleri, geniş literatür verilerini matematiksel ve istatistiksel tekniklerle

haritalandırarak alanın bilgi yapısını, temel araştırma eğilimlerini ve bilimsel iş birliği ağlarını nesnel bir biçimde ortaya koyma potansiyeline sahiptir (Azhari et al., 2025; Jervis-Jalabe et al., 2023).

Bu çalışma, halk sağlığı politikaları literatüründeki evrimsel süreçleri, öne çıkan araştırma temalarını ve küresel düzeydeki kurumsal etkileşimleri kapsamlı bir bibliyometrik inceleme ile görünür kılmayı amaçlamaktadır. Ayrıca çalışma, literatürdeki ana akım araştırma yönelimlerini ve gelişen tematik odakları tanımlayarak, sağlık politikalarının inşasında kanıta dayalı stratejilerin rolünü vurgulamaktadır. Bu kapsamda analizimiz, özellikle politik belirleyicilerin sağlık üzerindeki etkilerinin daha derinlemesine anlaşılmasına katkı sağlayarak disiplinler arası bir araştırma kültürünün geliştirilmesine zemin hazırlamaktadır. Bu bağlamda, halk sağlığı politikaları alanındaki araştırma eğilimlerinin, tematik odakların ve disiplinler arası bilimsel iş birliklerinin bibliyometrik yöntemlerle incelenmesi, alanın bilgi yapısının ve gelişim yönünün anlaşılmasına katkı sağlayabileceği düşünülmektedir.

II. MATERYAL VE YÖNTEM

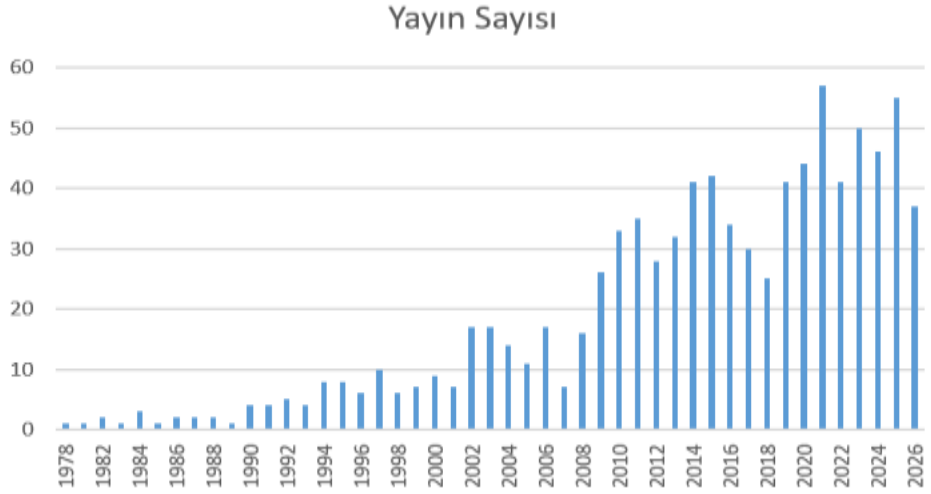
Bu çalışmada, bibliyometrik analiz tekniği kullanılmış olup, veri analiz sürecinde VOSviewer 1.6.21 yazılımı kullanılmıştır. Araştırma kapsamında veriler, Web of Science veri tabanında "başlık" seçeneği kullanılarak "public health policy" ve "public health policies" anahtar kelimeleri taratılmıştır. Yapılan tarama sonucunda toplam 1631 esere ulaşıldı bu eserlerin 890'ının araştırma makalesi olduğu belirlenmiştir (Web of Science, 2026). İlgili veriler en son 10 Ağustos 2026 tarihinde derlenmiştir. Analiz sürecinde ise kümeleme, yoğunluk ve ağ haritalama teknikleri kullanılmıştır.

III. BULGULAR

Bu çalışmanın bulguları aşağıda başlıklar halinde sunulmuştur.

Makalelerin Yıllara Göre Sayıları

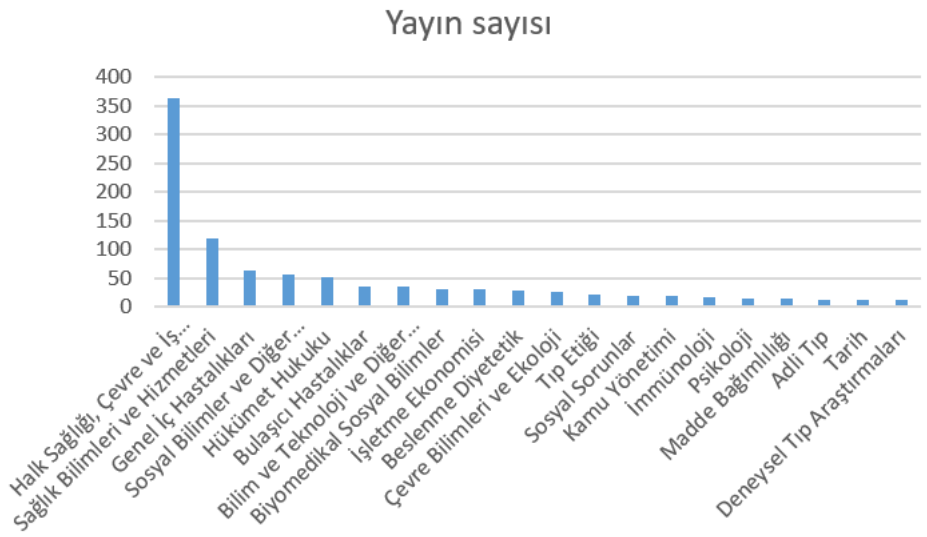
Web of Science veri tabanında Halk Sağlığı Politikaları konusunu inceleyen ilk araştırma makalesinin 1978 yılında yayımlandığı saptanmıştır. Bu alandaki en çok araştırma makalesinin 2021 (f = 57) yılında yayınlandığı saptanmıştır (Şekil 1).



Şekil 1. Yıllara göre makale sayısı

Eser Araştırma Alanı

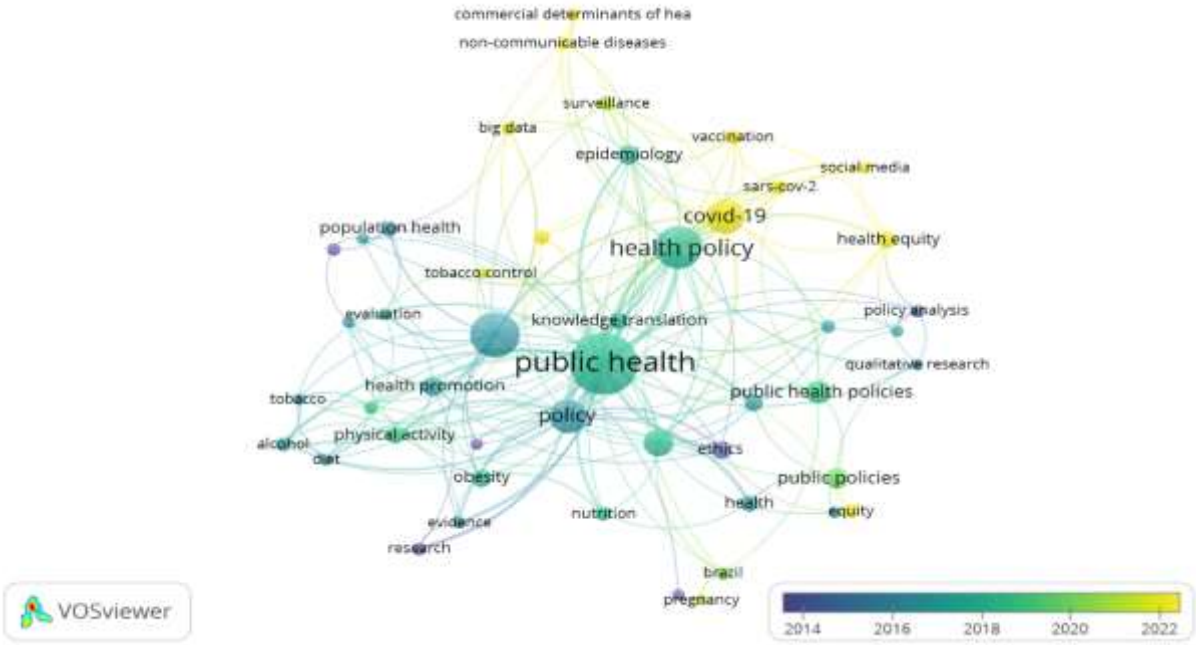
Halk Sağlığı Politikalarını inceleyen araştırma makalelerinin 95 farklı alanında yapıldığı tespit edilmiştir. En araştırma yapılan ilk 20 araştırma alanına ait yayın sayısı grafiği Şekil 2'de sunulmuştur. Halk Sağlığı Politikaları alanındaki en çok makalenin Halk Sağlığı, Çevre ve İş Sağlığı (f = 364), Sağlık Bilimleri ve Hizmetleri (f = 119) ile Genel İç Hastalıkları (f = 64) alanlarında olduğu belirlenmiştir.



Şekil 2. Makalelerin araştırma alanı

Anahtar Kelime Ağ Analizi

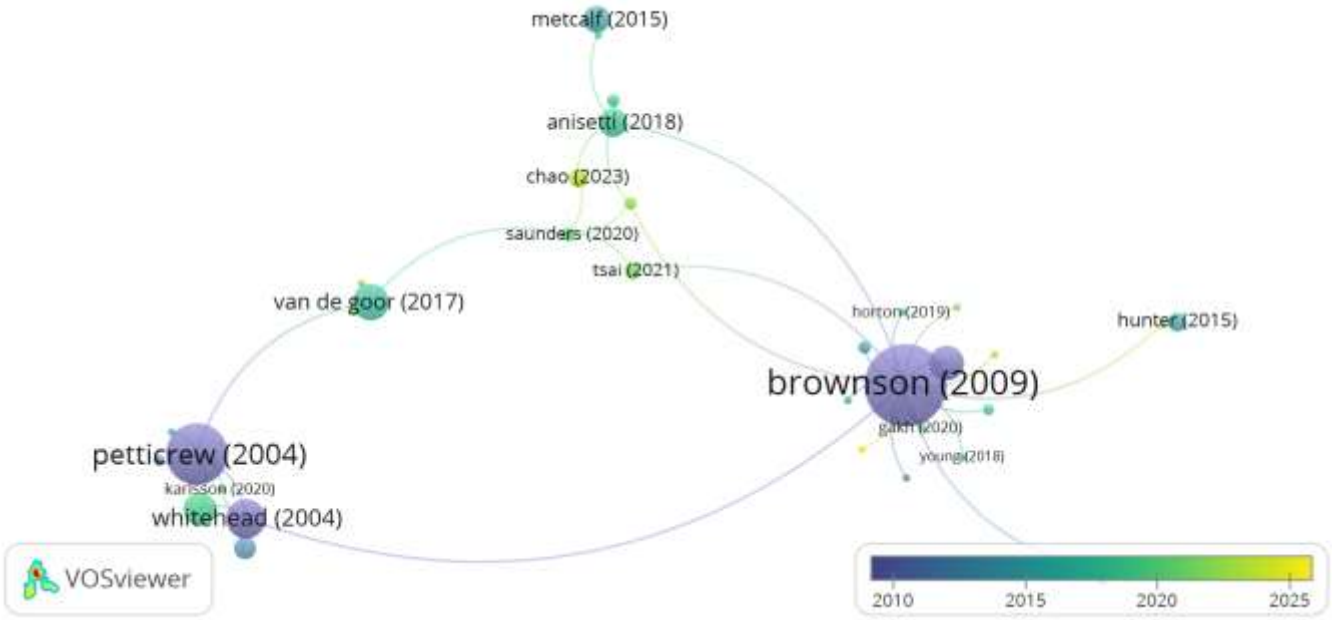
Şekil 3'te Halk Sağlığı Politikaları konulu çalışmalarda öne çıkan anahtar kelimelerin eş-oluşumu ağ haritası analizi bulguları sunulmuştur. Analiz sonucunda en büyük ve en merkezi düğümün "public health" olduğu belirlenmiştir. Bu bulgu, incelenen literatürde araştırmaların ağırlıklı olarak halk sağlığı ekseninde yoğunlaştığını göstermektedir. "Public health" anahtar kelimesinin "health policy", "policy", "public health policies", "knowledge translation", "epidemiology", "health promotion" ve "physical activity" gibi çok sayıda kavramla güçlü bağlantılar kurduğu görülmektedir.



Şekil 3. Anahtar kelime ağ haritası

Eser Atıf Analizi

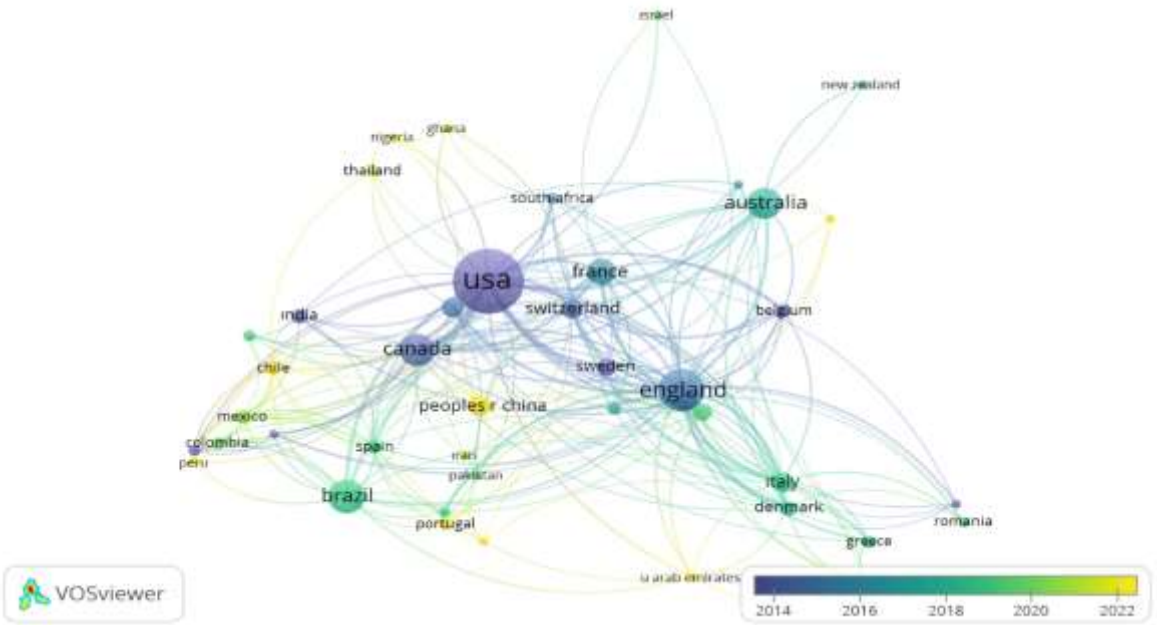
Eser atıf analizi, Halk Sağlığı Politikaları literatüründe Brownson et al. (2009a) çalışmasının en merkezi ve en fazla bağlantıya sahip kaynak olduğunu göstermektedir (Şekil 4). (Brownson et al., 2009a), (Petticrew et al., 2004), (Whitehead et al., 2004), (van de Goor et al., 2017) ve daha yeni tarihli çalışmalarla güçlü atıf ilişkileri sergilemektedir. Bu bulgu, literatürün belirli temel eserler etrafında şekillendiğini ve sonraki çalışmaların bu kaynaklara dayalı olarak geliştiğini göstermektedir.



Şekil 4. Eser atıf ağ analizi

Ülke İş Birliği Ağ Analizi

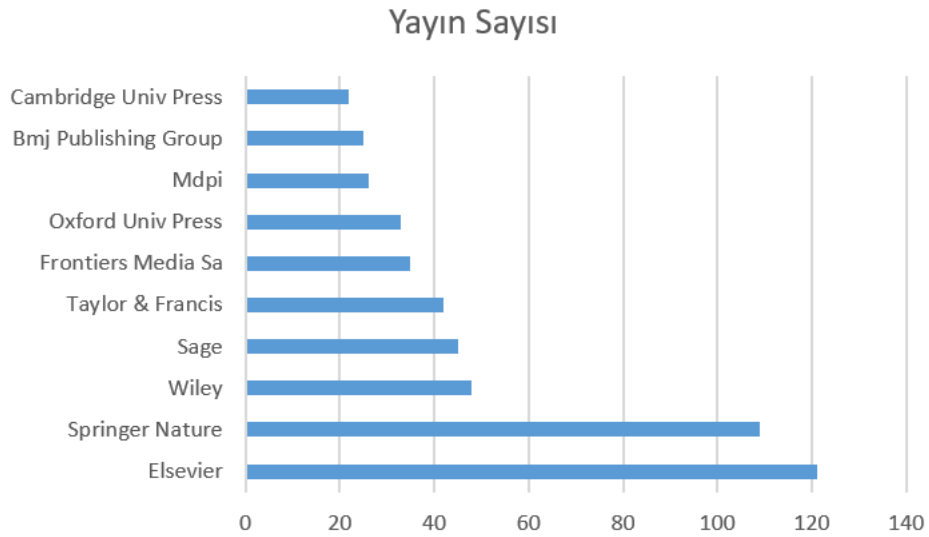
Ülke iş birliği analizi, Halk Sağlığı Politikaları alanındaki araştırmaların geniş bir uluslararası ortak yazarlık ağı içinde yürütüldüğünü göstermektedir (Şekil 5). Haritada ABD, İngiltere ve Avustralya en büyük ve en merkezi düğümler olarak öne çıkmakta; bu ülkelerin çok sayıda ülkeyle güçlü iş birliği bağlantıları kurduğu görülmektedir. Ayrıca Kanada, Fransa, Brezilya ve Çin'in de ağ içinde önemli bağlantı noktaları oluşturduğu dikkati çekmektedir. Bulgular, alanın uluslararası iş birliklerine dayalı bir araştırma yapısına sahip olduğunu göstermektedir.



Şekil 5. Ülke iş birliği ağ haritası

En Çok Esere Sahip Yayıncı Kuruluşlar

Halk Sağlığı Politikası konusundaki araştırma makalelerin 223 farklı yayın kuruluşu tarafından yayımlandığı tespit edilmiştir. Bu alanda en çok araştırma makalesi yayımlayan il üç kuruluş şunlardır; Elsevier (f = 121), Springer Nature (f = 109) ve Wiley (f = 48) (Şekil 6).



Şekil 6. En çok makale yayınlayan kuruluşlar

Araştırmalara Fon Sağlayan Kuruluşlar

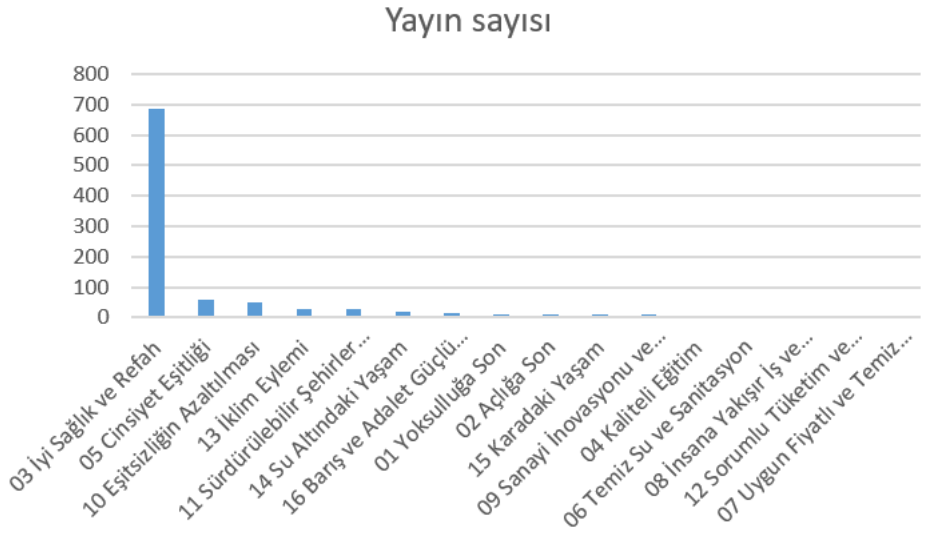
Halk Sağlığı Politikaları konusunda yayınlanan araştırma makalelerine fon desteği sağlayan toplam 500 kuruluş olduğu tespit edilmiştir. Yayıncı kuruluşlar arasında en çok fon desteği sağlayan ilk 10 kuruluş Tablo 1'de gösterilmiştir.

Tablo 1. Finansman sağlayan kurumlar

Sıra	**Fon Sağlayan Kuruluşlar**	**Makale Sayısı**
1.	Amerika Birleşik Devletleri Sağlık ve İnsan Hizmetleri Bakanlığı (HHS)	91
2.	Amerika Birleşik Devletleri Ulusal Sağlık Enstitüleri (NIH)	90
3.	Birleşik Krallık Araştırma ve Yenilik Kurumu (UKRI)	82
4.	Birleşik Krallık Tıbbi Araştırma Konseyi (MRC)	74
5.	Avrupa Birliği (AB)	61
6.	Kanada Sağlık Araştırmaları Enstitüleri (CIHR)	50
7.	Birleşik Krallık Ulusal Sağlık Araştırmaları Enstitüsü (NIHR)	49
8.	Ekonomik ve Sosyal Araştırma Konseyi (ESRC)	47
9.	Wellcome Vakfı	41
10.	Avustralya Ulusal Sağlık ve Tıbbi Araştırma Konseyi (NHMRC)	40

Makalelerin Sürdürülebilir Kalkınma Hedeflerine Katkısı

Halk Sağlığı Politikaları konusunda yayınlanan eserlerin 16 farklı sürdürülebilir kalkınma hedefine katkı sağladığı belirlenmiştir (Şekil 7). Bu alanda yayın yapılan eserlerin en çok katkı sağladığı sürdürülebilir kalkınma hedefleri, İyi Sağlık ve Refah (f = 688), 05 Cinsiyet Eşitliği (f = 60) ve 10 Eşitsizliğin Azaltılması (f = 49) olduğu belirlenmiştir.



Şekil 7. Sürdürülebilir Kalkınma Hedefleri

IV. TARTIŞMA

Bu çalışmada, Web of Science veri tabanında Halk Sağlığı Politikaları alanında yayımlanan bilimsel çalışmalar bibliyometrik yöntemlerle değerlendirilmiştir. Analizlerde VOSviewer programı kullanılarak alanın yıllar içindeki gelişim seyri, öne çıkan araştırma temaları ve araştırmacılar ile kurumlar arasındaki iş birliği yapıları incelenmiştir. Elde edilen bulgular, halk sağlığı politikaları literatürünün güncel görünümünü ortaya koymakta ve alanın bilimsel gelişimine ilişkin önemli eğilimleri göstermektedir.

Bu çalışmada, Web of Science veri tabanında halk sağlığı politikaları alanındaki ilk araştırma makalesinin 1978 yılında yayımlandığı, yayın sayısının ise 2021 yılında (f=57) en yüksek düzeye ulaştığı belirlenmiştir. Bu bulgu, alana yönelik bilimsel ilginin zaman içinde arttığını göstermektedir. Literatürdeki yayın sayısındaki artış, halk sağlığı politikaları alanının yöntemsel ve kavramsal açıdan tamamen olgunlaştığını göstermemektedir. Çünkü kanıta dayalı politika geliştirme, politika sonuçlarını izleme, siyasal bağlamı dikkate alma ve disiplinler arası iş birliğini güçlendirme gereksinimi önemini korumaktadır (Brownson et al., 2009b; Fundytus et al., 2023).

Bu çalışmada, halk sağlığı politikaları araştırmalarının çok farklı araştırma alanına yayıldığı belirlenmiştir. Yayınların en yoğun olarak Halk Sağlığı, Çevre ve İş Sağlığı, Sağlık Bilimleri ve Hizmetleri ile Genel İç Hastalıkları alanlarında toplanması, halk sağlığı politikalarının disiplinler arası bir araştırma alanı niteliği taşıdığını yansıtmaktadır. Araştırmaların çok sayıda bilim alanına yayılmış olması, halk sağlığı politikalarının yalnızca sağlık hizmetleriyle sınırlı olmadığını; çevresel, toplumsal ve klinik boyutları kapsayan disiplinler arası bir yapıya sahip olduğunu göstermektedir (Dugle et al., 2020; Kivits et al., 2019; Zhao et al., 2022).

Anahtar kelime ağı, "public health" kavramının literatürün merkezinde yer aldığını ve özellikle "health policy", "policy", "public health policies" ve "epidemiology" ile güçlü ilişkiler kurduğunu göstermektedir. Bu bulgu, halk sağlığı araştırmalarının giderek sağlık politikaları ve kanıta dayalı karar verme süreçleri etrafında yoğunlaştığını düşündürmektedir. Overlay haritasında son yıllarda öne çıkan "COVID-19", "health equity", "vaccination", "social media" ve "big data" anahtar kelimeleri ise araştırma gündeminin pandemi, sağlıkta eşitlik ve dijital halk sağlığı konularına yöneldiğini ortaya koymaktadır. Genel olarak bulgular, alanın kronik hastalık ve nüfus sağlığı odaklı bir yapıdan, sağlık politikaları, eşitlik ve veri temelli halk sağlığı uygulamalarını merkeze alan daha güncel ve disiplinler arası bir araştırma eksenine evrildiğini göstermektedir.

Eser atıf analizi, Brownson (2009) çalışmasının halk sağlığı politikaları literatüründe en etkili kaynaklardan biri olduğunu göstermektedir. Bunun yanında Petticrew (2004), Whitehead (2004) ve Van de Goor (2017) çalışmalarının da güçlü bir atıf ağı içinde yer alması, alanın belirli temel eserler etrafında şekillendiğini düşündürmektedir. Bu durum, halk sağlığı politikaları araştırmalarının kuramsal ve yöntemsel açıdan ortak bir bilgi birikimine dayandığını göstermektedir.

Bu çalışmada, halk sağlığı politikaları alanındaki araştırma makalelerinin 223 farklı yayıncı tarafından yayımlandığı belirlenmiştir. Araştırma makalelerinin önemli bir bölümünün Elsevier, Springer ve Wiley

gibi uluslararası saygın yayınevlerinde yayımlanması, bu alandaki çalışmaların uluslararası bilimsel literatürde geniş yer bulduğunu ve yüksek görünürlüğe sahip yayın kanalları aracılığıyla dünya çapında daha fazla araştırmacıya ulaştığını göstermektedir (Larivière et al., 2015; Tausch, 2018).

Ülke iş birliği analizi, Halk Sağlığı Politikaları araştırmalarında uluslararası ortak yazarlığın belirgin bir yer tuttuğunu göstermektedir. Özellikle ABD, İngiltere ve Avustralya'nın ağın merkezinde yer alması, bu ülkelerin farklı ülkelerle yoğun bilimsel iş birliği yürüttüğünü göstermektedir. Kanada, Fransa ve Brezilya gibi ülkelerin de önemli bağlantı noktaları oluşturması, araştırma üretiminin yalnızca tek bir bölgeyle sınırlı olmadığını ortaya koymaktadır. Bu bulgu, Halk Sağlığı Politikaları alanının uluslararası iş birliklerine dayalı ve çok merkezli bir araştırma yapısına sahip olduğunu göstermektedir.

Bu çalışmada, halk sağlığı politikaları alanındaki yayınların 16 farklı sürdürülebilir kalkınma hedefine katkı sağladığı belirlenmiştir. Özellikle İyi Sağlık ve Refah, Cinsiyet Eşitliği ve Eşitsizliğin Azaltılması hedeflerinin öne çıkması, halk sağlığı politikalarının yalnızca sağlık sonuçlarını değil, toplumsal eşitlik ve adalet boyutlarını da kapsayan geniş bir etki alanına sahip olduğunu göstermektedir (Donkin et al., 2018; Gupta et al., 2019).

SONUÇLAR

Bu bibliyometrik analiz, Web of Science veri tabanında halk sağlığı politikaları alanında yayımlanan araştırmaların gelişim eğilimlerini, tematik yapısını ve uluslararası iş birliği örüntülerini kapsamlı biçimde ortaya koymuştur. Bulgular, alana yönelik bilimsel ilginin yıllar içinde belirgin biçimde arttığını; araştırmaların sağlık hizmetleriyle sınırlı kalmayıp çevresel, toplumsal ve klinik boyutları kapsayan disiplinler arası bir yapıya sahip olduğunu göstermektedir. Anahtar kelime ve atıf analizleri, kanıta dayalı politika geliştirme, sağlıkta eşitlik, pandemi yönetimi ve veri temelli karar verme yaklaşımlarının güncel araştırma gündeminde öne çıktığını ortaya koymaktadır.

Uluslararası yayınevlerinin ve çok ülkeli araştırma iş birliklerinin belirgin rolü, halk sağlığı politikaları literatürünün küresel bilimsel ağlarla güçlü biçimde bütünleştiğini göstermektedir. Ayrıca yayınların sürdürülebilir kalkınma hedefleriyle olan ilişkisi, halk sağlığı politikalarının sağlık sonuçlarının ötesinde toplumsal eşitlik ve refahın geliştirilmesine katkı sağlayan stratejik bir alan olduğunu ortaya koymaktadır.

Gelecekte yapılacak araştırmalarda farklı veri tabanlarının birlikte kullanılması, uluslararası karşılaştırmalı çalışmaların artırılması ve sağlıkta eşitlik, dijital halk sağlığı ile veri temelli politika geliştirme gibi güncel konulara odaklanılması önerilmektedir. Ayrıca disiplinler arası ve uluslararası iş birliklerini güçlendiren araştırmalar, halk sağlığı politikaları alanının kuramsal ve uygulamalı gelişimine önemli katkılar sağlayabilir.

KAYNAKLAR

Azhari, S. W., Violeta, D. T., Sitorus, V., & Tanziha, I. (2025). Health policy towards obesity has an impact on non-communicable diseases (ncds) over the last 10 years: a bibliometric analysis. *Media Gizi Indonesia*, *20*(3), 255–279. https://doi.org/10.20473/mgi.v20i3.255-279

Brownson, R. C., Chriqui, J. F., & Stamatakis, K. A. (2009a). Understanding Evidence-Based Public Health Policy. *American Journal of Public Health*, *99*(9), 1576–1583. https://doi.org/10.2105/AJPH.2008.156224

Brownson, R. C., Chriqui, J. F., & Stamatakis, K. A. (2009b). Understanding Evidence-Based Public Health Policy. *American Journal of Public Health*, *99*(9), 1576–1583. https://doi.org/10.2105/AJPH.2008.156224

Burton, D. C., Burris, S., Mermin, J. H., Purcell, D. W., Zeigler, S. C., Bull-Otterson, L., & Dean, H. D. (2020). Policy and Public Health. *Public Health Reports®*, *135*(1), 5–9. https://doi.org/10.1177/0033354920932641

de Leeuw, E. (2020). Public policy for health promotion: a special curated collection. *Health Promotion International*, *35*(2), 381–385. https://doi.org/10.1093/heapro/daz118

Donkin, A., Goldblatt, P., Allen, J., Nathanson, V., & Marmot, M. (2018). Global action on the social determinants of health. *BMJ Global Health*, *3*(Suppl 1), e000603. https://doi.org/10.1136/bmjgh-2017-000603

Dugle, G., Wulifan, J. K., Tanyeh, J. P., & Quentin, W. (2020). A critical realist synthesis of cross-disciplinary health policy and systems research: defining characteristic features, developing an evaluation framework and identifying challenges. *Health Research Policy and Systems*, *18*(1), 79. https://doi.org/10.1186/s12961-020-00556-2

Fundytus, K., Santamaria-Plaza, C., & McLaren, L. (2023). Policy diffusion theory, evidence-informed public health, and public health political science: a scoping review. *Canadian Journal of Public Health*, *114*(3), 331–345. https://doi.org/10.17269/s41997-023-00752-x

Green, L., Ashton, K., Bellis, M. A., Clemens, T., & Douglas, M. (2021). 'Health in All Policies' A Key Driver for Health and Well-Being in a Post-COVID-19 Pandemic World. *International Journal of Environmental Research and Public Health*, *18*(18), 9468. https://doi.org/10.3390/ijerph18189468

Gupta, G. R., Oomman, N., Grown, C., Conn, K., Hawkes, S., Shawar, Y. R., Shiffman, J., Buse, K., Mehra, R., Bah, C. A., Heise, L., Greene, M. E., Weber, A. M., Heymann, J., Hay, K., Raj, A., Henry, S., Klugman, J., & Darmstadt, G. L. (2019). Gender equality and gender norms: framing the opportunities for health. *The Lancet*, *393*(10190), 2550–2562. https://doi.org/10.1016/S0140-6736(19)30651-8

Harris, P., & Wise, M. (2020). Healthy Public Policy. In *Oxford Bibliographies in Public Health*. Oxford University PressNew York, NY. https://doi.org/10.1093/obo/9780199756797-0196

Jervis-Jalabe, D. S., Olivares-Cerpa, F. D. J., Jervis-Pinzon, C. A., & Arteta-Olivares, N. X. (2023). Decision support techniques in public health policy formulation: bibliometric analysis. In *Research Square*. https://doi.org/10.21203/rs.3.rs-2399509/v1

Judijanto, L., Sabaruddin, E. E., Zani, B. N., Musiana, M., & Ibrahim, S. M. (2024). Trends in Public Health Research: a Bibliometric Review 2018-2022. *Journal of World Future Medicine, Health and Nursing*, *2*(1), 65–78. https://doi.org/10.70177/health.v2i1.711

Kelly, M. A., Puddy, R. W., Siddiqi, S. M., Nelson, C., Ntazinda, A. H., Kucik, J. E., Hall, D., Murray, C. T., & Tomoaia-Cotisel, A. (2025). Distilling the Fundamentals of Evidence-Based Public Health Policy. *Public Health Reports®*, *140*(1), 40–47. https://doi.org/10.1177/00333549241256751

Kivits, J., Ricci, L., & Minary, L. (2019). Interdisciplinary research in public health: the 'why' and the 'how.' *Journal of Epidemiology and Community Health*, *73*(12), 1061–1062. https://doi.org/10.1136/jech-2019-212511

Larivière, V., Haustein, S., & Mongeon, P. (2015). The Oligopoly of Academic Publishers in the Digital Era. *PLOS ONE*, *10*(6), e0127502. https://doi.org/10.1371/journal.pone.0127502

Petticrew, M., Whitehead, M., Macintyre, S. J., Graham, H., & Egan, M. (2004). Evidence for public health policy on inequalities: 1: The reality according to policymakers. *Journal of Epidemiology and Community Health*, *58*(10), 811–816. https://doi.org/10.1136/jech.2003.015289

Pollack Porter, K. M., Rutkow, L., & McGinty, E. E. (2018). The Importance of Policy Change for Addressing Public Health Problems. *Public Health Reports®*, *133*(1_suppl), 9S-14S. https://doi.org/10.1177/0033354918788880

Tausch, A. (2018). The Market Power of Global Scientific Publishing Companies in the Age of Globalization. An Analysis Based on the OCLC Worldcat. *Journal of Globalization Studies*, *9*(2), 63–91. https://doi.org/10.30884/jogs/2018.02.05

van de Goor, I., Hämäläinen, R.-M., Syed, A., Juel Lau, C., Sandu, P., Spitters, H., Eklund Karlsson, L., Dulf, D., Valente, A., Castellani, T., & Aro, A. R. (2017). Determinants of evidence use in public health policy making: Results from a study across six EU countries. *Health Policy*, *121*(3), 273–281. https://doi.org/10.1016/j.healthpol.2017.01.003

Web of Science. (2026). *Web of Science Core Collection*. Https://Www.Webofscience.Com/Wos/Woscc/Summary/7946bd30-02e0-4e91-8390-6f30a5b338fa-01c2b98ceb/Relevance/1. (Erişim Tarihi: 10.08.2026).

Whitehead, M., Petticrew, M., Graham, H., Macintyre, S. J., Bambra, C., & Egan, M. (2004). Evidence for public health policy on inequalities: 2: Assembling the evidence jigsaw. *Journal of Epidemiology and Community Health*, *58*(10), 817–821. https://doi.org/10.1136/jech.2003.015297

Zhao, L., Zhao, Y., Du, J., Desloge, A., Hu, Z., & Cao, G. (2022). Mapping the Research on Health Policy and Services in the Last Decade (2009–2018): A Bibliometric Analysis. *Frontiers in Public Health*, *10*. https://doi.org/10.3389/fpubh.2022.773668

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Machine and Deep Learning in Mechanics of Materials: The Evaluation Protocol Governs Reported Accuracy in Alloy Yield Strength Prediction

Hamdi Kuleyin[1] and Erkan Caner Ozkat[2*]

[1]*Department of Mechanical Engineering, Faculty of Engineering and Architecture, Recep Tayyip Erdogan University, Rize, Turkiye (ORCID: https://orcid.org/0000-0001-7507-5531)*

[2]*Department of Mechanical Engineering, Faculty of Engineering and Architecture, Recep Tayyip Erdogan University, Rize, Turkiye (ORCID: https://orcid.org/0000-0003-0530-5439)*

[*]*(erkancaner.ozkat@erdogan.edu.tr) Email of the corresponding author*

Abstract – Machine learning is now widely applied to structure and property problems in the mechanics of materials. The accuracy that a study reports, however, depends on how the test set is formed. This work quantifies that dependence for the yield strength of multi-principal element alloys. An open access dataset of 1545 experimental records was used, of which 1067 report both yield strength and test temperature. Eleven composition descriptors were derived from solid solution theory, including atomic size mismatch, mixing entropy and valence electron concentration. Six statistical and machine learning regressors and one deep neural network were trained under three cross-validation protocols. The first protocol shuffles records at random, which is the common practice. The second and the third protocols keep the folds disjoint in alloy composition and in source publication, respectively. A variance component analysis showed that alloy identity explains 64.4 per cent of the yield strength variance, while publication identity explains 61.0 per cent. Random splitting therefore places replicate records of one alloy in the training and the test fold at the same time. Gradient boosting reached a coefficient of determination of 0.796 under random splitting, 0.701 under composition-disjoint splitting and 0.581 under source-disjoint splitting. The deep network followed the same pattern, and the ranking of the models also changed with the protocol. Headline accuracies obtained by random splitting therefore overstate how well a model generalises to an unseen alloy. Grouped validation is recommended whenever a data driven model is intended to support alloy design.

Keywords – High-entropy alloys, yield strength, machine learning, deep learning, cross-validation

I. INTRODUCTION

Data driven methods have become a standard tool in continuum materials mechanics, where they complement constitutive modelling and experiment [1]. Regression models now predict stiffness, strength and ductility directly from composition and processing variables [2]. The appeal is practical, because a surrogate model screens candidate materials far faster than a mechanical test campaign.

Multi-principal element alloys illustrate both the promise and the difficulty of this approach. These alloys contain several elements in near equal proportion, rather than one base element [3], [4]. They can form single phase solid solutions with attractive strength and toughness [5]. Their composition space is enormous, so exhaustive testing is impossible. Empirical descriptors such as the atomic size mismatch, the ideal mixing entropy and the valence electron concentration were introduced to rationalise phase selection [6], [7]. These descriptors are computed from tabulated elemental radii and electronegativities [8]. The same descriptors are now used as machine learning features for property prediction [9].

A question that receives less attention is how such models are validated. Mechanical property datasets are usually compilations assembled from the published literature. One alloy is therefore measured many times, at several test temperatures, and often within a single study. When records are shuffled at random,

replicate measurements of one alloy enter the training fold and the test fold together. The test score then rewards recognition of a familiar alloy rather than genuine extrapolation. This mechanism is a form of data leakage [10]. It has been reported for superconductor discovery [11], for materials discovery campaigns in general [12], and for redundant crystal structure datasets [13]. Its magnitude has not been quantified for the mechanical properties of multi-principal element alloys.

This paper addresses that gap. An open access dataset of experimentally reported alloy properties [14] is combined with descriptors motivated by solid solution mechanics. Six statistical and machine learning regressors and one deep neural network are then evaluated under three cross-validation protocols of increasing strictness. The contributions are threefold. First, the replicate structure of the dataset is quantified by a variance component analysis. Second, the accuracy of every model is reported under random, composition-disjoint and source-disjoint splitting. Third, the effect of the protocol on model ranking is examined. The remainder of the paper is organised as follows. Section II describes the dataset, the descriptors, the models and the protocols. Section III reports the statistical structure and the accuracy results. Section IV discusses the implications for alloy design, and Section V states the conclusions.

II. Materials and Method

This section is organised in five parts. Part A describes the dataset. Part B defines the mechanics based descriptors. Part C lists the regression models. Part D defines the three cross-validation protocols. Part E describes the statistical analysis.

A. Dataset

The study uses the open access compilation of mechanical properties of multi-principal element alloys published by Borg and co-workers [14]. The compilation contains 1545 records extracted from the published literature. Each record reports an alloy composition, a processing route, an observed phase, a test condition and one or more mechanical properties. The present work models the room and elevated temperature yield strength. Records were retained when both the yield strength and the test temperature were reported, which gives 1067 records. These records cover 395 distinct compositions drawn from 191 source publications, and involve 26 chemical elements. Yield strength ranges from 24 to 3416 MPa, and the test temperature ranges from -269 to 1600 degrees Celsius.

B. Mechanics based descriptors

Each formula was parsed into atomic fractions and converted into eleven descriptors that are standard in solid solution theory [6]-[8]. The atomic size mismatch is defined in (1). Here c_i is the atomic fraction of element i, r_i is its atomic radius, and $\bar{r}$ is the composition average radius.

$$\delta = 100 \times [\, \Sigma\, c_i \,(1 - r_i / \bar{r}\,)^2 \,]^{1/2} \qquad (1)$$

The ideal configurational mixing entropy is given in (2), where R is the universal gas constant.

$$\Delta S_{mix} = - R \,\Sigma\, c_i \ln c_i \qquad (2)$$

The remaining descriptors are the number of elements and the mean atomic radius. They also include the rule of mixtures melting point and the valence electron concentration. The mean and the mismatch of Pauling electronegativity complete the chemical set. The mean molar mass and the mean and the mismatch of the elemental shear modulus were added as stiffness proxies. Two test variables follow, namely the test temperature and the homologous temperature. The homologous temperature is the absolute test temperature divided by the rule of mixtures melting point, so it carries the thermal activation argument of deformation mechanics. The reported density, the processing route, the observed phase and the test mode complete the feature set. Table 1 summarises the modelled variables.

C. Regression models

Seven predictors were compared. A mean predictor provides the null reference. Ordinary least squares and ridge regression represent the classical statistical baseline. Support vector regression with a radial basis kernel, a random forest of 500 trees [15] and gradient boosted trees with 600 estimators [16] represent classical machine learning. The deep model is a fully connected network with three hidden layers of 128, 64 and 32 units. It uses batch normalisation, rectified linear activation, dropout of 0.1 and

the Adam optimiser [17]. Training stops early when an inner validation loss stops improving for 40 epochs, and that inner split respects the same grouping as the outer protocol. Hyperparameters were held fixed across all protocols, so that any change in accuracy is caused by the split and not by tuning. The classical models were implemented with scikit-learn [18] and the deep model with PyTorch [19].

D. Cross-validation protocols

Three ten-fold protocols were applied to the same data and the same models. The random protocol shuffles the records, which is the usual practice. The composition-disjoint protocol assigns every record of one alloy composition to a single fold. The source-disjoint protocol assigns every record of one source publication to a single fold. The second protocol therefore tests generalisation to an unseen alloy. The third additionally removes the shared laboratory, the shared processing route and the shared measurement practice of a single study. Each protocol was repeated over five seeds, and the reported values are the mean and the standard deviation across those seeds. All preprocessing was fitted inside the training fold only.

E. Statistical analysis

Descriptive statistics summarise the modelled variables. Spearman rank correlation quantifies the monotone association between each descriptor and the yield strength. The Kruskal-Wallis test compares the yield strength across the reported phase, the processing route and the test mode. A one-way random effects analysis of variance gives the intraclass correlation coefficient for alloy identity and for source identity. That coefficient measures the share of the yield strength variance that lies between groups rather than within them, and it is the statistic that motivates grouped validation.

Table 1. Descriptive statistics of the modelled variables (n = 1067 records).

Variable	Mean	SD	Minimum	Maximum
Yield strength, MPa	890	570	24	3416
Test temperature, C	258	408	-269	1600
Homologous temperature	0.25	0.18	0.00	0.90
Atomic size mismatch, %	4.79	2.15	0.14	19.40
Mixing entropy, J/(mol K)	12.73	1.61	6.44	18.21
Valence electron concentration	6.16	1.69	1.77	9.60
Mean melting point, K	2097	424	702	3261
Mean shear modulus, GPa	66.4	18.1	13.5	116.6
Density, g/cm3	8.08	1.78	1.40	13.70
Number of elements	4.97	0.87	3.00	9.00

III. Results

This section reports the results in three parts. Part A characterises the replicate structure of the dataset. Part B reports the accuracy of every model under the three protocols. Part C examines the effect of the protocol on model ranking.

A. Replicate structure and descriptive statistics

The 1067 modelled records describe only 395 distinct compositions. A composition that appears more than once accounts for 76.3 per cent of the records, and the largest single composition contributes 40 records. The corresponding figures for source publications are 96.9 per cent and 78 records. Repetition is not simple duplication, because the median spread of test temperature within a repeated composition is 775 degrees Celsius. The one-way analysis of variance gives an intraclass correlation of 0.437 for alloy identity and 0.529 for source identity, both with p below 0.001. Group identity alone explains 64.4 per cent of the yield strength variance for compositions and 61.0 per cent for sources. Fig. 1 shows the resulting replication structure and the strength distribution.

The descriptive statistics of the modelled variables are given in Table 1. Two mechanics effects are visible in the raw data. The test mode separates the population, because the median yield strength is 1010 MPa in compression and 521 MPa in tension. A Mann-Whitney test confirms the difference, with p below

0.001 and a Cliff delta of 0.41. The reported phase also matters, since body centred cubic alloys are stronger and less ductile than face centred cubic alloys. Fig. 1(b) displays both effects.

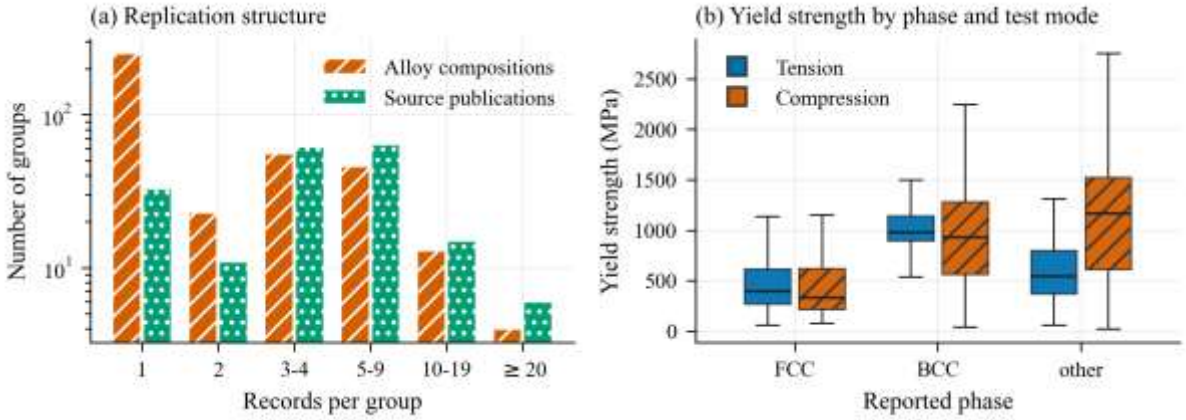


Fig. 1 Structure of the modelled dataset. (a) Number of records per alloy composition and per source publication. (b) Yield strength by reported phase and test mode.

B. Accuracy under the three protocols

Table 2 reports the coefficient of determination and the root mean squared error for every model and protocol. Under random splitting the best model is gradient boosting, which reaches an R squared of 0.796. That value is the kind of headline number usually reported for this class of problem. The same model and the same hyperparameters give 0.701 when the folds are disjoint in composition and 0.581 when they are disjoint in source. The loss of explained variance is therefore 0.095 and 0.215, respectively. The root mean squared error rises from 257 MPa to 369 MPa over the same comparison. The deep network behaves in the same way, falling from 0.793 to 0.613, a loss of 0.179. Fig. 2 shows the effect for every model, and Fig. 3 shows the corresponding parity plots.

Table 2. Out-of-fold accuracy by model and cross-validation protocol. Values are the mean over five seeds, with the standard deviation in brackets. RMSE is in MPa.

Model	Random		Composition-disjoint		Source-disjoint	
	R^2	RMSE	R^2	RMSE	R^2	RMSE
Mean baseline	-0.002 (0.001)	571	-0.007 (0.001)	572	-0.007 (0.004)	572
Linear regression	0.589 (0.002)	365	0.565 (0.007)	376	0.525 (0.015)	393
Ridge regression	0.589 (0.002)	366	0.569 (0.004)	374	0.536 (0.012)	388
Support vector regression	0.685 (0.003)	320	0.644 (0.002)	340	0.616 (0.006)	353
Random forest	0.786 (0.008)	264	0.681 (0.009)	322	0.570 (0.016)	373
Gradient boosting	0.796 (0.008)	257	0.701 (0.011)	311	0.581 (0.019)	369
Deep neural network	0.793 (0.007)	260	0.671 (0.022)	327	0.613 (0.028)	354

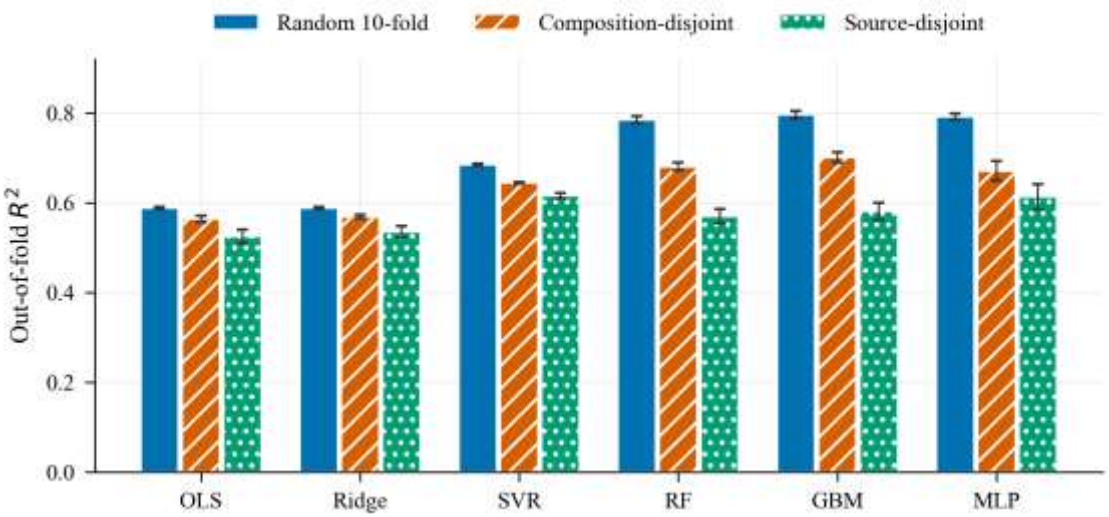


Fig. 2 Out-of-fold coefficient of determination for each model under the three cross-validation protocols. Bars are the mean over five seeds and error bars are one standard deviation.

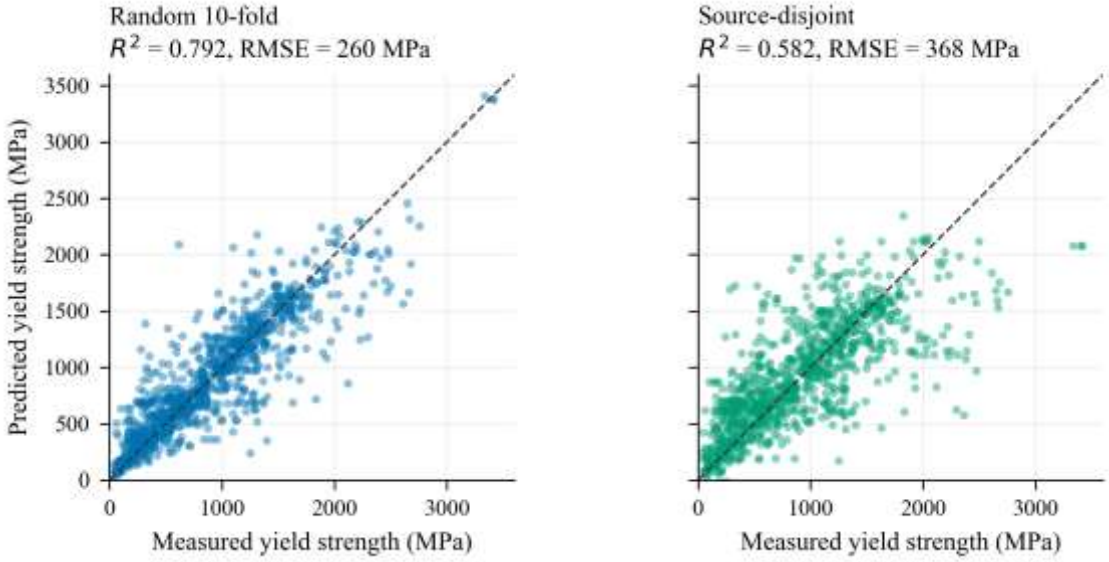


Fig. 3 Out-of-fold parity plots for gradient boosting under the loosest and the strictest protocol. The dashed line is the ideal prediction.

C. Effect of the protocol on model ranking

The protocol also changes which model appears best. Under random splitting the ranking is led by gradient boosting, followed by deep neural network. Under source-disjoint splitting the leader is support vector regression. The gap between the flexible learners and the linear baseline narrows sharply. The flexible learners lose the most when replicate records leave the training fold. Linear regression reaches 0.525 under the strictest protocol, against 0.589 under random splitting. A comparison made under random splitting would therefore recommend a different model from the one that generalises best to an unseen alloy.

In summary, the results establish three findings. Alloy identity and source identity explain most of the yield strength variance. Random splitting inflates the reported accuracy of every learned model, and the inflation is largest for the most flexible models. The choice of protocol also alters the ranking of the models.

IV. Discussion

The results show that the evaluation protocol, and not only the learning algorithm, determines the accuracy that a study reports. The mechanism is specific to the way mechanical property datasets are

built. A literature compilation records one alloy many times, at several test temperatures and often inside one study. Random splitting then asks the model to interpolate between two measurements of a material it has already seen. That task is much easier than predicting the strength of a new alloy, which is the task that alloy design actually requires.

The variance component analysis explains why the effect is so large. Group identity accounts for the majority of the yield strength variance, so knowing the alloy is close to knowing the answer. This observation agrees with reports from other branches of materials informatics, where clustered validation and redundancy control produced similar reductions [11], [13]. The present study adds a quantitative estimate for the mechanical properties of multi-principal element alloys, and it separates two distinct grouping levels. The source-disjoint protocol is the stricter of the two, which indicates that a shared laboratory and a shared processing route leak information beyond composition alone.

The practical implication is not that machine learning fails in mechanics of materials. The models still explain a useful share of the variance under the strict protocols, and they remain far better than the mean predictor. The implication is that a headline score obtained by random splitting is the wrong quantity to report when the intended use is design. A grouped score answers the question that a designer asks, which is how the model will behave on an alloy that has never been tested. Reporting both values is inexpensive and makes the claim verifiable.

Three limitations should be stated. The dataset mixes tension and compression tests, and the two populations differ significantly, so the test mode was included as a feature rather than modelled separately. Hyperparameters were fixed rather than tuned inside each protocol, which keeps the comparison controlled but may understate the achievable accuracy. Finally, the descriptors are composition averages, so they do not capture grain size, texture or precipitate state, all of which govern strength in real microstructures.

V. Conclusion

This study combined an open access dataset of multi-principal element alloys with descriptors from solid solution mechanics, and evaluated seven regression models under three cross-validation protocols. The main conclusions are the following.

1) Alloy identity explains 64.4 per cent and source identity explains 61.0 per cent of the yield strength variance, so the records are strongly replicated.

2) Gradient boosting falls from an R squared of 0.796 under random splitting to 0.701 under composition-disjoint splitting and 0.581 under source-disjoint splitting.

3) The deep neural network shows the same loss, so the effect is a property of the evaluation protocol and not of one model family.

4) The ranking of the models changes with the protocol, so a model chosen under random splitting need not be the model that generalises best.

5) Grouped validation should be reported whenever a data driven model in mechanics of materials is intended to guide the design of a new material.

Acknowledgment

The authors thank the compilers of the open access dataset of multi-principal element alloys for making the records publicly available. The analysis code and the derived results files of this study are available from the corresponding author on reasonable request.

References

[1] F. E. Bock, R. C. Aydin, C. J. Cyron, N. Huber, S. R. Kalidindi, and B. Klusemann, “A review of the application of machine learning and data mining approaches in continuum materials mechanics,” *Front. Mater.*, vol. 6, art. 110, 2019.

[2] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, “Machine learning for molecular and materials science,” *Nature*, vol. 559, pp. 547–555, 2018.

[3] J. W. Yeh *et al.*, “Nanostructured high-entropy alloys with multiple principal elements: novel alloy design concepts and outcomes,” *Adv. Eng. Mater.*, vol. 6, no. 5, pp. 299–303, 2004.

[4] B. Cantor, I. T. H. Chang, P. Knight, and A. J. B. Vincent, “Microstructural development in equiatomic multicomponent alloys,” *Mater. Sci. Eng. A*, vol. 375–377, pp. 213–218, 2004.

[5] E. P. George, D. Raabe, and R. O. Ritchie, “High-entropy alloys,” *Nat. Rev. Mater.*, vol. 4, pp. 515–534, 2019.

[6] Y. Zhang, Y. J. Zhou, J. P. Lin, G. L. Chen, and P. K. Liaw, “Solid-solution phase formation rules for multi-component alloys,” *Adv. Eng. Mater.*, vol. 10, no. 6, pp. 534–538, 2008.
[7] S. Guo and C. T. Liu, “Phase stability in high entropy alloys: formation of solid-solution phase or amorphous phase,” *Prog. Nat. Sci. Mater. Int.*, vol. 21, no. 6, pp. 433–446, 2011.
[8] A. Takeuchi and A. Inoue, “Classification of bulk metallic glasses by atomic size difference, heat of mixing and period of constituent elements,” *Mater. Trans.*, vol. 46, no. 12, pp. 2817–2829, 2005.
[9] C. Wen *et al.*, “Machine learning assisted design of high entropy alloys with desired property,” *Acta Mater.*, vol. 170, pp. 109–117, 2019.
[10] S. Kaufman, S. Rosset, C. Perlich, and O. Stitelman, “Leakage in data mining: formulation, detection, and avoidance,” *ACM Trans. Knowl. Discov. Data*, vol. 6, no. 4, art. 15, 2012.
[11] B. Meredig *et al.*, “Can machine learning identify the next high-temperature superconductor? Examining extrapolation performance for materials discovery,” *Mol. Syst. Des. Eng.*, vol. 3, no. 5, pp. 819–825, 2018.
[12] C. K. H. Borg *et al.*, “Quantifying the performance of machine learning models in materials discovery,” *Digital Discovery*, vol. 2, no. 2, pp. 327–338, 2023.
[13] Q. Li, N. Fu, S. S. Omee, and J. Hu, “MD-HIT: machine learning for material property prediction with dataset redundancy control,” *npj Comput. Mater.*, vol. 10, art. 245, 2024.
[14] C. K. H. Borg *et al.*, “Expanded dataset of mechanical properties and observed phases of multi-principal element alloys,” *Sci. Data*, vol. 7, art. 430, 2020.
[15] L. Breiman, “Random forests,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[16] T. Chen and C. Guestrin, “XGBoost: a scalable tree boosting system,” in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining*, San Francisco, CA, USA, 2016, pp. 785–794.
[17] D. P. Kingma and J. Ba, “Adam: a method for stochastic optimization,” in *Proc. 3rd Int. Conf. Learning Representations*, San Diego, CA, USA, 2015.
[18] F. Pedregosa *et al.*, “Scikit-learn: machine learning in Python,” *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.
[19] A. Paszke *et al.*, “PyTorch: an imperative style, high-performance deep learning library,” in *Advances in Neural Information Processing Systems 32*, 2019, pp. 8024–8035.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Comparative Performance Analysis of PID and Active Disturbance Rejection Control for a Nonlinear Magnetic Levitation System

Mehmet Latif LEVENT[1*], Halit HÜLAKO[2] and Sertaç YAMAN [1]

[1]*Department of Electrical-Electronics/ Faculty of Engineering, Hakkari University, Turkey*
[2]*Department of Mechanical Engineering /Faculty of Engineering, Hakkari University, Turkey*

[*]*(mehmetlatiflevent@hakkari.edu.tr)*

Abstract **–** Magnetic levitation (MagLev) systems represent a challenging control problem due to their inherently nonlinear dynamics, open-loop instability, and high sensitivity to external disturbances. In this study, a comparative performance analysis between a conventional Proportional-Integral-Derivative (PID) controller and an Active Disturbance Rejection Controller (ADRC) is carried out for the position control of a nonlinear magnetic levitation system. The complete nonlinear mathematical model of the MagLev system is developed and implemented in MATLAB/Simulink, and both control strategies are tested under identical simulation conditions to ensure a fair comparison.

Two distinct operating scenarios are considered. In the first scenario, the system is evaluated under no-load conditions in order to examine the pure reference-tracking performance of both controllers. In the second scenario, an external load disturbance is applied at $(t = 7\text{s})$ to assess the robustness and disturbance-rejection capability of the controllers. The simulation results indicate that both PID and ADRC are capable of stabilizing the unstable MagLev system and successfully tracking the desired position reference. However, significant differences are observed in transient behavior. The ADRC controller demonstrates a noticeably lower overshoot, improved damping characteristics, and faster convergence to the reference value compared to the conventional PID controller. In addition, when subjected to the load disturbance, ADRC responds more rapidly, effectively compensates for the disturbance, and achieves a shorter recovery time. Overall, the obtained results clearly show that ADRC provides superior transient response and enhanced disturbance-rejection performance for nonlinear magnetic levitation systems when compared to classical PID control.

Keywords – Active disturbance rejection control, ADRC, PID control, magnetic levitation, nonlinear control, disturbance rejection, MATLAB/Simulink.

I. INTRODUCTION

Magnetic levitation systems are based on maintaining a ferromagnetic object at a desired position by balancing gravitational and electromagnetic forces. Since the electromagnetic force is a nonlinear function of both the coil current and the distance between the electromagnet and the levitated object, magnetic levitation systems exhibit highly nonlinear dynamic characteristics. Furthermore, the equilibrium point of the system is inherently unstable in open-loop operation. Consequently, an effective closed-loop controller is required to maintain the levitated object at the desired position [1], [2].

The Proportional-Integral-Derivative (PID) controller is one of the most widely used feedback control methods because of its simple structure, ease of implementation, and satisfactory performance in many industrial applications. The PID control signal is generated from the proportional, integral, and derivative components of the tracking error. Nevertheless, the performance of a PID controller may deteriorate in

the presence of strong nonlinearities, model uncertainties, parameter variations, and external disturbances [3].

Active Disturbance Rejection Control (ADRC) provides an alternative control strategy by treating internal uncertainties and external disturbances as a generalized disturbance. An Extended State Observer (ESO) is employed to estimate this total disturbance in real time, and the estimated disturbance is compensated through the control input. Therefore, ADRC can provide effective control performance without requiring a highly accurate mathematical model of the plant [4].

In this study, conventional PID and classical ADRC controllers are comparatively investigated for the position control of a nonlinear magnetic levitation system. Unlike optimization-based or intelligent control approaches, the proposed study focuses exclusively on the comparison of the two classical control structures under identical simulation conditions.

The nonlinear MagLev system is implemented in MATLAB/Simulink. The complete Simulink implementation including the PID and ADRC controller branches is presented in Fig. 1. Two operating scenarios are investigated. First, the controllers are compared under nominal no-load conditions. Second, an external load disturbance is applied at $(t = 7\text{s})$, and the disturbance-rejection performances of the controllers are evaluated.

The major contributions of this study can be summarized as follows:

- A nonlinear MagLev system is implemented directly from its nonlinear dynamic equations in MATLAB/Simulink.
- PID and ADRC controllers are implemented on the same nonlinear plant under identical operating conditions.
- Reference-tracking performances of the controllers are compared under no-load conditions.
- Robustness and disturbance-rejection performances are investigated by applying an external load disturbance.
- The transient responses of the controllers are quantitatively evaluated using rise time, settling time, maximum overshoot, steady-state error, and error-based performance indices.

II. Mathematical Model of the Magnetic Levitation System

The magnetic levitation system considered in this study consists of an electromagnetic actuator and a ferromagnetic levitated object. The vertical motion of the object is determined by the gravitational force, electromagnetic force, and external disturbance force.

According to Newton's second law, the vertical motion of the levitated object can be expressed as [5]

$$[m\ddot{h} = mg - F_m + F_d] \tag{1}$$

where (m) represents the mass of the levitated object, (h) is the vertical position, (g) is gravitational acceleration, (F_m) represents the electromagnetic attraction force, and (F_d) denotes the external disturbance force.

The electromagnetic force is modeled as

$$F_m = \frac{\alpha i^2}{h^2} \tag{2}$$

where (i) is the electromagnetic coil current and (α) represents the magnetic force coefficient. Substituting the electromagnetic force into the mechanical equation gives

$$mg - \frac{\alpha i^2}{h^2} + F_d \tag{3}$$

and consequently,

$$g - \frac{\alpha}{m}\frac{i^2}{h^2} + \frac{F_d}{m}. \tag{4}$$

This equation clearly indicates the nonlinear nature of the magnetic levitation system because the electromagnetic force depends on both (i^2) and ($1/h^2$).

The nominal parameters used in the simulations are given in Table 1.

Table 1. Parameters of the nonlinear magnetic levitation system

Parameter	Symbol	Value	Unit
Mass of levitated object	m	0.02	kg
Gravitational acceleration	g	9.81	m/s^2
Magnetic coefficient	α	2.4832×10^{-5}	—
Initial position	h_0	0.15	m
Initial velocity	$\dot{h}_0$	0	m/s
Operating-point current	i_0	13.3332	A

At the equilibrium point, $\ddot{h} = 0$ and $F_d = 0$. Therefore,

$$mg = \frac{\alpha i_0^2}{h_0^2}. \quad [5]$$

The equilibrium current can be determined as

$$h_0 \sqrt{\frac{mg}{\alpha}}. \quad [6]$$

Using the parameters given in Table 1, the operating-point current is obtained as approximately $i_0 = 13.3332\text{A}$. The controller output is defined as the variation of the coil current around the operating point, $u = \Delta i$. Thus, the total current applied to the nonlinear MagLev model is $i = i_0 + \Delta i$. This formulation enables both PID and ADRC controllers to operate around the same equilibrium point and provides identical plant-input conditions for a fair comparison.

III. MATLAB/SIMULINK IMPLEMENTATION

The complete nonlinear magnetic levitation system was implemented in MATLAB/Simulink using the differential equation presented in the previous section. The implemented model, including the PID and ADRC controller branches, is shown in Figure 1.

In the Simulink model, the total coil current is first obtained as $i = i_0 + \Delta i$. The square of the current, i^2, and the square of the measured position, h^2 are calculated separately. The nonlinear electromagnetic term is then obtained using

$$-\frac{\alpha}{m}\frac{i^2}{h^2}. \quad [7]$$

The acceleration of the levitated object is calculated as

$$\frac{\alpha}{m}\frac{i^2}{h^2} + \frac{F_d}{m}. \quad [8]$$

Two consecutive integrators are used to obtain the velocity and position:

$$\dot{h}(t) = \int \ddot{h}(t)dt \quad [9]$$

and

$$h(t) = \int \dot{h}(t)dt. \qquad [10]$$

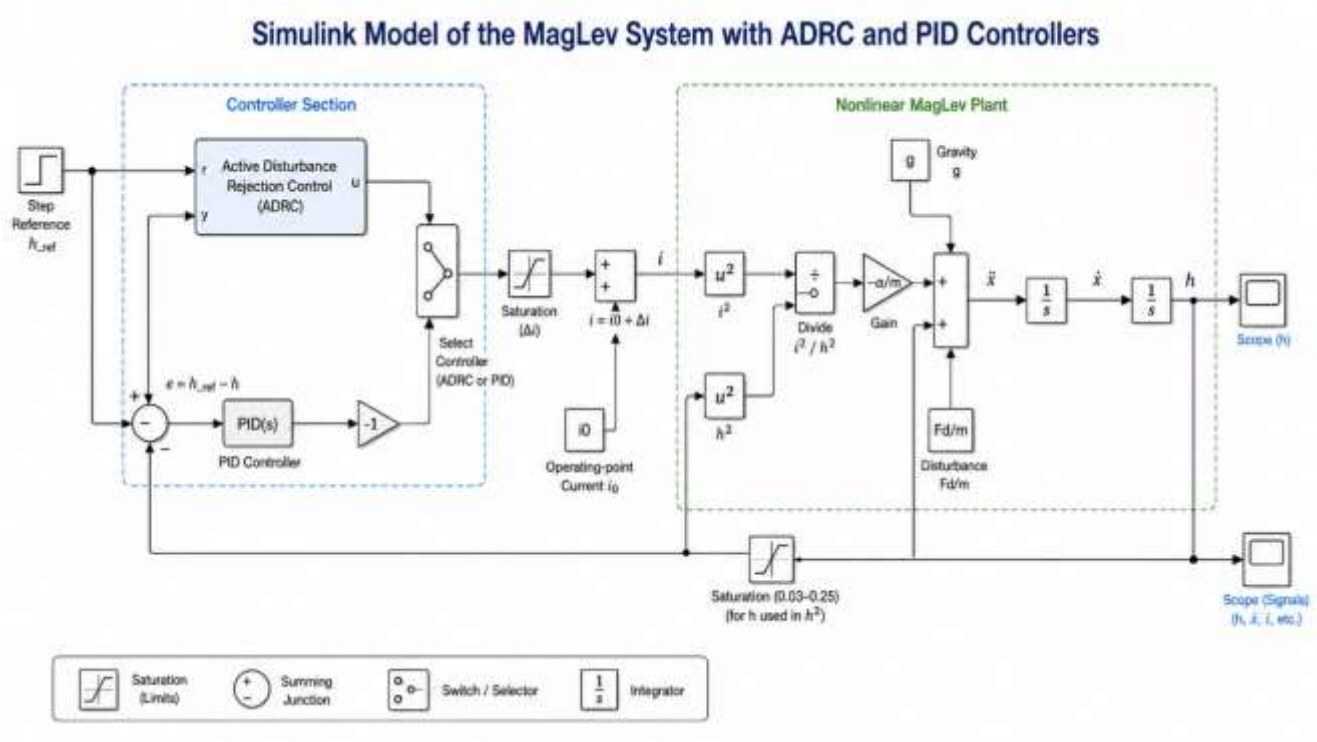


Figure 1. MATLAB/Simulink implementation of the nonlinear MagLev system with PID and ADRC controllers.

Since the nonlinear model contains the term $(1/h^2)$, the position signal used in the denominator was limited to prevent numerical singularities. The position used for the nonlinear electromagnetic force calculation was restricted within the interval $0.03 \le h \le 0.25\text{m}$. The actual position output was fed back independently to both the PID and ADRC structures.

IV. Simulation Results and Discussion

A. NO-LOAD PERFORMANCE

The reference-tracking responses obtained with the PID and ADRC controllers under no-load conditions are presented in Figure 2.

As shown in Figure 2, both controllers successfully stabilize the nonlinear magnetic levitation system and track the desired reference position of 0.17m. However, significant differences can be observed in their transient responses. The PID-controlled system exhibits a relatively large initial overshoot following the reference change at $(t = 1\text{s})$. After reaching its first maximum value, the response crosses the reference and exhibits a noticeable underdamped oscillation before converging to the desired position. In contrast, the ADRC-controlled response exhibits considerably smaller overshoot. Although a limited oscillatory transient is observed immediately following the reference change, the oscillations are rapidly attenuated and the system reaches the desired position significantly faster than the PID-controlled system. Another important observation is that both controllers provide approximately zero steady-state tracking error. Therefore, the primary difference between the two controllers is observed in their transient performance rather than their steady-state accuracy. The quantitative no-load results should be summarized in Table 2.

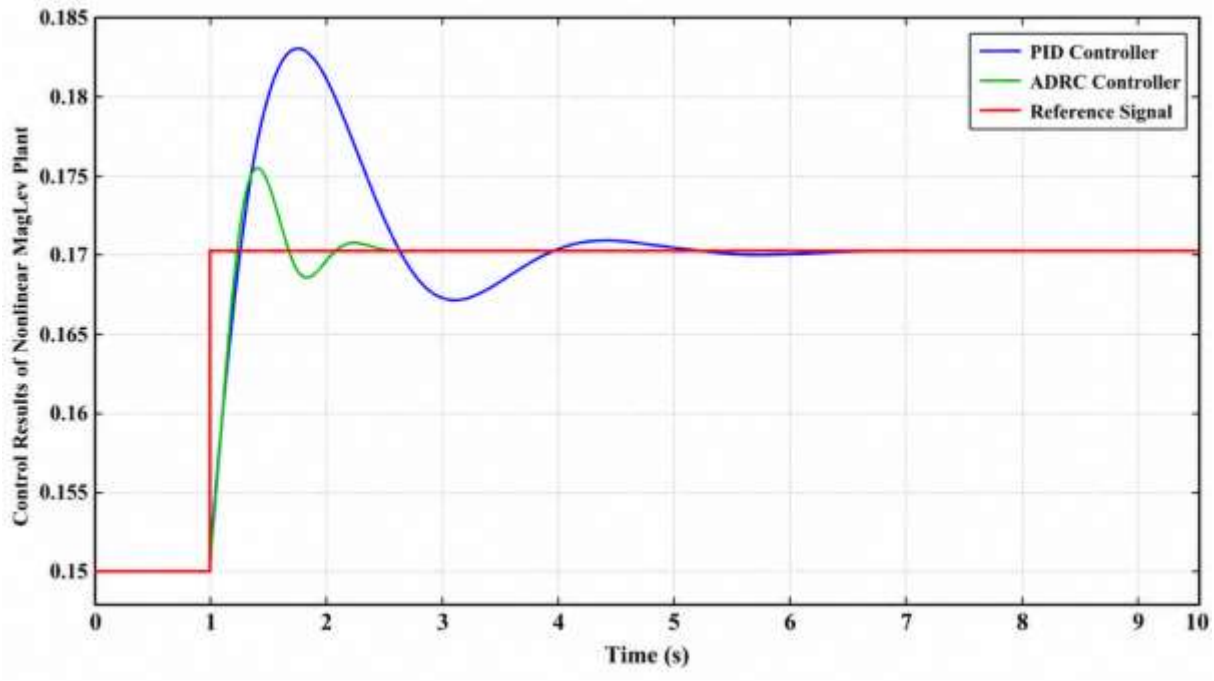


Figure 2. Reference-tracking responses of the PID and ADRC controllers under no-load conditions.

Table 2. Performance comparison under no-load conditions

Performance criterion	PID	ADRC	Improvement (%)
Rise time (T_r)	0.2254	0.1793	20.44
Settling time (T_s)	3.7036	1.2618	65.93
Maximum overshoot (M_p) (%)	64.06	26.58	58.50
Steady-state error (e_{ss})	3.39×10^{-6}	1.08×10^{-12}	100.00
IAE	0.0164	0.0048	71.02
ISE	1.48×10^{-4}	4.86×10^{-5}	67.25
ITAE	0.0174	0.0014	91.97

The response shown in Figure 2 qualitatively demonstrates that ADRC provides improved transient response, particularly in terms of maximum overshoot and settling behavior.

B. PERFORMANCE UNDER LOAD DISTURBANCE

The robustness of the controllers was evaluated by applying an external load disturbance at $t = 7$s. The responses obtained under this condition are shown in Figure 3.

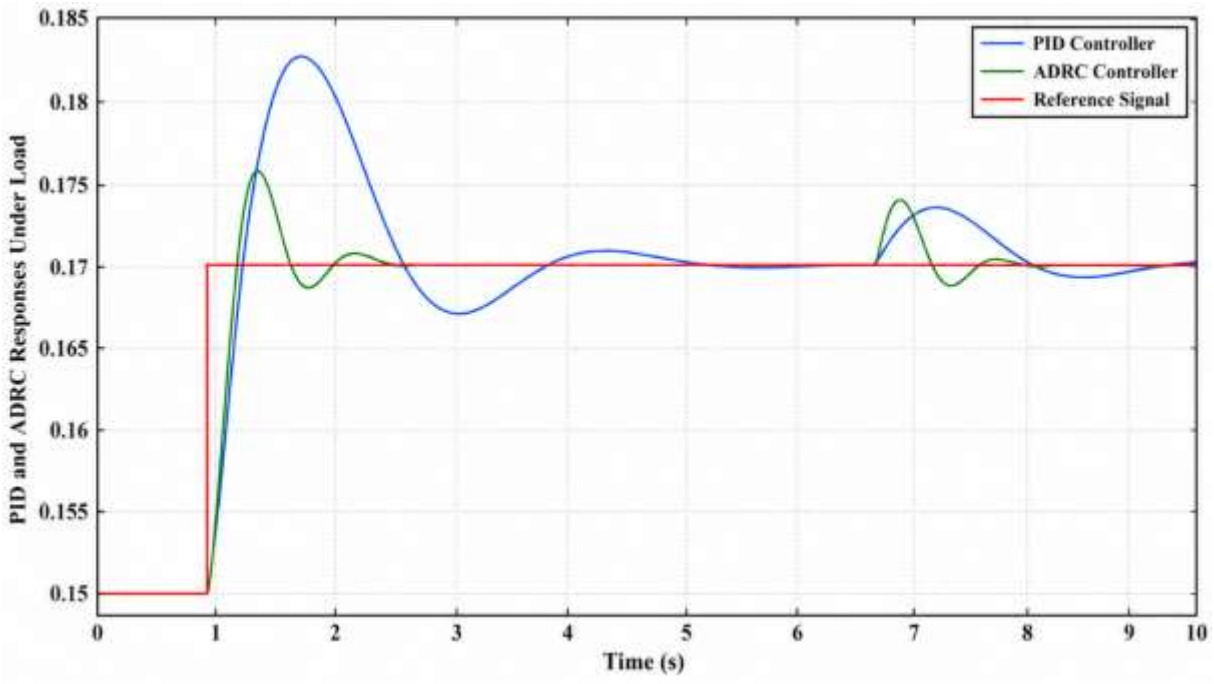


Figure 3. PID and ADRC responses under an external load disturbance applied at $(t = 7\text{s})$.

Before the disturbance is introduced, both controllers maintain the levitated object close to the desired position. At exactly $t = 7s$ the external load produces a deviation from the reference position for both controllers. The ADRC-controlled system responds rapidly to the disturbance. A relatively short transient oscillation occurs, after which the output quickly returns to the desired reference value. The amplitude and duration of the disturbance-induced oscillation are limited. The PID controller also compensates for the disturbance and eventually restores the desired position. However, its response exhibits a slower and more weakly damped recovery. The effect of the disturbance persists for a longer period compared with the ADRC response. This difference is consistent with the operating principle of ADRC. The Extended State Observer estimates the generalized disturbance and allows its effect to be actively compensated through the control input. The quantitative disturbance-rejection results can be presented in Table 3.

Table 3. Performance Comparison Under Load Disturbance

Performance criterion	PID	ADRC	Improvement of ADRC (%)
Maximum disturbance deviation (Δh_{max}) (m)	0.0036331	0.0045432	-25.052%
Disturbance recovery time (Tdr) (s)	2.3025	0.77475	66.352%
Maximum value after disturbance (m)	0.17363	0.17454	-25.052%
Minimum value after disturbance (m)	0.16913	0.16867	-53.74%
IAE after (t = 7 s)	0.0037055	0.001699	54.15%
ISE after (t = 7 s)	8.7526×10^{-6}	4.691×10^{-6}	46.404%

For the disturbance analysis, it is preferable to calculate the error indices only over the interval $7 \le t \le 10\text{s}$ so that the initial reference-tracking transient does not affect the disturbance-rejection comparison. Accordingly,

$$\int_7^{10} |e(t)|\,dt,] \quad \int_7^{10} e^2(t)dt,] \quad \text{and} \int_7^{10} (t-7)\,|e(t)|\,dt \qquad [11]$$

The simulation results obtained from both operating scenarios demonstrate the distinct dynamic characteristics of the two controllers. Under nominal no-load conditions, the PID controller successfully stabilizes the nonlinear MagLev system but exhibits a relatively high overshoot and a longer oscillatory transient. ADRC reduces the magnitude of the initial overshoot and provides faster attenuation of the oscillations. Under the external load condition, both controllers restore the levitated object to the

reference position. Nevertheless, ADRC rejects the disturbance more rapidly and returns to the reference with a shorter transient response.

V. CONCLUSION

In this study, conventional PID control and Active Disturbance Rejection Control were comparatively investigated for the position control of a nonlinear magnetic levitation system. The nonlinear dynamics of the MagLev system were modeled directly in MATLAB/Simulink using the relationship between gravitational force, electromagnetic force, coil current, and levitation position. Both controllers were implemented on the same nonlinear plant and evaluated under identical initial conditions, reference signals, and control-input constraints. Two simulation scenarios were considered. In the first scenario, the reference-tracking performances were investigated under no-load conditions. Both PID and ADRC successfully stabilized the nonlinear MagLev system and achieved the desired steady-state position. However, the PID controller produced a larger overshoot and a longer oscillatory transient, whereas ADRC achieved faster convergence with reduced overshoot. In the second scenario, an external load disturbance was applied at $(t = 7\text{s})$. Both controllers were capable of rejecting the disturbance and restoring the output to the desired reference value. Nevertheless, ADRC demonstrated faster disturbance attenuation and a shorter recovery period than the conventional PID controller. The overall simulation results therefore demonstrate that classical ADRC provides superior transient performance and disturbance-rejection capability compared with conventional PID control for the nonlinear magnetic levitation system considered in this study. Future work may include experimental implementation of the proposed control structures, evaluation under parameter uncertainties, different disturbance amplitudes, measurement noise, actuator constraints, and real-time hardware implementation.

REFERENCES

[1] Mrotek, M., Michalski, J., Zurita-Bustamante, E. W., Madonski, R., & Pazderski, D. (2025). Embedding System Knowledge in Nonlinear Active Disturbance Rejection Control: Insights from a Magnetic Levitation System. Electronics, 14(24), 4811.

[2] El-Sayed, A. T., Hussein, R. K., Amer, Y. A., & El-Sayed, M. A. (2026). Nonlinear dynamics, stability, and vibration suppression of magnetically levitated systems under a combined control strategy. Journal of Low Frequency Noise, Vibration and Active Control, 14613484261434302.

[3] Vázquez, C. G., Cotfas, D. T., Santos, A. G., Cotfas, P. A., Ávila, L., & Garrido, V. P. (2025). Optimal tuning of PID controllers focused on energy efficiency in a multivariable and coupled system. Case study: An AHU in the biopharmaceutical industry. Applied Thermal Engineering, 269, 126053.

[4] Hong, Z., Cao, Y., Jin, C., Xu, B., Zhou, J., & Gao, J. (2026). Research on Disturbance Rejection of Active Magnetic Bearing System Based on Generalized Integrator-Extended State Observer. Journal of Vibration Engineering & Technologies, 14(1), 50.

[5] Zeng, X., Li, S., Kuo, C. H., & Sariyildiz, E. (2025). Passive levitation in long vertical range using two permanent magnets. IEEE/ASME Transactions on Mechatronics.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Retrieval-Augmented Forecasting: A Structured Review of Time-Series Foundation Models, Retrieval Strategies, and Application Domains

Ali YILDIZ[1]*, Ali AKDAĞLI[2]

[1]*Department of Electrical and Electronics Engineering, Mersin University, Mersin, Türkiye*

**(yildiz@mersin.edu.tr)*

ORCID ID: 0000-0003-3904-6017

[2]*Department of Electrical and Electronics Engineering, Mersin University, Mersin, Türkiye*

akdagli@mersin.edu.tr

ORCID ID: 0000-0003-3312-992X

Abstract – Retrieval-Augmented Forecasting (RAF) augments parametric forecasting with dynamically retrieved historical sequences, analogous entities, or external context. This study presents a structured thematic synthesis of 28 full-text studies identified between 2022 and 11 August 2026 and organizes them around knowledge-base construction, query representation, candidate retrieval, relevance control, context integration, forecasting backbone, and calibration. Publication status is considered separately for peer-reviewed papers, preprints, and conference submissions. The evidence indicates that RAF evolved from cross-entity and reference-based forecasting in 2022 toward learnable retrievers, time-series foundation models, and retrieval-guided diffusion after 2024. Recent studies clarify three emerging directions: task-aligned retrieval trained with LLM feedback; diffusion guidance using channel dependence, physical constraints, and visual-semantic representations; and reliability-aware prediction revision based on retrieved residual memories. Reported results suggest benefits under data scarcity, long horizons, distribution shift, and probabilistic forecasting. However, excessive or irrelevant retrieval, temporal leakage, unprecedented regimes, computational overhead, and fragmented evaluation remain major concerns. Reliable RAF therefore requires regime- and physics-aware retrieval, calibrated fallback mechanisms, leakage-free benchmarks, uncertainty evaluation, and traceable evidence use.

Keywords – Retrieval-Augmented Forecasting, time-series forecasting, time-series foundation models, zero-shot forecasting, information retrieval, retrieval augmentation

I. INTRODUCTION

Time-series forecasting supports decision making in energy production and demand, financial risk, traffic flow, meteorology, health monitoring, and industrial maintenance. Classical statistical models offer explicit assumptions and relatively low computational cost, whereas deep learning models capture nonlinear relationships and long dependencies. Time-Series Foundation Models (TSFMs) extend this line of work by pretraining on large multi-domain collections and then transferring to new datasets in zero-shot or few-shot settings. A fixed parametric model, however, cannot reliably encode every rare event, unseen regime, or domain-specific dependency that may appear at inference time.

Retrieval-Augmented Forecasting (RAF) provides a complementary solution. Instead of increasing model capacity only through more parameters or retraining, RAF retrieves relevant examples from an external memory at forecast time and combines them with the current query. In time series, the retrieved

evidence may consist of similar historical windows, related entities, covariate regimes, textual context, or physical descriptors. The central promise is to combine general knowledge compressed into model parameters with explicit case-based evidence stored outside the model.

The retrieval idea predates the RAF label. MQRetNN used approximate retrieval across more than two million products and reported approximately 3% lower test loss at small scale and about 1% at large scale [1]. ReTime combined relational retrieval with content synthesis and argued that reference series can reduce learning uncertainty, particularly when the forecast horizon is long relative to the observation window [2]. These studies established two assumptions that later RAF methods make explicit: useful analogues can exist in an external memory, and the futures associated with those analogues can reduce predictive uncertainty.

RAF became a distinct paradigm after 2024. Tire et al. formalized a model-agnostic framework in which retrieved motifs and their futures are supplied to TSFMs [3]. Retrieval-augmented diffusion conditioned probabilistic generation on references [4]. Two studies named TimeRAG then explored different routes: Yang et al. [5] transformed DTW-retrieved sequences into numerical text prompts for an LLM, whereas Xiao et al. [6] used LLM feedback to train a task-aligned retriever. Later work introduced end-to-end retriever learning [7], adaptive retrieval mixing [8], multiscale lightweight retrieval [9], input-aware cross-attention [10], stationarity-aware diversity [11], semantic retrieval [12], invariant-dynamic decomposition [13], and several forms of retrieval-guided diffusion [14]-[16].

Despite this rapid expansion, terminology remains fragmented. 'Retrieval', 'reference', 'analogy', 'in-context learning', and 'post-hoc revision' are sometimes used for overlapping mechanisms. High reported accuracy does not by itself establish that the temporal split is leakage-free, that retrieved candidates were actually available at forecast time, or that gains are attributable to retrieval rather than additional context or computation. This review therefore compares RAF studies through a common system-level pipeline and evaluates both reported benefits and reliability risks.

II. MATERIALS AND METHOD

The study was conducted as a structured thematic review of Retrieval-Augmented Forecasting. The search window covered publications and public versions from 1 January 2022 through 11 August 2026. Search expressions included 'retrieval-augmented forecasting', 'retrieval augmented time series forecasting', 'retrieval-guided forecasting', 'retrieval-based time series forecasting', 'time-series RAG', 'reference-guided diffusion forecasting', 'retrieval-based revision', and closely related formulations. arXiv, OpenReview, ACL Anthology, NeurIPS, ICML, KDD, AAAI, DOI records, and publisher pages were cross-checked when publication status or version history was ambiguous.

A study was included when it performed an explicit retrieval operation during training or inference; selected a historical window, another entity, textual context, physical regime, or comparable external record; used the retrieved information in a way that could directly affect the forecast; and reported an empirical evaluation. Papers discussing generic RAG without forecasting, using retrieval only as a data-collection synonym, or focusing solely on classification or question answering were excluded. Multiple public versions of the same study were treated as a single record.

All 28 identified records were examined in full text. Each paper was coded by knowledge-base type, query representation, similarity or retrieval mechanism, number of retrieved candidates, relevance control, integration mechanism, forecasting backbone, datasets, horizon, baselines, metrics, ablation findings, computational implications, and stated limitations. Publication status was recorded separately because a withdrawn submission or a paper under review should not carry the same evidential weight as a peer-reviewed journal or conference publication. Numerical improvements were reproduced only when explicitly reported in the source study; no meta-analytic average was computed because data splits, error scales, and compute budgets were not sufficiently homogeneous.

The synthesis was organized into four themes: chronological development; knowledge-base and retrieval design; integration and forecasting architecture; and application, evaluation, and reliability. This structure was chosen to compare methods at the level of the complete RAF pipeline rather than simply

listing model names. The chronological scope and changing methodological emphasis of the included literature are summarized in Table 1.

Table 1. Scope and chronological development of the reviewed RAF literature.

First public year	Full-text studies	Dominant direction
2022	2	Cross-entity retrieval and reference-series uncertainty reduction.
2023	0	No direct new record; TSFM and LLM forecasting infrastructure matured.
2024	6	Explicit RAF formulations; early LLM-, TSFM-, and diffusion-based architectures.
2025	10	Learnable retrieval, adaptive integration, tabular and multimodal expansion.
2026*	10	Relevance control, stationarity, semantics, industrial and agent-based methods.

**2026 count is through 11 August 2026; years refer to first public versions.*

III. RESULTS

A. *End-to-End RAF Pipeline*

A RAF system does not rely only on a parametric mapping $f(x)$ from an observed sequence x to a future horizon H. Instead, an external knowledge base $D = \{(k_i, v_i)\}$ stores keys representing historical windows or learned representations and values representing subsequent futures or additional context. A retriever $r(x,D)$ selects candidate records; a relevance stage can rerank or filter them; an integrator combines the selected evidence with the query; and the forecasting head produces the final output. This decomposition separates the quality of the memory, the retrieval decision, and the forecasting model itself.

Offline knowledge-base preparation is often a hidden but decisive stage. Historical sequences are divided into past-future pairs, normalized or embedded, indexed, and accompanied by time and entity metadata. Candidate windows that overlap the query horizon or contain information not available at the forecast origin must be removed. In cross-series retrieval, differences in units, sampling frequency, sensor quality, and entity identity must also be aligned. Without these controls, apparently strong retrieval gains can be artifacts of leakage or near-duplicate matching.

At inference time, four questions dominate retrieval design: how the query is represented, how similarity is measured, how many candidates are retrieved, and how their relevance is controlled. Raw Euclidean distance and Dynamic Time Warping (DTW) preserve local shape, while learned embeddings compress trend, periodicity, channel interactions, or semantic information. Crucially, input similarity and future similarity are not identical; two near-identical historical windows can transition into different future regimes. Therefore, robust RAF increasingly couples top-k retrieval with reranking, diversity, stationarity, semantics, physics, or rejection mechanisms. Figure 1 summarizes this end-to-end workflow from knowledge-base preparation to forecast and uncertainty output.

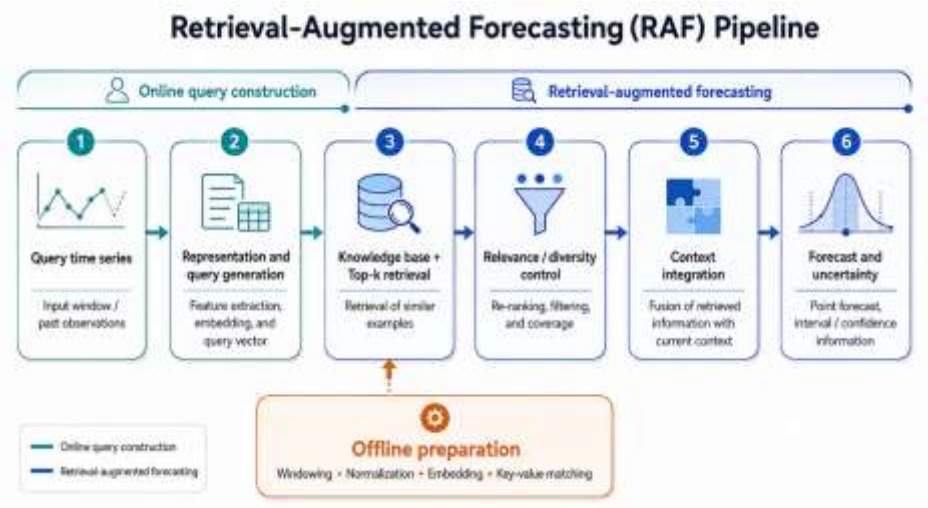


Fig. 1. End-to-end RAF pipeline from offline knowledge-base preparation to forecast and uncertainty output.

B. *Chronological Development and Retrieval Strategies*

The literature can be divided into four broad phases. The first phase, represented by MQRetNN and ReTime in 2022, treated retrieval primarily as a mechanism for using the memory of related entities or reference series under computational constraints [1], [2]. The second phase began in 2024, when RAF was explicitly articulated and combined with foundation models, LLM prompting, and diffusion [3]-[6]. The third phase, concentrated in 2025, made both retrieval and integration increasingly learnable through task-specific knowledge bases, adaptive mixers, multiscale retrieval, schema-aware attention, and urban transfer [7]-[9], [18]-[21]. The fourth phase, visible in 2026, focuses on whether retrieved evidence is trustworthy: cross-attention, stationarity-aware diversity, semantic channels, physics constraints, visual representations, decomposition, gated revision, sequence-native in-context calibration, and reference-token fusion all aim to prevent irrelevant analogues from dominating the forecast [10]-[16], [22]-[24], [28].

Within-series motif retrieval is the simplest memory construction. A key is a past window of length L and the corresponding value is the following H-step future. This setting is attractive because labels arise naturally and no external dataset is required. It is effective when seasonal motifs or regimes recur, but it fails when the history contains few useful analogues or when the system enters a genuinely new regime. Cross-entity retrieval expands the pool to other products, sensors, stations, or cities, which can reduce data scarcity but introduces scale, calendar, sensor, and domain mismatch risks.

Retrieval representation has evolved from raw distance to learned and multimodal similarity. DTW requires no training and tolerates local temporal shifts. Learned retrievers instead optimize an embedding for task utility: TimeRAF learns retrieval with forecasting loss [7], Ning et al.'s TS-RAG uses pretrained time-series encoders and adaptive mixing [8], and Xiao et al.'s TimeRAG attempts to align retriever scores with the forecasting model's own directional confidence [6]. TRACE aligns text and time series in a shared space [17], whereas VARDiff transforms series into Gramian Angular Field images and retrieves using visual features [16]. These approaches broaden the meaning of 'similarity' from pointwise closeness to structural, semantic, and task-relevant proximity.

The choice of k is similarly not monotonic. PhysDiffWind obtained its best errors at moderate k values and degraded when too many references were supplied [15]. VARDiff also favored small reference sets [16], while MiDDiR showed that excessive guidance can pull the generated distribution too strongly toward the retrieval memory [14]. These findings support query-adaptive retrieval and explicit rejection rather than a fixed large top-k policy.

C. *Integration Mechanisms and Forecasting Backbones*

Retrieved evidence can enter the forecasting system at different stages. Prompt-based approaches serialize retrieved numerical sequences and append them to an LLM context [5], [6]. Foundation-model approaches insert dedicated tokens, channel prompts, or learned mixtures into latent representations [7], [8], [10]. Han et al.'s Retrieval Augmented Time Series Forecasting method (RAFT) demonstrates that retrieval is not limited to very large foundation models; multiscale retrieval can also strengthen an MLP backbone when retrieval supplies a useful inductive bias [9]. Tabular RAF further incorporates schema metadata and learned temporal lags [18]. Baguan-TS uses a sequence-native 3D Transformer with retrieval-based local calibration [24], whereas Xiao et al.'s 2026 TS-RAG introduces dedicated reference tokens for fusing retrieved sequences [28]. Spatio-temporal methods combine retrieval with graph or spatial encoders [25].

Table 2. Representative RAF architectures, reported findings, and principal limitations.

Architecture / study	Integration	Representative reported finding	Principal limitation
MQRetNN [1]	Cross-entity attention + retrieval	~3% test-loss improvement at small scale and ~1% at >2M products.	Gain may shrink with scale; entity similarity is critical.
TimeRAG [5]	DTW analogues in LLM prompt	Average 2.97% increase in base LLM forecasting accuracy.	Token cost and numerical text representation.
TS-RAG [8]	Adaptive Retrieval Mixer	Up to 6.84% improvement over existing TSFMs on seven datasets.	Extra encoder and integration training.
RAFT [9]	Multiscale retrieval + MLP	Average 86% win rate against contemporary baselines across ten datasets.	Sensitive to retrieval pool and k.
Tabular RAF [18]	Schema mixing + temporal cross-attention	19.1–26.5% sMAPE reduction; stronger robustness to schema changes.	Higher training cost.
Retrieval-guided diffusion [14]-[16]	Guided denoising / cross-attention	Large reported gains in CRPS/MSE/MAE across distinct domains.	Protocols are not directly comparable; diffusion and reference quality add cost.
Decompose-Learn-Revise [22]	Similarity-gated residual revision	Best average RMSE/MAE rank across 14 univariate datasets.	Total gain also reflects decomposition and selective learning.

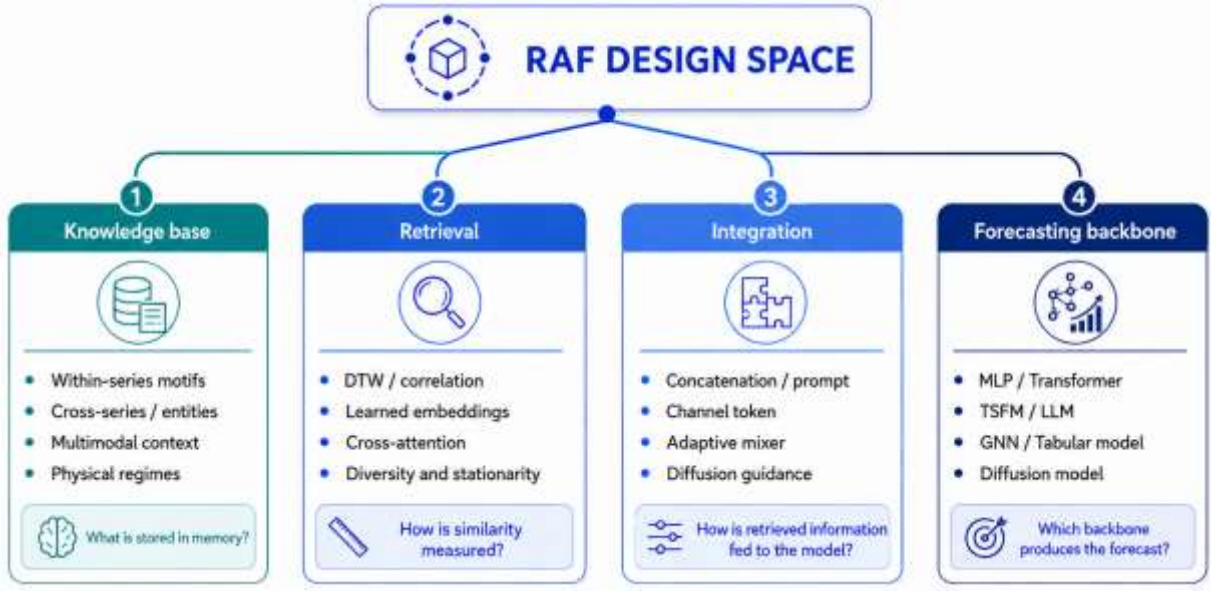


Fig. 2. RAF design space across knowledge base, retrieval, integration, and forecasting backbone.

Diffusion-based RAF conditions an entire denoising trajectory rather than only adjusting an initial input. RATD supplies references as conditions [4]. MiDDiR introduces retrieved futures through score guidance and combines channel-dependent history encoding with channel-independent denoising [14]. PhysDiffWind couples retrieved meteorological analogues with pressure-gradient and air-density constraints [15]. VARDiff retrieves visual analogues and injects them through decoder cross-attention [16]. This family can improve probabilistic forecasting but amplifies the consequences of a poor reference because the retrieved signal can influence many denoising steps.

Retrieval can also operate after the base forecast. Decompose-Learn-Revise stores historical residual errors together with training windows and applies a similarity-gated residual correction only when the retrieved neighborhood is sufficiently reliable [22]. This design changes retrieval from a generative context provider into a reliability-aware revision layer and illustrates a broader trend toward conditional use of retrieved evidence. Figure 2 organizes these alternatives across the knowledge base, retrieval, integration, and forecasting-backbone dimensions. Table 2 compares representative architectures, reported findings, and principal limitations.

D. Application Domains

General benchmark datasets such as Electricity, Weather, Traffic, and the ETT family remain the dominant test bed because they support common horizons and multivariate settings. Yet identical dataset names do not guarantee comparable experiments: normalization, input length, retrieval-pool construction,

random seeds, sample counts, and compute budgets can differ substantially. Consequently, cross-paper ranking should be interpreted with caution.

Demand and economic forecasting provide a natural cross-entity setting. MQRetNN retrieves across large product populations [1], while tabular RAF uses macroeconomic and financial tables [18]. The tabular study reported 19.1%-26.5% lower sMAPE and improved directional accuracy under market shocks, but economic forecasting requires special care because revised indicators, changing constituent definitions, and transaction costs complicate truly real-time evaluation.

Financial applications illustrate the importance of retrieval representation. The withdrawn TimeRAG submission reported that an LLM-feedback-trained retriever outperformed generic text embeddings for directional prediction on four U.S. stock datasets [6]. Peer-reviewed VARDiff reported average MSE and MAE reductions of 16.27% and 8.12%, respectively, on nine ACL18 StockNet stocks, with the retrieval database restricted to training-period examples [16]. Decompose-Learn-Revise used retrieved historical residuals for gated corrections on AAPL, AMZN, and NVDA [22]. These predictive improvements should not be conflated with trading profitability because price-level error, direction classification, and economic return are distinct outcomes.

Meteorology and offshore wind forecasting benefit from both spatial analogues and physical side information. Resolution-Aware RAF changes the geographical neighborhood according to temporal frequency [19]. PhysDiffWind combines similar multivariate meteorological histories with pressure-gradient and air-density constraints and reported up to 55.1% MSE and 49.3% CRPS reduction relative to its baselines [15]. The authors also identify data quality, unseen extreme conditions, horizons beyond 24 hours, and diffusion cost as deployment constraints.

Traffic and urban systems add graph structure and transfer across cities. RAST retrieves fine-grained patterns from spatio-temporal repositories [25], whereas STRATA-TS selects source-city slices for a data-scarce target city and distills structured LLM reasoning into a smaller model [20]. Industrial maintenance prioritizes physical consistency and false-alarm control: Liang et al.'s Retrieval-Augmented Generation with Covariate Time Series framework (RAG4CTS) stores raw historical regimes in a hierarchical memory and has been deployed with Apache IoTDB for pressure-regulating and shutoff-valve monitoring [23]. Event forecasting expands the retrieval concept further by searching time-stamped textual evidence rather than numerical motifs [26]. Table 3 consolidates these application domains and the validation requirements specific to each one.

Table 3. RAF application domains and domain-specific validation requirements.

Domain	Representative methods / data	Priority metrics	Domain-specific validation
General time series	ETT, Electricity, Weather, Traffic; MiDDiR	MSE, MAE, CRPS, QICE, long-horizon error	Unseen datasets, multiple seeds, horizon-wise analysis
Demand	MQRetNN; large product populations	Quantile loss, scalability	New products, sparse demand, out-of-stock periods
Finance / economy	Tabular RAF, TimeRAG, VARDiff, Decompose-Learn-Revise	sMAPE, MCC, MSE/MAE, CRPS, direction accuracy	Real-time data vintages, leakage control, transaction-aware testing
Meteorology	Resolution-Aware RAF; PhysDiffWind	MSE, CRPS, calibration, physical consistency	Unseen locations, extreme weather regimes, long horizons
Traffic / urban	RAST, STRATA-TS	MAE, RMSE, transfer performance	New cities, sensor failures, calendar shifts
Industrial maintenance	RAG4CTS; PRSOV monitoring	False alarms, early-warning lead time	Physical consistency, field-expert review, multi-site validation
Event forecasting	Retrieval-supported LLM systems	Brier score, calibration	Publication time, source reliability, auditability

E. Evaluation Problems and Reliability Risks

RAF evaluation should answer three counterfactual questions. First, how strong is the forecasting backbone without retrieval? Second, what happens when the model receives the same amount of randomly selected context? Third, can the retriever identify useful evidence when the true future of the query is unknown? Reporting only the difference between a RAF system and its base model cannot determine whether the gain comes from retrieval quality, a longer context, additional parameters, or extra training data. Random retrieval, nearest-neighbor retrieval, oracle retrieval, past-only retrieval, and retrieved-future-only ablations provide complementary evidence about where the improvement originates.

Oracle conditions are useful for diagnosing headroom but must be clearly separated from deployable retrieval because they use information unavailable at forecast time.

Temporal information leakage is the most serious validity risk. During knowledge-base construction, windows overlapping the query forecast horizon, observations added after the forecast origin, or retrospectively revised economic indicators must not be accessible to the retriever. When dense sliding windows are created from one long series, train and test examples can share most of their observations even when their nominal start times differ. A very small retrieval distance can then represent near-copying rather than genuine analogue-based generalization. Chronological splits should therefore be defined by forecast origin, and entity-level separation should be added when cross-entity generalization is claimed. For continuously updated systems, the time at which an index was built or refreshed should also be recorded.

A second structural risk is that similar histories need not imply similar futures. Regime change, policy intervention, maintenance, equipment replacement, market shocks, or extreme weather can redirect a process after two nearly identical observation windows. This limitation is visible in recent ablations: PhysDiffWind and VARDiff obtain their strongest results with moderate reference counts, whereas excessive retrieval can dilute the conditioning signal [15], [16]; MiDDiR shows that overly strong guidance can pull the predictive distribution too close to the retrieval memory [14]. Stationarity-Aware Retrieval-Augmented Forecasting (SARAF), Semantics-Enhanced Retrieval-Augmented Forecasting (SERAF), and invariant-dynamic decomposition address the problem through stationarity, semantic evidence, and component separation [11]-[13], while Decompose-Learn-Revise suppresses the revision when similarity is insufficient [22]. These mechanisms support explicit calibration of retrieval confidence rather than assuming that every retrieved neighbor is informative.

Reliability assessment should extend beyond average error. Mean MSE or MAE can hide rare but operationally important failures, so distribution shift, long-tail errors, horizon-specific degradation, and difficult regimes should be reported separately. Probabilistic forecasting additionally requires CRPS, QICE, empirical coverage, and calibration diagnostics. For point forecasting, worst-slice errors and retrieval-confidence strata are useful complements to average performance. Statistical testing is also valuable when many datasets are compared: Decompose-Learn-Revise, for example, reports Friedman and pairwise Wilcoxon tests across 14 univariate datasets [22]. However, statistical significance does not remove design limitations; a result restricted to univariate point forecasting cannot automatically be generalized to multivariate or probabilistic RAF.

Traceability and explainability are further requirements for production systems. Displaying a retrieved sequence is not sufficient if users cannot determine which part of the evidence matched the query, how much it influenced the forecast, or whether the model would have produced a similar output without it. Semantic alignment in TRACE [17], structured relational hints in STRATA-TS [20], and per-reference cross-attention in Cross-RAG [10] provide useful signals, but attention weight alone is not a causal explanation. Counterfactual removal of retrieved examples, perturbation of the retrieval set, and influence analysis are stronger ways to verify that an explanation corresponds to an actual forecast mechanism. Provenance should therefore accompany each retrieved item with source identity, timestamp, similarity or relevance score, and the decision path by which it affected the final prediction.

F. Interpreting the Strength of Evidence

Reported improvements should be interpreted in relation to publication status and experimental scope. The review contains peer-reviewed journal and conference papers, preprints, workshop papers, a withdrawn ICLR submission, and a study submitted to ICLR 2026. The withdrawn TimeRAG work is methodologically informative because it illustrates retriever training with LLM feedback, but its MCC values should not be treated as independently peer-reviewed evidence [6]. Likewise, MiDDiR provides useful preliminary evidence for retrieval-guided probabilistic diffusion, yet its reported CRPS and QICE gains remain submission-stage results [14]. Separating methodological novelty from evidential maturity prevents a fast-moving literature from being summarized as if all claims had passed the same review process.

The same caution applies to cross-domain numerical comparisons. A 55.1% MSE reduction in offshore wind forecasting [15], a 16.27% average MSE reduction in stock forecasting [16], and an 86% win rate across time-series benchmarks [9] answer different questions under different data scales, baselines, and metrics. They demonstrate that retrieval can be valuable in multiple settings, but they do not define a universal effect size. For this reason, the present review emphasizes architectural mechanisms, ablations, failure modes, and validation requirements rather than ranking all RAF methods in one league table. This system-level interpretation is especially important for a field in which the retrieval pool itself can function as an additional source of training information unless its construction is carefully controlled.

IV. DISCUSSION

The central conclusion across architectures is that retrieval is useful only when relevance is controlled. The field is moving from nearest-neighbor augmentation to selective evidence use. Cross-attention [10], stationarity-aware diversity [11], semantic retrieval [12], LLM-feedback alignment [6], physics-constrained diffusion [15], and gated residual revision [22] all address the same structural mismatch: similar historical inputs can have dissimilar futures because of interventions, maintenance, policy decisions, regime change, or extreme events.

Evaluation should therefore distinguish retrieval benefit from the effects of extra context, model capacity, and training data. A convincing RAF study should compare no retrieval, random retrieval, and meaningful retrieval under identical chronological splits and comparable compute. Oracle retrieval can provide an upper bound but is not a deployable baseline. The knowledge base must be frozen at the forecast origin, and candidate windows overlapping the query horizon or using data revisions unavailable at that time must be excluded. Dense sliding windows can otherwise create near-duplicate training and test examples that mimic strong generalization.

Average MSE or MAE also hides where retrieval fails. Long-tail errors, rare events, horizon-specific performance, and regime shifts should be reported separately. Probabilistic models require CRPS, QICE, coverage, and calibration in addition to average accuracy. Retrieval confidence should be treated as distinct from forecast confidence: disagreement among retrieved futures or weak nearest-neighbor similarity should reduce retrieval influence and trigger fallback to the base model. This is consistent with the conditional revision philosophy of instance-aware post-hoc methods [27] and Decompose-Learn-Revise [22]. Figure 3 synthesizes the components, safeguards, and expected benefits required for a production-oriented RAF system.

Computational overhead remains important for deployment. Retrieval introduces index construction, encoding, top-k search, reranking, integration, and sometimes additional denoising. In tabular RAF, training time per epoch increased relative to the baseline [18]. In VARDiff, image transformation and visual encoding dominate retrieval overhead even though FAISS search itself is very fast [16]. PhysDiffWind is more resource-intensive than simple baselines because diffusion inference adds latency and memory demand [15]. Future reports should therefore include total latency, index refresh cost, token or denoising budgets, memory, and energy use alongside forecasting accuracy.

Five research priorities follow from the evidence. First, retrieval should incorporate regime, causal, and physical information rather than rely on shape alone. Second, leakage-free benchmarks should mimic production-time information availability and explicitly test new entities, unseen locations, long horizons, and abrupt shifts. Third, retrieval uncertainty and predictive uncertainty should be calibrated jointly, with dynamic fallback. Fourth, multimodal knowledge bases should connect numerical series with text events, maintenance notes, physical variables, schema, images, or graph position, while counterfactual ablations measure each modality's contribution. Fifth, RAF must become efficient and continuously updateable through incremental indexing, compact prototype or residual memories, approximate search, and dynamic k.

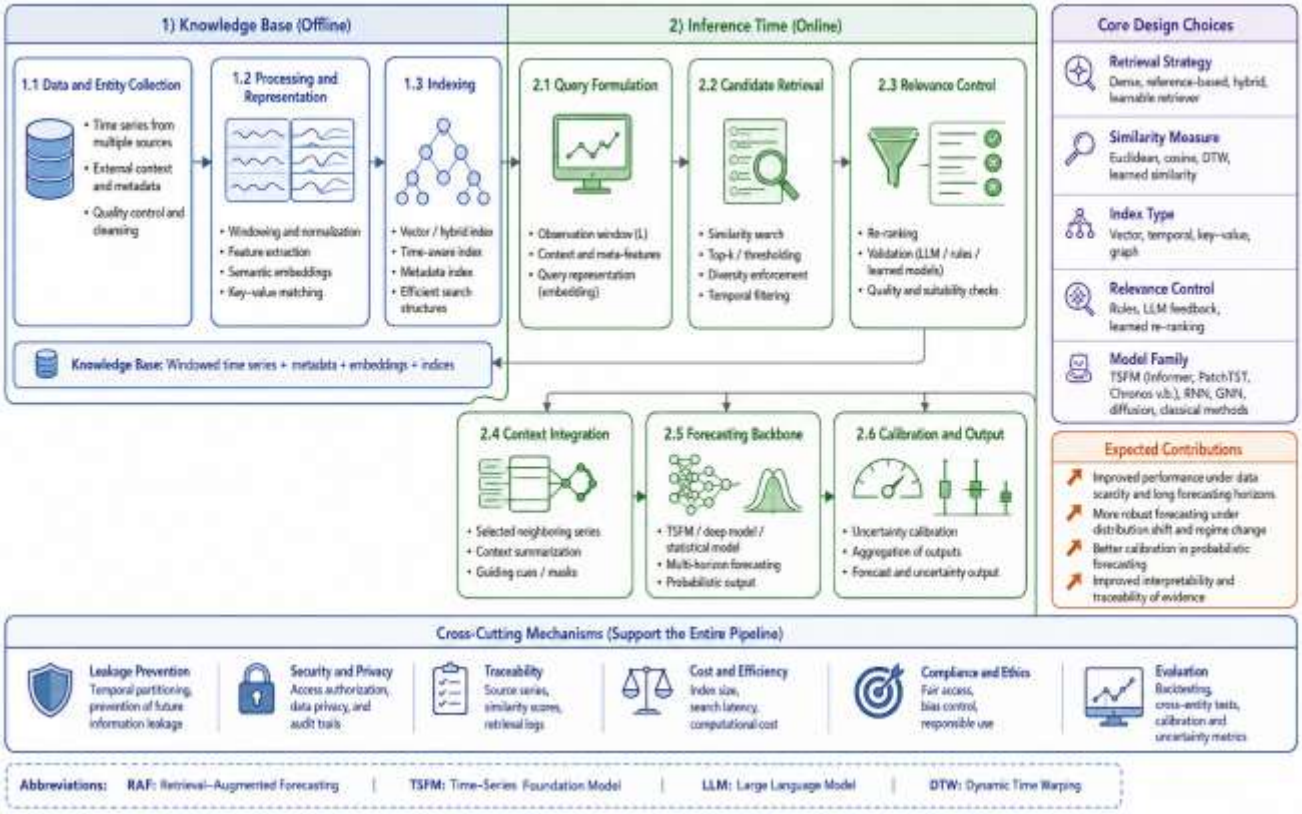


Fig. 3. Core components, design options, cross-cutting safeguards, and expected benefits of production-oriented RAF systems.

V. CONCLUSION

This review treats Retrieval-Augmented Forecasting not as a single similarity-search operation but as a compound system comprising knowledge-base preparation, query representation, retrieval, selection, integration, forecasting, revision, and calibration. Across 28 full-text studies, the field progresses from retrieval-based precursors in 2022 to explicit RAF formulations in 2024 and then to learnable, semantic, visual, physics-aware, diffusion-based, and reliability-oriented methods. Withdrawn or under-review submissions were retained only to document methodological development and were not interpreted as equivalent to peer-reviewed evidence.

Reported gains can be substantial, but they arise from heterogeneous tasks, data splits, metrics, and computational budgets and therefore should not be interpreted as a single meta-analytic effect or universal model ranking. The most important practical lesson is that retrieved evidence should be treated as uncertain evidence. Reliable RAF systems must quantify relevance, reject weak references, prevent temporal leakage, calibrate uncertainty, and preserve source traceability. When retrieval is combined with regime and physics knowledge, multimodal context, transparent evaluation, and human oversight, it can become a robust forecasting layer rather than an uncontrolled source of extra context.

REFERENCES

[1] S. Yang, C. Eisenach, and D. Madeka, “MQRetNN: Multi-horizon time series forecasting with retrieval augmentation,” 8th KDD Workshop on Mining and Learning from Time Series, 2022.
[2] B. Jing et al., “Retrieval based time series forecasting,” CIKM 2022 Workshop on Applied Machine Learning Methods for Time Series Forecasting, 2022.
[3] K. Tire, E. O. Taga, M. E. Ildiz, and S. Oymak, “Retrieval augmented time series forecasting,” arXiv:2411.08249, 2024.
[4] J. Liu, L. Yang, H. Li, and S. Hong, “Retrieval-augmented diffusion models for time series forecasting,” Advances in Neural Information Processing Systems, vol. 37, 2024.
[5] S. Yang, D. Wang, H. Zheng, and R. Jin, “TimeRAG: Boosting LLM time series forecasting via retrieval-augmented generation,” arXiv:2412.16643, 2024.

[6] M. Xiao et al., “TimeRAG: It’s time for retrieval-augmented generation in time-series forecasting,” withdrawn ICLR 2025 submission, OpenReview GvzL4LuycW, 2024.
[7] H. Zhang et al., “TimeRAF: Retrieval-augmented foundation model for zero-shot time series forecasting,” IEEE Transactions on Knowledge and Data Engineering, 2025.
[8] K. Ning et al., “TS-RAG: Retrieval-augmented generation based time series foundation models are stronger zero-shot forecaster,” Advances in Neural Information Processing Systems, vol. 38, 2025.
[9] S. Han, S. Lee, M. Cha, S. Ö. Arik, and J. Yoon, “Retrieval augmented time series forecasting,” Proc. 42nd International Conference on Machine Learning, PMLR 267, 2025.
[10] S. Lee and J. Lee, “Not all retrievals are useful: Cross-attention for input-aware RAG in time series forecasting,” arXiv:2603.14709, 2026.
[11] S. Zhou et al., “Stationarity-aware retrieval-augmented time series forecasting,” Proc. 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2026.
[12] S. Zhou et al., “Semantics-enhanced retrieval-augmented time series forecasting,” ICML 2026 Workshop on Forecasting as a New Frontier of Intelligence, arXiv:2606.14941, 2026.
[13] J. Chi et al., “Factorize to generalize: Retrieval-guided invariant-dynamic decomposition for time series forecasting,” arXiv:2605.24911, 2026.
[14] Y.-H. Huang and W.-J. Li, “Mixed channel dependency diffusion model with retrieval guidance for time series forecasting,” submitted to ICLR 2026, OpenReview 8inm5tmw9g, 2025.
[15] W. Dong, J. Che, Y. Mei, L. Zhou, and Y. Zhang, “PhysDiffWind: A physics-constrained retrieval-augmented diffusion framework for offshore wind speed forecasting,” Applied Soft Computing, vol. 190, 114647, 2026.
[16] T.-T. Nguyen, X.-T. Truong, T.-B. Nguyen-Khac, and N.-H. Nguyen, “VARDiff: Vision-augmented retrieval-guided diffusion for stock forecasting,” Information Sciences, vol. 736, 123113, 2026.
[17] J. Chen et al., “TRACE: Grounding time series in context for multimodal embedding and retrieval,” Advances in Neural Information Processing Systems, vol. 38, 2025.
[18] Z. Li, “Retrieval-augmented forecasting with tabular time series data,” Proc. 4th Table Representation Learning Workshop, pp. 192–199, 2025.
[19] I. Deznabi, P. Kumar, and M. Fiterau, “Resolution-aware retrieval augmented zero-shot forecasting,” arXiv:2510.16695, 2025.
[20] Y. Jiang et al., “STRATA-TS: Selective knowledge transfer for urban time series forecasting with retrieval-guided reasoning,” arXiv:2508.18635, 2025.
[21] G. Jalori, P. Verma, and S. Ö. Arik, “FLAIRR-TS: Forecasting LLM-agents with iterative refinement and retrieval for time series,” arXiv:2508.19279, 2025.
[22] X. Zhang and Z. Chen, “Decompose-learn-revise: A reliability-aware framework for robust time series forecasting,” Computers & Electrical Engineering, vol. 138, 111291, 2026.
[23] K. Y. Liang et al., “Retrieval-augmented generation with covariate time series,” arXiv:2603.04951, 2026.
[24] L. Yang et al., “Baguan-TS: A sequence-native in-context learning model for time series forecasting with covariates,” arXiv:2603.17439, 2026.
[25] W. Ruan, X. Dang, Z. Zhou, S. Lyu, and Y. Liang, “A retrieval augmented spatio-temporal framework for traffic prediction,” arXiv:2508.16623, 2025.
[26] D. Halawi, F. Zhang, C. Yueh-Han, and J. Steinhardt, “Approaching human-level forecasting with language models,” Advances in Neural Information Processing Systems, vol. 37, 2024.
[27] Z. Liu, M. Cheng, G. Zhao, J. Yang, Q. Liu, and E. Chen, “Improving time series forecasting via instance-aware post-hoc revision,” Advances in Neural Information Processing Systems, vol. 38, 2025.
[28] Y. Xiao, C. Xiao, and J. Zhou, “TS-RAG: Retrieval augmented generation for time series forecasting,” arXiv:2608.06223, 2026.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Liquid–Liquid Extraction of Chicago Sky Blue Using a Diamide–Diether Calix[4]arene Receptor

İbrahim Uyanık

Metallurgical and Materials Engineering Department, Technology Faculty, Selçuk University, Türkiye

iuyanik@selcuk.edu.tr

Abstract – Azo dyes are among the most common organic pollutants released from textile and related industries because of their extensive production and widespread use. Their high-water solubility and resistance to degradation make their removal from aqueous media an important environmental challenge. In the present study, a diamide–diether functionalized calix[4]arene receptor (C4-DADE) was synthesized and evaluated as an extractant for the liquid–liquid extraction of the carcinogenic azo dye Chicago Sky Blue (CSB). The extraction performance was examined by investigating the effects of solution pH, NaCl concentration, and contact time. Among the parameters studied, pH had the greatest influence on extraction efficiency. The highest extraction efficiency (89.1%) was obtained at pH 3.0, whereas the parent calix[4]arene (C4) exhibited negligible extraction under the same experimental conditions. Increasing the NaCl concentration from 0.05 M to 0.10 M produced only a slight improvement in extraction efficiency, while no further enhancement was observed at 0.20 M. The extraction equilibrium was reached within 30 min, indicating that longer contact times were unnecessary. Based on the experimental findings, a possible extraction mechanism involving hydrogen bonding, π–π interactions, and hydrophobic interactions at the liquid–liquid interface is proposed. These results show that C4-DADE can be used as an effective extractant for CSB under acidic conditions.

Keywords – Calix[4]arene, Liquid–liquid extraction, Chicago Sky Blue, Azo dye, Supramolecular chemistry

I. INTRODUCTION

Azo dyes are among the most widely used synthetic colorants because of their wide color range, good chemical stability, and relatively low production cost. They are extensively employed in textile, paper, leather, printing, and several other industries. It has been reported that azo dyes account for approximately 65–75% of the total synthetic dye production worldwide [1], [2], [3]. During dyeing processes, a considerable fraction of the applied dyes does not bind to fibers and is released into industrial wastewater [1], [2], [4]. Because of their resistance to biological degradation and photolysis, these compounds may remain in aquatic environments for long periods [1], [5]. Azo dyes not only cause visible color pollution but may also form aromatic amines during degradation. Some of these degradation products are toxic or carcinogenic [1], [4]. Therefore, the efficient removal of azo dyes from wastewater remains an important research topic. Several methods have been developed for the treatment of dye-containing wastewater, including adsorption, membrane separation, biological treatment, coagulation–flocculation, and advanced oxidation processes [4], [5]. Although these methods are effective, each has some limitations. Adsorbents require regeneration after repeated use, biological treatment is often less effective for highly stable azo dyes, and advanced oxidation processes usually require high energy input or expensive oxidants [5], [6]. Therefore, alternative separation methods are still being explored. Liquid–liquid extraction is a simple and effective separation technique based on the distribution of a solute between two immiscible liquid phases.

The method has been used in many separation processes because it requires simple equipment, provides rapid phase separation, and can be adapted to different target compounds [7]. The efficiency of liquid–liquid extraction, however, depends largely on the extractant. Conventional extractants mainly rely on differences in solubility between the aqueous and organic phases. A different strategy is to use supramolecular receptors capable of recognizing target molecules through non-covalent interactions. This approach may improve both extraction efficiency and selectivity. Calixarenes are well-known host molecules in supramolecular chemistry. Their three-dimensional cavities and easily functionalized upper and lower rims allow the introduction of different binding sites without changing the macrocyclic framework. As a result, calixarene derivatives have been investigated for the recognition and extraction of metal ions, organic compounds, and dye molecules [8], [9], [10]. Previous studies have shown that amino-functionalized, carboxylated, pyrazine-substituted, magnetic, and polymer-supported calixarene derivatives can efficiently remove different azo dyes from aqueous media [11], [12], [13], [14]. These studies also indicate that extraction performance is influenced not only by solution conditions such as pH and ionic strength, but also by the structure of the receptor and the nature of the dye molecule. In the present study, C4-DADE was evaluated as an extractant for CSB. The effects of pH, sodium chloride concentration, and contact time on the extraction efficiency were systematically investigated in order to assess the extraction performance of C4-DADE toward a representative sulfonated azo dye.

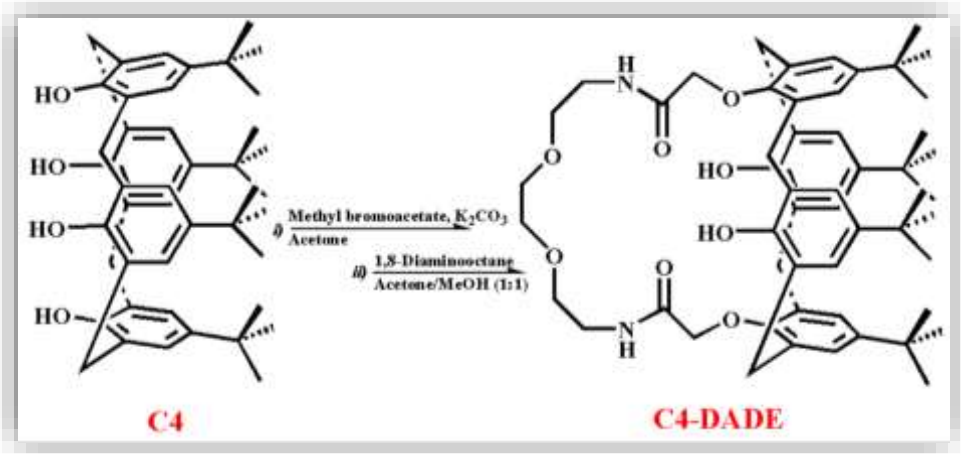
II. Materials and Method

A. Materials and Reagents

Chicago Sky Blue (CSB), dichloromethane (DCM), sodium chloride (NaCl), hydrochloric acid (HCl), sodium hydroxide (NaOH), and deionized water were used without further purification. All chemicals and solvents were obtained from Merck or Sigma-Aldrich. ^{1}H NMR spectra were recorded using a Bruker 400 MHz spectrometer. UV–vis absorption spectra were measured with a Shimadzu UV-1280. The pH values of the aqueous solutions were measured with a digital pH meter.

B. Synthesis of C4-DADE

C4-DADE was synthesized according to the literature procedure [15], [16]. The synthetic route is presented in Scheme 1. The product was obtained as a white solid in 49% yield; m.p. 238–240 °C. ^{1}H NMR (400 MHz, $CDCl_3$, δ, ppm): 0.85 (s, 18H, tert-Bu), 1.32 (s, 18H, tert-Bu), 3.37 (d, 4H, J = 13.49 Hz, $ArCH_2Ar$), 3.63 (s, 8H, $CH_2OCH_2CH_2OCH_2$), 3.76 (t, 4H, CH_2NH), 4.19 (d, 4H, J = 13.49 Hz, $ArCH_2Ar$), 4.48 (s, 4H, OCH_2), 5.99 (s, 2H, OH), 6.66 (s, 4H, ArH), 7.12 (s, 4H, ArH), 8.26 (t, 2H, NH). The ^{1}H NMR spectrum of the synthesized C4-DADE receptor is shown in Figure 1.



Scheme 1 Synthesis of the diamide–diether calix[4]arene receptor (C4-DADE).

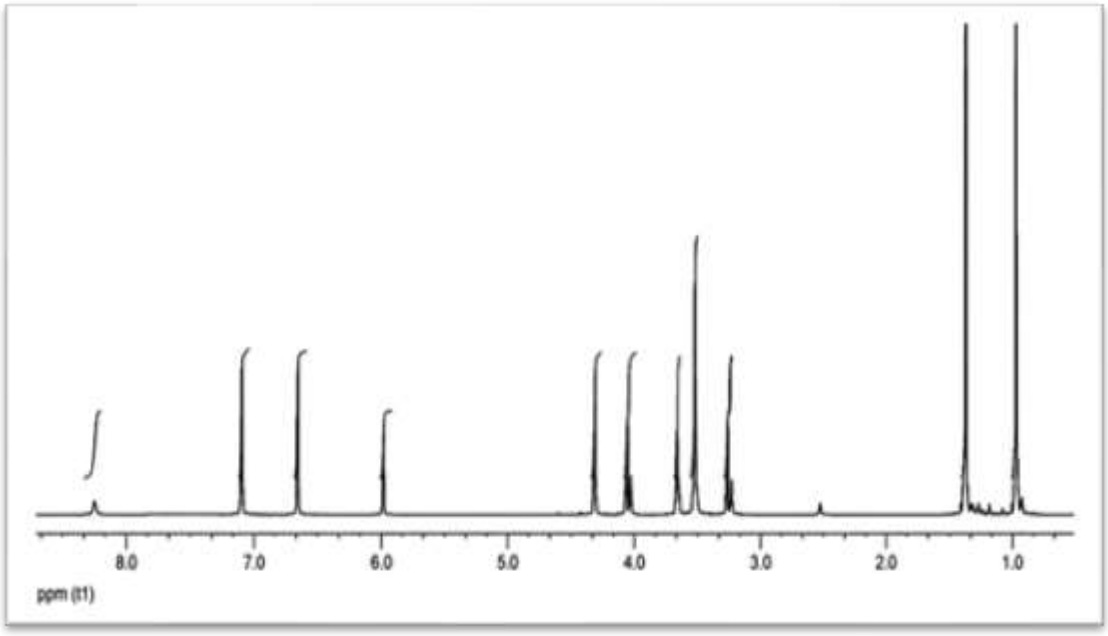


Fig. 1 ^{1}H NMR spectrum of C4-DADE.

C. *Liquid–Liquid Extraction Procedure*

Liquid–liquid extraction experiments were carried out to investigate the extraction behavior of C4-DADE toward CSB. A 1.0×10^{-3} M solution of C4-DADE was prepared in dichloromethane. Aqueous CSB solutions (2.0×10^{-5} M) were prepared at pH 3.0, 5.0, 7.0, 9.0, and 11.0. For each experiment, 10 mL of the aqueous dye solution was transferred into a glass vial, and 0.1 g NaCl was added. Subsequently, 10 mL of the C4-DADE solution was introduced, and the biphasic mixture was shaken for 1 h at 25 °C. The two phases were then allowed to separate naturally. The residual concentration of CSB in the aqueous phase was determined by UV–Vis spectrophotometry. For comparison, identical extraction experiments were also performed using the parent p-tert-butyl-calix[4]arene (C4) under the same experimental conditions.

D. *Determination of Extraction Efficiency*

The extraction efficiency (%) was calculated from the decrease in the absorbance of the aqueous phase before and after extraction according to the following equation:

$$\%E=[(A_0-A)/(A_0)]\times 100$$

where A_0 and A are the absorbance values of the aqueous dye solution before and after extraction, respectively

III. Results and Discussion

A. *Synthesis and Characterization of C4-DADE*

The synthesis of C4-DADE was accomplished in two consecutive steps starting from p-tert-butyl-calix[4]arene, as outlined in Scheme 1. The target receptor was isolated as a white solid in 49% yield with a melting point of 238–240 °C. The structure of C4-DADE was confirmed by ^{1}H NMR spectroscopy.

The ^{1}H NMR spectrum was consistent with the proposed structure. Two singlets observed at δ 0.85 and 1.32 ppm were assigned to the tert-butyl groups. The bridging methylene protons appeared as two doublets at δ 3.37 and 4.19 ppm (J = 13.49 Hz), indicating the preservation of the cone conformation of the calix[4]arene framework. The signals at δ 3.63, 3.76, and 4.53 ppm corresponded to the methylene protons of the diamide–diether bridge. The phenolic hydroxyl protons appeared as a singlet at δ 5.99 ppm, while the amide NH protons were observed as a triplet at δ 8.26 ppm. These spectroscopic data confirmed the successful synthesis of C4-DADE.

B. *Effect of Solution pH on CSB*

The effect of solution pH on CSB by C4-DADE was investigated over the pH range of 3.0–11.0, and the results are presented in Table 1. For comparison, identical experiments were also carried out using the parent p-tert-butyl-calix[4]arene under the same conditions.

Table 1. Effect of solution pH on the extraction of CSB by C4 and C4-DADE.

Receptor	pH 3.0	pH 5.0	pH 7.0	pH 9.0	pH 11.0
C4	<1.0	<1.0	<1.0	<1.0	<1.0
C4-DADE	89.1	81.0	61.5	13.0	5.0

Note: C4 represents the parent p-tert-butyl-calix[4]arene receptor.

As shown in Table 1, the parent calix[4]arene exhibited almost no extraction ability throughout the investigated pH range, with extraction efficiencies remaining below 1%. In contrast, C4-DADE showed a pronounced dependence on solution pH. The highest extraction efficiency (89.1%) was obtained at pH 3.0. The extraction gradually decreased with increasing pH, reaching 81.0% at pH 5.0 and 61.5% at pH 7.0. At alkaline pH values, the extraction efficiency dropped sharply to 13.0% at pH 9.0 and 5.0% at pH 11.0. The negligible extraction observed for the parent calix[4]arene indicates that the hydrophobic cavity alone is not sufficient for efficient transfer of CSB into the organic phase. The introduction of the diamide–diether bridge therefore plays an important role in the extraction process. The amide groups may interact with the sulfonate groups of CSB through hydrogen-bonding and other non-covalent interactions, while the overall geometry of the receptor may facilitate host–guest association. The strong pH dependence suggests that protonation of the receptor and the ionization state of the dye influence the extraction process. Under acidic conditions, these interactions appear to be more favorable, resulting in higher extraction efficiencies. As the pH increases, changes in the interaction between the receptor and the dye are likely to weaken the extraction process, leading to lower extraction efficiencies. Similar pH-dependent behavior has also been reported for other calixarene-based receptors used for azo dye extraction [12]. The results demonstrate that C4-DADE is considerably more effective than the parent calix[4]arene for the extraction of CSB, particularly under acidic conditions. Therefore, pH 3.0 was selected for the subsequent experiments.

C. *Effect of NaCl Concentration*

The influence of sodium chloride concentration on the extraction of CSB by C4-DADE was investigated at pH 3.0, and the results are summarized in Table 2.

Table 2. Effect of NaCl concentration on the extraction efficiency of CSB by C4-DADE (pH 3.0).

[NaCl]/ (M)	Extraction efficiency (%)
0.05	85.7
0.10	88.3
0.20	88.3

The extraction efficiency increased slightly from 85.7% at 0.05 M NaCl to 88.3% at 0.10 M NaCl. No further improvement was observed when the NaCl concentration was increased to 0.20 M, and the extraction efficiency remained unchanged (88.3%). The slight increase in extraction efficiency may be attributed to the increase in ionic strength of the aqueous phase, which can reduce the solubility of the dye

in water and facilitate its transfer into the organic phase. However, the nearly identical extraction efficiencies obtained at 0.10 M and 0.20 M NaCl indicate that further addition of electrolyte does not significantly influence the extraction process. Therefore, 0.10 M NaCl was selected for the subsequent experiments.

D. Effect of Contact Time

The extraction of CSB by C4-DADE was monitored at different contact times, and the results are summarized in Table 3. At 15 min, the extraction efficiency reached 87.7%. A slight increase was observed after 30 min, giving a maximum extraction efficiency of 89.1%. Increasing the contact time to 60 min did not produce any further improvement. These results indicate that extraction equilibrium was reached within 30 min under the selected experimental conditions. Therefore, extending the extraction time beyond 30 min was unnecessary, and this contact time was used in the subsequent experiments.

Table 3. Effect of contact time on the extraction efficiency of CSB by C4-DADE (pH 3.0).

Contact time (min)	Extraction efficiency (%)
15	87.7
30	89.1
60	89.1

E. Proposed Extraction Mechanism

Figure 2 illustrates a possible extraction mechanism for CSB by C4-DADE. During liquid–liquid extraction, the dye is initially present in the aqueous phase, whereas C4-DADE is dissolved in dichloromethane.

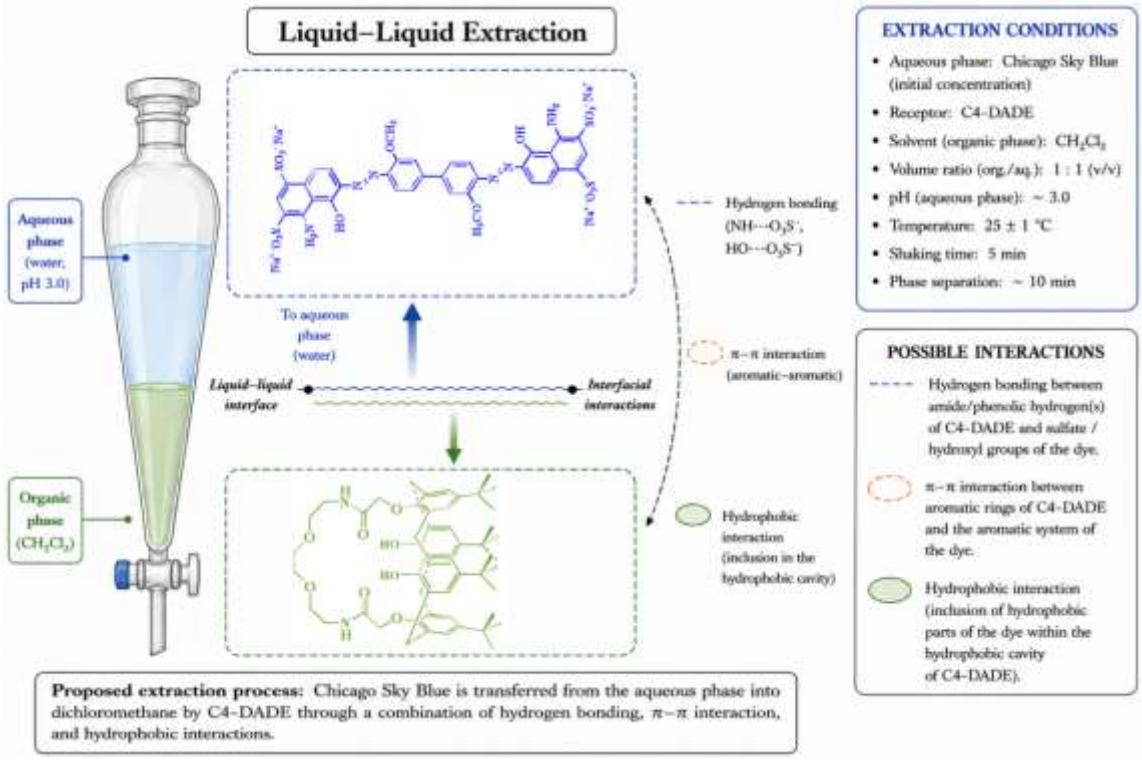


Fig. 2 Proposed liquid–liquid extraction mechanism of CSB by C4-DADE.

The two species are expected to interact at the liquid–liquid interface before the dye is transferred into the organic phase. In contrast to C4, which showed almost no extraction under the same conditions, C4-DADE exhibited a much higher extraction efficiency. This difference suggests that the diamide–diether moiety contributes to the extraction process. CSB contains sulfonate groups together with aromatic rings and azo linkages. These functional groups provide several possible interaction sites for C4-DADE. Hydrogen bonding involving the amide NH and phenolic OH groups is likely to contribute to the extraction process. In addition, π–π interactions between the aromatic rings and hydrophobic interactions within the calix[4]arene cavity may also assist the transfer of the dye into the organic phase. The hydrophobic cavity of the calix[4]arene may facilitate the transfer of the hydrophobic regions of the dye into the organic phase, while the aromatic framework of the receptor may contribute through π–π interactions with the aromatic rings of CSB. However, the relative contribution of each interaction cannot be established from the present study. The pronounced decrease in extraction efficiency with increasing pH suggests that the interaction between C4-DADE and CSB is sensitive to the solution environment. Under acidic conditions, the receptor provides a more favorable environment for extraction, whereas increasing pH weakens the overall interaction between the receptor and the dye. Similar pH-dependent behavior has also been reported for calixarene-based receptors used for azo dye extraction [12], [17]. Therefore, the extraction process is proposed to arise from the combined contribution of hydrogen bonding, π–π interactions, and hydrophobic interactions, as schematically illustrated in Figure 2.

IV. CONCLUSION

A diamide–diether calix[4]arene receptor (C4-DADE) was successfully applied to the liquid–liquid extraction of CSB from aqueous solution. Compared with the parent calix[4]arene (C4), C4-DADE exhibited substantially higher extraction efficiency under all experimental conditions investigated. The extraction process was strongly influenced by solution pH, with the highest extraction efficiency (89.1%) obtained at pH 3.0. In contrast, sodium chloride concentration had only a limited effect, and extraction equilibrium was established within 30 min. The experimental results suggest that the extraction process is governed by the combined contribution of hydrogen bonding, π–π interactions, and hydrophobic interactions at the liquid–liquid interface. Overall, C4-DADE provided higher extraction efficiency than the parent calix[4]arene and can be considered an effective receptor for the extraction of CSB from aqueous solution.

ACKNOWLEDGMENT

İU would like to thank the Chemistry Department, Faculty of Science, Selcuk University, Konya, Türkiye for utilizing their facilities to conduct the analyses.

REFERENCES

[1] H. Alzain, V. Kalimugogo, and K. Hussein, "A Review of Environmental Impact of Azo Dyes," *Int J Res Rev*, vol. 10, no. 6, pp. 64–689, Jun. 2023, doi: 10.52403/ijrr.20230682.

[2] D. Rawat, V. Mishra, and R. S. Sharma, "Detoxification of azo dyes in the context of environmental processes," *Chemosphere*, vol. 155, pp. 591–605, Jul. 2016, doi: 10.1016/j.chemosphere.2016.04.068.

[3] S. Benkhaya, S. M'rabet, and A. El Harfi, "Classifications, properties, recent synthesis and applications of azo dyes," *Heliyon*, vol. 6, no. 1, p. e03271, Jan. 2020, doi: 10.1016/j.heliyon.2020.e03271.

[4] D. Goswami, J. Mukherjee, C. Mondal, and B. Bhunia, "Bioremediation of azo dye: A review on strategies, toxicity assessment, mechanisms, bottlenecks and prospects," *Science of The Total Environment*, vol. 954, p. 176426, Dec. 2024, doi: 10.1016/j.scitotenv.2024.176426.

[5] F. Sevim, Ö. Laçin, F. Demir, and Ö. F. Erkiliç, "Adsorption Capacity, Isotherm, Kinetics, and Thermodynamics Examinations on the Removal of a Textile Azo Dye by Local Natural Adsorbent," *Global Challenges*, vol. 9, no. 5, p. 2500024, May 2025, doi: 10.1002/gch2.202500024.

[6] V. Vaiano and I. De Marco, “Removal of Azo Dyes from Wastewater through Heterogeneous Photocatalysis and Supercritical Water Oxidation,” *Separations*, vol. 10, no. 4, p. 230, Apr. 2023, doi: 10.3390/separations10040230.
[7] Ó. Martínez-Rico, A. Asla, Á. Domínguez, and B. González, “Reversible dye extraction from aqueous matrices using ammonium salt-based deep eutectic solvents,” *Separation and Purification Technology*, vol. 335, p. 126208, May 2024, doi: 10.1016/j.seppur.2023.126208.
[8] G. Srinivasan, P. Anaikutti, S. Mohan, and M. Arukkani, “Versatile Application of Calixarenes and Their Derivatives: From Drug Delivery to Industrial Catalysis and Environmental Remediation,” *Critical Reviews in Analytical Chemistry*, vol. 0, no. 0, pp. 1–58, Aug. 2025, doi: 10.1080/10408347.2025.2538731.
[9] G. Yu and X. Chen, “Host-Guest Chemistry in Supramolecular Theranostics,” *Theranostics*, vol. 9, no. 11, pp. 3041–3074, May 2019, doi: 10.7150/thno.31653.
[10] O. Gungor, A. Yilmaz, S. Memon, and M. Yilmaz, “Evaluation of the performance of calix[8]arene derivatives as liquid phase extraction material for the removal of azo dyes,” *Journal of Hazardous Materials*, vol. 158, no. 1, pp. 202–207, Oct. 2008, doi: 10.1016/j.jhazmat.2008.01.060.
[11] E. Akceylan, M. Bahadir, and M. Yılmaz, “Removal efficiency of a calix[4]arene-based polymer for water-soluble carcinogenic direct azo dyes and aromatic amines,” *Journal of Hazardous Materials*, vol. 162, no. 2–3, pp. 960–966, Mar. 2009, doi: 10.1016/j.jhazmat.2008.05.127.
[12] M. A. Kamboh, E. Akoz, S. Memon, and M. Yilmaz, “Synthesis of Amino-Substituted p-tert-Butylcalix[4]arene for the Removal of Chicago Sky Blue and Tropaeolin 000 Azo Dyes from Aqueous Environment,” *Water Air Soil Pollut*, vol. 224, no. 2, p. 1424, Jan. 2013, doi: 10.1007/s11270-012-1424-x.
[13] A. A. Bhatti, M. Oguz, and M. Yilmaz, “Magnetizing Calixarene: Azo Dye Removal from Aqueous Media by Fe3O4 Nanoparticles Fabricated with Carboxylic-Substituted Calix[4]arene,” no. 62, pp. 2819–2825, Jun. 2017.
[14] Ö. Güngör, “Efficient removal of carcinogenic azo dyes by novel pyrazine-2-carboxylate substituted calix[4, 8]arene derivatives,” *Supramolecular Chemistry*, vol. 31, no. 12, pp. 776–783, Nov. 2019, doi: 10.1080/10610278.2020.1711908.
[15] V. Böhmer *et al.*, “Synthesis and X-ray molecular structures of p-tert-butylcalix[4]arenes with diamide bridges spanning the 1,3-(distal) positions on the lower rim,” *J. Chem. Soc., Perkin Trans. 1*, no. 13, pp. 1521–1527, 1993, doi: 10.1039/P19930001521.
[16] A. Uyanik, N. Sen, and M. Yilmaz, “Improvement of catalytic activity of lipase from Candida rugosa via sol–gel encapsulation in the presence of calix(aza)crown,” *Bioresource Technology*, vol. 102, no. 6, pp. 4313–4318, Mar. 2011, doi: 10.1016/j.biortech.2010.12.105.
[17] E. Yilmaz Ozmen, S. Erdemir, M. Yilmaz, and M. Bahadir, “Removal of Carcinogenic Direct Azo Dyes from Aqueous Solutions Using Calix[n]arene Derivatives,” *CLEAN – Soil, Air, Water*, vol. 35, no. 6, pp. 612–616, 2007, doi: 10.1002/clen.200700033.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Land Surface Temperature as a Landslide Conditioning Factor: Mechanisms, Applications, and Implications for Susceptibility Assessment

Seda Çellek*

[1]*Department of Geological Engineering Kırşehir Ahi Evran University, Turkey*

**(sedacellek@ahievran.edu.tr) Email of the corresponding author*

Abstract **–** Land Surface Temperature (LST) is not merely a climatic indicator but an environmental parameter associated with surface energy balance, soil moisture, evapotranspiration, vegetation dynamics, and the thermo-hydro-mechanical behavior of soils. This study reviews the role of LST in landslide susceptibility assessment by synthesizing findings from theoretical and applied studies in the literature. The results indicate that the relationship between LST and landslide occurrence is neither unidirectional nor linear. In areas characterized by low surface temperatures, dense vegetation, and high moisture conditions, increased soil moisture and pore-water pressure may reduce slope stability. Conversely, elevated temperatures may contribute to thermal weakening, particularly in clay-rich soils, by affecting permeability, compressibility, creep behavior, and shear strength. Long-term increases in LST, when accompanied by declining Normalized Difference Vegetation Index (NDVI), land degradation, deforestation, and intense rainfall, may also provide valuable information for evaluating climate-related changes in landslide susceptibility. The reviewed studies demonstrate that LST has been used as an auxiliary or conditioning factor in susceptibility models, a climatic variable interacting with rainfall, an input to dynamic models, and an indicator for post-landslide environmental monitoring. Thermal remote sensing further enables the identification of moisture-related variations, surface-cover changes, and thermal anomalies in landslide-affected areas. However, LST alone has limited explanatory power and should be evaluated in combination with lithology, clay mineralogy, rainfall, soil moisture, vegetation, topography, and hydrological conditions. Overall, LST represents a promising complementary parameter for landslide susceptibility assessment, but its influence is strongly site-dependent and requires integrated evaluation with other environmental and hydro-mechanical factors.

Keywords – Land Surface Temperature, LST, landslide susceptibility, soil moisture, thermal remote sensing, climate change

I. INTRODUCTION

Land Surface Temperature (LST) is not only an indicator of climatic conditions but also a physically meaningful environmental parameter representing thermo-hydro-mechanical processes that can be incorporated into large-scale landslide susceptibility models [1]. However, rather than being considered a direct landslide-triggering factor, LST has been regarded as an auxiliary parameter that provides indirect information on environmental conditions associated with landslide processes, such as surface energy balance, vegetation, and soil moisture [2]. Similarly, LST has also been considered a spatial indicator of climate change and environmental degradation [3].

The relationship between LST and landslide occurrence can be explained through the effects of surface temperature on soil moisture and slope stability. The literature suggests an indirect relationship between low surface temperature, high soil moisture, and reduced slope stability, indicating that these conditions may increase the probability of landslide occurrence [2, 3]. Conversely, landslide activity has been shown

to persist for longer periods on slopes with warmer surfaces [4]. This difference indicates that the influence of temperature on landslide processes is not limited to its effects on evapotranspiration. Temperature can alter hydro-mechanical soil properties, including permeability, water-retention capacity, compressibility, consolidation, pore-water pressure, and shear strength, thereby affecting slope stability [1, 5]. Therefore, the influence of LST may vary depending on the soil and hydrological characteristics of the study area.

Changes in thermal properties are also important for the identification and monitoring of landslide-affected areas. Following a landslide, the removal of vegetation and exposure of bare ground can alter the surface energy balance and consequently lead to changes in LST. Rose et al. [6] showed that LST values within the landslide-affected area were higher than those in the surrounding areas. Temperature variations were relatively limited before the landslide but became considerably more pronounced after the event. This pattern was attributed to the greater absorption of solar radiation by exposed bare surfaces. In the same study, ecological recovery following the 2019 Puthumala landslide in Kerala, India, was assessed using changes in the Normalized Difference Vegetation Index (NDVI) and LST during the pre-landslide (2017–2019) and post-landslide (2020–2024) periods [6].

Thermal remote sensing products can provide information not only on surface temperature but also indirectly on soil moisture conditions. Therefore, LST can be used as a proxy for soil moisture, contributing to the assessment of changes in evapotranspiration and rainfall conditions capable of triggering landslides [5]. Nevertheless, LST should not be evaluated independently but rather in combination with lithology, clay mineralogy, moisture conditions, rainfall, and other hydrological processes. Putro et al. [2] evaluated slope stability by integrating LST derived from Landsat 8 imagery with the Soil Moisture Index (SMI), demonstrating the potential of LST as an auxiliary environmental parameter.

The application of LST in landslide research has expanded in recent years through different methodological approaches. Loche et al. [4] conducted one of the first studies to directly incorporate LST into a landslide susceptibility model to explain the temporal and spatial evolution of landslides following the 2008 Wenchuan earthquake. Karmakar et al. [7] investigated the landslides that occurred on 7 September 2007 in the Himalayan region of India by jointly evaluating daily rainfall and maximum (Tmax) and minimum (Tmin) surface temperatures. Aziz et al. [3] developed two different models based on the Information Value method to evaluate the influence of LST on landslide susceptibility mapping in the Ramban District of Jammu and Kashmir, India.

The assessment of long-term LST variations is also important for understanding the role of temperature in the relationship between climate change and landslide processes. Islam et al. [8] investigated changes in LST and NDVI between 2001 and 2022 together with landslide intensity in the Chattogram Hill Tracts region of Bangladesh. Based on a 22-year time series, the study is particularly important because it evaluated LST not only in relation to a single landslide event or susceptibility model but also in the context of long-term environmental changes and landslide intensity. The findings indicate that LST can be used in landslide research not only as a conditioning factor but also as an indicator of spatial and temporal environmental changes associated with climate change [8].

Nevertheless, temperature has received relatively limited attention in landslide susceptibility studies compared with morphological, geological, and hydrological factors. Scaringi and Loche [5] comprehensively evaluated the effects of temperature on soil behavior and highlighted that, despite the use of hundreds of different input variables across numerous landslide susceptibility models, temperature has been only rarely represented in modeling studies. This indicates that LST remains an insufficiently investigated environmental variable in landslide susceptibility research.

Within this framework, LST can be considered a proxy for soil moisture and hydro-mechanical behavior, an auxiliary or conditioning factor in landslide susceptibility models, a spatial indicator of climate change, and a remote sensing parameter for monitoring post-landslide environmental changes. However, the fact that the relationship between LST and landslides does not exhibit the same direction under different environmental and soil conditions highlights the need for an integrated assessment of the findings reported in the literature. Accordingly, this study aims to evaluate the effects of land surface temperature on landslide occurrence and susceptibility and to review the applications of LST in landslide research, its physical mechanisms of influence, and its interactions with other environmental parameters.

II. Effects of LST on Landslides

The influence of LST on landslides does not occur independently but rather in combination with soil moisture, vegetation, rainfall, lithology, and the hydro-mechanical properties of soils. The literature indicates that both low and high surface temperatures can influence landslide susceptibility through different mechanisms.

- Low LST → high soil moisture → increased pore-water pressure → reduced slope stability → increased landslide susceptibility [2, 3].
- In areas characterized by low surface temperatures and dense vegetation, greater moisture retention may increase soil water content and consequently reduce slope stability [2].
- High LST → thermal weakening → increased deformation and creep → reduced shear strength. This mechanism becomes particularly important in clay-rich soils [1, 5].
- Increasing temperature can affect the hydro-mechanical behavior of soils by altering pore-water viscosity, permeability, consolidation, water-retention capacity, and pore-water movement [5].
- Increasing LST + decreasing NDVI + land degradation + deforestation + intense rainfall may collectively create environmental conditions that increase landslide susceptibility over the long term.

Spatial Variations in LST

- High LST: Commonly observed in bare-ground areas, settlements, valley floors, and areas with sparse vegetation [2].
- Low LST: Commonly observed on shaded and humid slopes with dense vegetation [2].
- Cool and humid slopes may exhibit increased landslide susceptibility due to higher soil moisture conditions.
- The hydro-mechanical effects of temperature may become more pronounced in clay-rich soils [1].

LST as an Environmental Indicator

LST provides indirect information on:

- surface energy balance,
- soil moisture,
- evapotranspiration,
- hydrological processes, and
- vegetation dynamics.

In addition, the following general relationships have been observed:

- Decreasing vegetation cover → increasing LST
- Increasing bare-ground exposure → increasing LST
- Vegetation recovery → decreasing LST

Overall effect: The relationship between LST and landslides is not unidirectional. Low LST may increase landslide susceptibility through high-moisture conditions, whereas high LST may contribute to susceptibility through thermal and hydro-mechanical weakening, particularly in clay-rich soils. Therefore, the influence of LST should be evaluated in combination with other environmental and soil-related parameters.

III. Identification and Monitoring of Landslides Using LST

Thermal remote sensing data can be used to identify landslide-affected areas and monitor post-landslide changes.

- Disturbed and active landslide masses may exhibit different thermal characteristics from the surrounding stable areas [5].
- Differences in porosity, water content, and thermal conductivity within landslide masses may result in surface temperature anomalies.
- These thermal anomalies can help distinguish landslide-affected areas from the surrounding stable terrain and facilitate their identification using thermal remote sensing data [5].

- The combined use of LST and NDVI can support the monitoring of vegetation recovery and ecological restoration following landslides [6].
- The post-landslide decrease in NDVI accompanied by an increase in LST is associated with the disturbance of surface cover and exposure of bare ground.
- Rose et al. [6] considered LST not as an independent landslide susceptibility factor but as a remote sensing indicator for post-landslide monitoring and the assessment of terrain recovery.

Overall: LST can be used not only for landslide susceptibility assessment but also for the identification of active or disturbed landslide areas and the monitoring of post-landslide environmental changes.

IV. Surface Temperature Values

In the literature, LST has been evaluated in terms of specific temperature thresholds, pre- and post-landslide variations, and long-term temperature trends. The findings indicate that spatial and temporal variations in temperature should be considered rather than relying on a single LST value.

Rose et al. [6]; LST variations during the 2017–2024 period were examined, revealing a marked increase in surface temperature following the landslide:

- 2017: 28.97 °C
- 2018: 29.34 °C
- 2019: 32.99 °C
- 2020: 34.30 °C
- 2021: 34.05 °C
- 2022: 34.50 °C — highest value
- 2023: 31.19 °C
- 2024: 30.43 °C

The increase in LST following the landslide and the subsequent declining trend indicate that LST can serve as an environmental indicator for monitoring the temporal evolution and recovery of landslide-affected areas [6].

Alparslan et al. [9]; Surface temperatures in the study area ranged from 7 to 38 °C, with the 24–26 °C range identified as a critical temperature interval for landslide occurrence.

- <24 °C: 72.87% of the study area
- 24–26 °C: 24.34% of the study area — critical range
- >26 °C: 2.79% of the study area

These findings demonstrate that LST can be related to landslide susceptibility not only through generally high or low temperature conditions but also through site-specific temperature ranges and threshold values [9].

Islam et al. [8]; evaluated LST and NDVI variations over a 22-year period (2001–2022).

- An increasing LST trend was identified in approximately 83–94% of the grid cells.
- The mean LST increase was approximately 0.34 °C/year.
- The most pronounced increase in LST was observed in Khagrachari.
- A decreasing NDVI trend was identified in approximately 53% of the grid cells.
- Vegetation loss was particularly pronounced in areas with steep slopes.
- Khagrachari and Rangamati were identified as the areas most susceptible to vegetation loss.

These findings demonstrate that long-term increases in LST may occur simultaneously with decreases in NDVI and that surface temperature variations should therefore be evaluated in relation to vegetation dynamics, land-use changes, and landslide intensity [8].

Overall evidence from the literature: There is no single critical LST value that can be universally applied to all study areas. The relationship between LST and landslides varies according to regional climate, land cover, soil moisture, lithology, and soil properties. Therefore, rather than relying solely on absolute temperature values, site-specific thresholds and the spatial and temporal variability of LST should be considered.

V. CALCULATION OF LST

Landsat thermal bands and MODIS temperature products are widely used for the determination of LST. When LST is derived from Landsat imagery, the digital numbers (DN) of the thermal bands are first converted into spectral radiance (Lλ), after which Brightness Temperature is calculated from the resulting radiance values [9, 10].

The main parameters used for temperature calculation from Landsat thermal data include:

- Lλ: Spectral radiance at the sensor aperture,
- G: Gain coefficient,
- B: Bias coefficient,
- K1: First thermal calibration constant,
- K2: Second thermal calibration constant,
- T: Temperature in Kelvin,

as described in [9, 10].

Main Data Sources Used in the Literature

- Landsat thermal bands: Calculation of LST and brightness temperature,
- MODIS LST (MOD11A2): Assessment of temporal variations in LST,
- MODIS NDVI (MOD13C2): Assessment of temporal variations in vegetation,
- Google Earth Engine (GEE): Processing of long-term satellite imagery and LST analyses,
- Multi-temporal satellite imagery: Assessment of temporal variations in LST,
- Field observations and official landslide inventories: Comparison and validation of remote sensing results with landslide distribution.

In long-term studies, LST and NDVI data can be evaluated together to investigate the surface temperature–vegetation–landslide relationship. In some studies, these datasets have been analyzed at a spatial resolution of 1 km.

Temperature Variables Used

Surface temperature can be evaluated as a single value or through variables representing daily temperature variations, including:

- Tmax: Daily maximum surface temperature,
- Tmin: Daily minimum surface temperature.

Analytical Methods

Several statistical methods have been used to determine temporal variations in LST and its relationships with other environmental parameters, including:

- Linear regression,
- Ordinary Least Squares (OLS),
- Mann–Kendall trend test,
- Pearson correlation analysis.

These methods are used to identify long-term trends in LST and NDVI and subsequently compare these trends with landslide intensity and spatial distribution.

In landslide susceptibility studies, the resulting susceptibility levels can be classified depending on the methodology applied and are generally expressed as:

- Low,
- Moderate,
- High.

VI. DISCUSSION

The literature indicates that the influence of LST on landslide susceptibility is neither independent nor linear, nor does it operate in the same direction across all study areas. The significance of LST varies depending on its interactions with rainfall, soil moisture, vegetation, lithology, clay mineralogy, and other hydrological conditions. The main differences among the reviewed studies can be summarized as follows:

- Putro et al. [2] found that the influence of LST on landslides was relatively weak and considered it a secondary variable that indirectly represents slope stability. They emphasized that LST should

be used in combination with geology, topography, soil properties, slope, and hydrological conditions and that remote sensing results should be validated using field data.

- Karmakar et al. [7] evaluated LST not as an independent susceptibility parameter but in interaction with rainfall. Their findings indicate that the combined assessment of temperature and rainfall provides a more meaningful explanation of landslide occurrence.
- Scaringi and Loche [5] stated that a single mean temperature value is insufficient to explain soil and landslide behavior and emphasized the importance of considering thermal variables such as seasonal temperature variations, daily temperature fluctuations, and temperature differences. They further highlighted that thermal parameters should be selected according to landslide and soil type.
- Loche et al. [4] demonstrated that the influence of LST may vary over time. In 2008, when the Wenchuan earthquake occurred, LST had no pronounced influence on landslide activity. However, as the influence of the earthquake diminished over time, a positive relationship between LST and landslide activity emerged. Particularly in 2013 and 2015, landslide activity persisted for longer periods on slopes characterized by warmer surfaces. Moreover, the better performance of the LST-based model compared with the NDVI-based model demonstrates the potential contribution of LST to landslide susceptibility modeling.
- Loche and Scaringi [1] showed that temperature may become a more significant conditioning factor, particularly in clay-rich soils, through hydro-mechanical and thermal processes. This finding indicates that the influence of LST may vary depending on soil properties.
- Aziz et al. [3] developed a dynamic model by incorporating LST into a static model based on conventional topographic and environmental parameters. The inclusion of LST improved model performance and revealed a relationship between low surface temperature, high moisture conditions, and increased landslide susceptibility.
- Islam et al. [8] demonstrated that long-term increases in LST should not be considered independently. The combined effects of increasing LST, decreasing NDVI, land degradation, deforestation, and intense rainfall provide a more meaningful explanation of changes in landslide susceptibility.
- Unlike the other studies, Rose et al. [6] primarily demonstrated the potential of LST for post-landslide monitoring rather than its role in landslide occurrence. The combined use of LST and NDVI enables the monitoring of vegetation recovery, changes in surface conditions, and ecological restoration.

When these studies are evaluated collectively, the reported associations of "low LST–high landslide susceptibility" and "high LST–high landslide activity" should not necessarily be interpreted as contradictory. Low LST values may be associated with reduced slope stability through high soil moisture and increased pore-water pressure in cool, humid, and densely vegetated areas. Conversely, high temperatures may influence landslide behavior through thermal weakening and changes in hydro-mechanical properties, particularly in clay-rich soils. Therefore, both the direction and magnitude of the LST–landslide relationship depend on climate, soil properties, vegetation, moisture conditions, rainfall, and temporal scale.

Accordingly, rather than treating LST as an independent and universally applicable landslide conditioning factor, it may be more appropriate to consider it as a dynamic environmental variable whose influence emerges through interactions with other environmental and hydro-mechanical parameters.

VII. CONCLUSION

Based on the collective evaluation of the literature, the main conclusions regarding the role of LST in landslide studies can be summarized as follows:

- Rather than being an independent factor that explains landslide occurrence on its own, LST is an environmental parameter that interacts with soil moisture, rainfall, vegetation, lithology, and soil properties.
- The incorporation of LST into conventional landslide conditioning factors has improved the performance of landslide susceptibility models in some studies. Dynamic models incorporating LST have been shown to outperform static models [3].

- Low LST values may be associated with reduced slope stability and increased landslide susceptibility through higher soil moisture and increased pore-water pressure [2, 3].
- Conversely, high LST values may increase landslide activity, particularly in clay-rich soils, through thermal weakening and changes in hydro-mechanical properties [1, 5].
  Evaluating LST in combination with rainfall may provide a better explanation of landslide occurrence than considering temperature alone [7].
- Long-term increases in LST, when evaluated together with decreasing NDVI, land degradation, deforestation, and intense rainfall, provide important information for explaining changes in landslide intensity [8].
- Following landslides, increasing LST and decreasing NDVI may indicate surface-cover disturbance, whereas decreasing LST associated with vegetation recovery can serve as a useful indicator for monitoring terrain recovery [6].
- The direction and magnitude of the LST–landslide relationship vary depending on the climate, soil properties, vegetation, moisture conditions, rainfall regime, and temporal scale of the study area. Therefore, defining a single LST threshold or a universal linear relationship applicable to all study areas is not appropriate.

Overall, Land Surface Temperature (LST) is a relatively underutilized parameter in landslide susceptibility studies, despite its ability to provide valuable physical and environmental information. LST can serve multiple roles, including as an auxiliary or conditioning factor in landslide susceptibility models, a climatic variable, a proxy for soil moisture, and an indicator for post-landslide monitoring. Future studies should therefore evaluate LST in combination with rainfall, soil moisture, NDVI, lithology, clay mineralogy, and hydrological parameters, rather than considering it independently, and should further investigate its effects under different climatic and soil conditions.

REFERENCES

[1] Loche, M., & Scaringi, G. (2024). Assessing the influence of temperature on slope stability in a temperate climate: A nationwide spatial probability analysis in Italy. EarthArXiv. [Preprint].

[2] Putro, S. T. J., Arif, N., & Sarastika, T. (2022). Land surface temperature (LST) and soil moisture index (SMI) to identify slope stability. IOP Conference Series: Earth and Environmental Science, 986(1), 012022. https://doi.org/10.1088/1755-1315/986/1/012022

[3] Aziz, K., Singh, H., Sarkar, S., Sahu, P., & Rawat, D. (2022). Study of land surface temperature on landslide susceptibility zonation in Ramban District of Jammu and Kashmir, India. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2022) (pp. 453–456). IEEE. https://doi.org/10.1109/IGARSS46834.2022.9884515

[4] Loche, M., Scaringi, G., Yunus, A. P., Catani, F., Tanyaş, H., Frodella, W., Fan, X., & Lombardo, L. (2022). Surface temperature controls the pattern of post-earthquake landslide activity. Scientific Reports, 12, 988. https://doi.org/10.1038/s41598-022-04992-8.

[5] Scaringi, G., & Loche, M. (2022). A thermo-hydro-mechanical approach to soil slope stability under climate change. Geomorphology, 401, 108108. https://doi.org/10.1016/j.geomorph.2022.108108

[6] Rose, S., Pradeep, G. S., & Vijith, H. (2026). An investigation of terrain recovery by analyzing the Puthumala landslide-impacted region in Kerala, India, utilizing both pre and post disaster NDVI and land surface temperature data. Discover Hazards, 2, 10. https://doi.org/10.1007/s44475-026-00016-5.

[7] Karmakar, S., Tiwari, P. N., & Basak, T. (2023). A combined influence of surface temperature and daily rainfall to the historical landslides occurred on 7th September 2007 over sub-Himalayan region, India. Journal of Earth System Science, 132, 42. https://doi.org/10.1007/s12040-023-02054-9

[8] Islam, S., Adil, M., Ganguli, S., & Akter, N. (2023). Spatiotemporal variations of land surface temperature and vegetation dynamics impacts on landslide intensity in Chattogram Hill Tracts Region of Bangladesh. Grassroots Journal of Natural Resources, 6(2), 19–37. https://doi.org/10.33002/nr2581.6853.060203

[9] *Chander, G., & Markham, B. (2003). Revised Landsat-5 TM radiometric calibration procedures and postcalibration dynamic ranges. IEEE Transactions on Geoscience and Remote Sensing, 41(11), 2674–2677. https://doi.org/10.1109/TGRS.2003.818464*

[10] Alparslan E. vd. 2006, "Büyük ve Küçük Çekmece gölleri arasındaki bölgede heyelan duyarlılığının uzaktan algılama ve coğrafi bilgi sistemleri ile çok kıstaslı analizi", 4. Coğrafya Bilgi Sistemleri Bilişim Günleri, 13–16 Eylül 2006, Fatih Üniversitesi, İstanbul

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Solar Radiation As A Landslide Conditioning Factor: A Literature Review

Seda Çellek*

[1]*Department of Geological Engineering Kırşehir Ahi Evran University, Turkey*

**(sedacellek@ahievran.edu.tr) Email of the corresponding author*

Abstract – Solar radiation is a secondary topographic parameter representing the spatial distribution of solar energy reaching the Earth's surface. It is primarily controlled by slope, aspect, latitude, solar position, and topographic shading. Solar radiation can indirectly influence landslide susceptibility through its effects on evapotranspiration, land surface temperature, soil moisture, vegetation growth, weathering, and freeze–thaw processes. However, despite its potential importance, solar radiation is less frequently employed in landslide susceptibility studies than slope and aspect because of its computational complexity, inconsistent units and classification schemes adopted across studies, and its climate-dependent influence.

This review comprehensively examines solar radiation parameters reported in the literature under various names, including Area Solar Radiation (ASR), Annual Solar Radiation, Potential Solar Radiation, Incoming Solar Radiation, and Insolation, by integrating their conceptual definitions and methodological characteristics into a unified framework. The reviewed studies indicate that areas receiving lower solar radiation generally experience reduced evapotranspiration, allowing soil moisture to be retained for longer periods, which may increase landslide susceptibility in certain environments. Nevertheless, the relationship between solar radiation and landslide occurrence is not universally linear because solar radiation also affects vegetation cover, desiccation cracking, weathering intensity, and soil properties. Consequently, the contribution of solar radiation to landslide susceptibility models should be interpreted in conjunction with climatic conditions, latitude, topographic resolution, and other environmental factors.

Keywords – Landslide susceptibility, conditioning factor, Area Solar Radiation, Annual Solar Radiation

I. INTRODUCTION

Annual solar radiation has been defined as the average amount of solar radiation received over one year by a specific pixel [1], by each pixel [2, 3], or by a pixel over an annual period [4]. Although these definitions are similar, they do not necessarily refer to exactly the same computational approach. Area Solar Radiation (ASR) represents the spatial estimation of direct and diffuse solar energy reaching a topographic surface over a specified time interval, whereas Annual Solar Radiation refers specifically to applications in which the calculation period is one year. Potential Solar Radiation, in contrast, generally represents the theoretical or potential amount of incoming solar energy under idealized atmospheric conditions, where factors such as cloud cover are minimized or excluded.

This parameter is commonly calculated using the Area Solar Radiation tool in ArcGIS and is expressed in kWh m^{-2} (kilowatt-hours per square meter). Solar radiation is regarded as an indirect conditioning factor influencing landslide susceptibility [4]. It is one of the secondary terrain derivatives calculated primarily from slope and aspect [5]. The underlying reason is that solar radiation directly controls evapotranspiration while simultaneously being strongly influenced by topographic conditions [4]. Krassakis et al. [6] incorporated Solar Radiation as one of the conditioning factors in their Rock Engineering System (RES)-based landslide susceptibility analysis. They constructed an interaction matrix among conditioning factors and subsequently calculated their weighting coefficients.

Ali et al. [1] reported a normalized ReliefF feature importance weight of 0.0670 for annual solar radiation. Among the sixteen conditioning factors considered, annual solar radiation ranked approximately seventh in importance. The ranking of the most influential variables was as follows: slope (0.1516), lithology (0.1182), aspect (0.0966), land cover (0.0900), distance to faults (0.0874), Topographic Wetness Index (TWI) (0.0797), and annual solar radiation (0.0670).

Gorsevski et al. [7] identified Annual Solar Radiation (ASR) as one of the principal topographic variables in their landslide susceptibility model. It was incorporated as an input variable to an Artificial Neural Network (ANN) model together with slope, profile curvature, plan curvature, upslope drainage area, and the Topographic Wetness Index (TWI). All of these variables were derived from a LiDAR-based Digital Elevation Model (DEM).

Ghasemian et al. [2] employed the Information Gain Ratio (IGR) technique for feature selection and retained only the sixteen most informative variables out of twenty-five conditioning factors. However, solar radiation was not identified among the most influential predictors. Instead, distance to roads, road density, lithology, and land use were reported as the dominant variables. Therefore, although solar radiation was included in the modelling process, it was not highlighted as a leading conditioning factor.

Similarly, Nhu et al. [4] used the IGR algorithm to select twelve conditioning factors for model development, and Solar Radiation was retained among the selected variables. In contrast, NDVI, lithology, fault density, road density, stream density, distance to roads, distance to faults, and distance to streams were excluded because they contributed little predictive information. Solar radiation, however, provided sufficient information gain and was therefore retained in both the Random Forest and ensemble models.

In another study, Nhu et al. [8] developed landslide susceptibility models using twenty conditioning factors, including slope, aspect, elevation, curvature, plan curvature, profile curvature, slope length, rainfall, annual solar radiation, Stream Power Index (SPI), Topographic Wetness Index (TWI), distance to rivers, river density, lithology, distance to faults, fault density, land use, NDVI, distance to roads, and road density.

Pawluszek and Borkowski [5] evaluated ten DEM-derived topographic parameters, namely elevation, slope, morphological gradient, aspect, Area Solar Radiation (ASR), roughness, Topographic Position Index (TPI), Topographic Wetness Index (TWI), Stream Power Index (SPI), and shaded relief.

Shirzadi et al. [3] employed twenty landslide conditioning factors, among which Annual Solar Radiation was included within the hydrological parameter group together with rainfall, Stream Power Index (SPI), Topographic Wetness Index (TWI), distance to rivers, and river density.

Likewise, Krassakis et al. [6] selected Annual Solar Radiation as one of the ten principal conditioning factors for landslide susceptibility assessment. The parameter was evaluated together with lithology, elevation, slope, curvature, distance to tectonic structures, distance to the drainage network, distance to the road network, seismic parameters, and annual precipitation.

II. Calculatİon and Applİcatİon of the Parameter

All landslide conditioning factors are initially derived from Digital Elevation Models (DEMs) of varying spatial resolutions and subsequently resampled to a common raster resolution for susceptibility analysis [6]. Area Solar Radiation (ASR) is one of the secondary terrain derivatives and is calculated primarily from slope and aspect. In ArcGIS, the parameter is generated using the Area Solar Radiation tool based on slope and aspect datasets [3].

ASR integrates two principal components of solar exposure: the solar incidence angle, which is slope-dependent, and solar orientation, which is aspect-dependent [5]. The parameter is commonly calculated using the Area Solar Radiation tool available in ArcGIS. Gorsevski et al. [7] reported that annual solar radiation was derived from a LiDAR-based Digital Elevation Model (DEM) and incorporated as one of the fundamental terrain variables in topographic analyses.

Tekin [9] generated annual solar radiation maps using a Solar Radiation extension developed by Parks [10] for ArcView GIS 3.2. The extension implements regression equations proposed by McCune and Keon [11] to estimate annual solar radiation. By applying these regression equations to the study area, Tekin [9] calculated annual solar radiation using three different approaches. The calculations considered latitude, slope, and aspect as the primary input variables.

The three regression equations recommended for estimating annual solar radiation differ in their applicability and limitations. The first equation was designed for global application regardless of geographic location. However, Parks [10] suggested that this equation provides relatively low accuracy and therefore did not recommend its use. The second equation can be applied at all latitudes but is suitable only for areas where slope angles range between 0° and 60°. Parks [10] reported that, under these conditions, the second equation produces more reliable results than the first. The third equation requires more restrictive assumptions, namely that the study area is located between 30° and 60° latitude and that terrain slopes do not exceed 60°. When these conditions are satisfied, the third equation yields the highest prediction accuracy and was therefore recommended by Parks [10].

Solar radiation has been incorporated as an input variable in a wide range of machine learning and hybrid modelling approaches for landslide susceptibility assessment. Shirzadi et al. [3] included solar radiation as one of the input variables in several machine learning models, including ADTree, MB-ADTree, BA-ADTree, RS-ADTree, and RF-ADTree. Likewise, Ali et al. [1] employed annual solar radiation as an input parameter in Random Forest (RF), Naïve Bayes (NB), and FDEMATEL-ANP models. The authors generated the annual solar radiation layer from a LiDAR-derived DEM within a GIS environment and also incorporated it as one of the input variables in an Artificial Neural Network (ANN) model.

Similarly, Ghasemian et al. [2] used Solar Radiation as one of the predictor variables in their hybrid deep learning framework (DBN–ELM–BP–GA) and compared its performance with Support Vector Machine (SVM), REPTree, and NBTree models. Following the generation of the conditioning factors, continuous solar radiation values are commonly classified into discrete categories using the Natural Breaks (Jenks) classification method before being incorporated into landslide susceptibility models [4].

III. Classification of the Parameter

In landslide susceptibility studies, solar radiation is reported under a variety of terms. The most frequently used expressions in the literature include:

- Area Solar Radiation (ASR)
- Annual Solar Radiation
- Incoming Solar Radiation
- Potential Solar Radiation
- Solar Insolation
- Terrain Solar Radiation
- Insolation

Area Solar Radiation (ASR) refers to the spatial distribution of solar energy reaching a topographic surface over a specified time interval. During the calculation process, each raster cell is assigned a solar radiation value according to its topographic position and orientation relative to the sun. Consequently, ASR is regarded as a secondary terrain variable derived from a Digital Elevation Model (DEM).

ASR fundamentally integrates two topographic components: the solar incidence angle, which depends on slope, and the orientation effect, which depends on aspect. However, modern solar radiation models incorporate considerably more variables than slope and aspect alone. These include latitude, the position of the sun in the sky, day length, topographic shading, direct solar radiation, diffuse solar radiation, and atmospheric transmissivity.

The total amount of solar radiation reaching a topographic surface generally consists of three components:

Direct radiation is the solar energy reaching the Earth's surface without atmospheric scattering. It is strongly controlled by slope, aspect, solar elevation angle, and solar azimuth.

Diffuse radiation is the solar energy scattered by atmospheric gases, aerosols, and clouds before reaching the surface. Consequently, even shaded areas receive a certain amount of diffuse radiation.

Reflected radiation represents the energy reflected from surrounding surfaces toward the target topographic surface. This component is influenced by surface albedo, snow cover, rock characteristics, and soil properties.

In most ASR calculations, total solar radiation is represented as the sum of direct and diffuse radiation, whereas some models additionally incorporate reflected radiation.

After calculation, continuous solar radiation values are generally classified into discrete categories using the Natural Breaks (Jenks) classification method. Nhu et al. [4] classified solar radiation into seven classes (3.015–6.563, 6.563–6.747, 6.747–6.849, 6.849–6.930, 6.930–7.073, 7.073–7.236, and 7.236–8.215). In another study, Nhu et al. [8] evaluated the predictive importance of Solar Radiation using the Information Gain Ratio (IGR) algorithm and reported an Average Merit (AM) value of 55.618. This ranking placed Solar Radiation ninth among the twenty conditioning factors, indicating that it was one of the important variables selected for landslide susceptibility modelling. Solar Radiation ranked below slope, Topographic Wetness Index (TWI), plan curvature, LS factor, curvature, land use, Stream Power Index (SPI), and profile curvature, but showed greater predictive importance than elevation, aspect, and rainfall [8].

Ali et al. [1] divided the annual solar radiation layer into five classes, with radiation values ranging from approximately 817,967 to 1,319,930 Wh m^{-2}. Similarly, Gorsevski et al. [7] classified the parameter into five intervals: 80,000–430,000, 440,000–540,000, 550,000–630,000, 640,000–700,000, and 710,000–810,000 Wh m^{-2}.

Krassakis et al. (2023) also categorized Annual Solar Radiation into five classes: >50,094.64, 42,347.82–50,094.64, 35,350.69–42,347.82, 26,104.48–35,350.69, and <26,104.48 Wh m^{-2}. In contrast, Shirzadi et al. [3] classified Solar Radiation into seven classes, whereas Ghasemian et al. (2022) adopted five classes. Regarding measurement units, Krassakis et al. [6] expressed solar radiation in Wh m^{-2}, while Ghasemian et al. [2] reported values in kWh m^{-2}.

Using the Rock Engineering System (RES) approach, Krassakis et al. [6] calculated a weighting coefficient (Wi) of 0.264 for Annual Solar Radiation. Although this value was lower than those obtained for lithology (0.595), slope (0.546), and proximity to tectonic structures (0.479), Annual Solar Radiation was still identified as an important conditioning factor contributing to landslide susceptibility modelling.

IV. Influence on Landslide Susceptibility

Most studies have reported an inverse relationship between solar radiation and landslide susceptibility, suggesting that the probability of landslide occurrence generally decreases as solar radiation increases. Consequently, slopes receiving higher annual solar radiation are considered less susceptible to landslides [6]. Similarly, lower solar radiation values have been associated with a greater likelihood of landslide occurrence [3], while Ali et al. [1] likewise concluded that areas receiving higher annual solar radiation generally exhibit lower landslide susceptibility.

This relationship is primarily attributed to the influence of solar radiation on soil moisture through evapotranspiration. Slopes exposed to high levels of solar radiation throughout the year tend to lose moisture more rapidly [6]. The resulting reduction in soil water content and pore water contributes to a drier subsurface environment, thereby increasing soil suction and reducing the likelihood of slope failure [1]. Krassakis et al. [6] further suggested that moisture loss creates additional void spaces between soil and rock particles as water is removed, which contributes to greater slope stability. Overall, increased solar radiation promotes drier ground conditions and consequently lowers landslide susceptibility [1].

Effects of low solar radiation on soil moisture

Pawluszek and Borkowski [5] reported that areas receiving lower amounts of solar radiation tend to maintain higher soil moisture, thereby increasing the likelihood of landslide occurrence. On shaded slopes where solar radiation is limited and evaporation is reduced, soil moisture can be retained for longer periods. This prolonged moisture retention may increase pore water pressure, reduce effective stress, and consequently decrease the shear strength of slope materials [1].

An increase in soil moisture may:

- increase the unit weight of the soil,
- increase pore water pressure,
- reduce effective stress, and
- decrease shear strength.

Collectively, these mechanisms may explain the higher landslide susceptibility frequently observed in low-radiation environments.

Snowmelt and freeze–thaw processes

In mountainous or cold-climate regions, incoming solar radiation strongly influences snowmelt timing and the frequency of freeze–thaw cycles. Snow cover generally persists longer on shaded slopes, whereas sun-exposed slopes experience earlier and more rapid snowmelt. Rapid snowmelt may introduce large volumes of water into the ground within a relatively short period, increasing soil saturation and pore water pressure. Therefore, solar radiation can indirectly influence landslide occurrence through snowmelt-related hydrological processes, particularly in alpine environments.

Vegetation development

Solar radiation also affects vegetation growth, although the relationship with landslide susceptibility is not straightforward. Pawluszek and Borkowski [5] suggested that greater solar radiation may promote vegetation development, resulting in more stable slopes. In contrast, Ghasemian et al. [2] reported that denser vegetation may also develop on shaded, low-radiation slopes under favorable moisture conditions, thereby enhancing slope stability. Consequently, the relationship between solar radiation, vegetation, and landslide susceptibility is highly dependent on local climatic and ecological conditions.

Weathering and desiccation processes

High solar radiation increases temperature fluctuations at the ground surface. Repeated heating and cooling cycles may generate thermal stresses that accelerate physical weathering of rock and soil materials. In clay-rich soils, alternating drying and wetting cycles can promote the formation of desiccation cracks. These cracks subsequently facilitate rapid infiltration of rainfall, increasing local pore water pressure and potentially reducing slope stability.

Evapotranspiration

Solar radiation represents the primary energy source driving evapotranspiration. Under favorable atmospheric and surface conditions, increased solar radiation enhances both soil evaporation and plant transpiration, thereby reducing soil moisture accumulation on hillslopes. Consequently, higher evapotranspiration rates generally contribute to improved slope stability by limiting prolonged saturation.

Topographic characteristics

Gorsevski et al. [7] observed that areas receiving higher solar radiation generally correspond to drier and more convex landforms, whereas low-radiation areas are commonly associated with steeper, more concave landslide scarps. This observation further supports the indirect influence of solar radiation on landslide susceptibility through its interaction with terrain morphology and surface hydrological processes.

Overall, the findings indicate that the influence of solar radiation on landslide susceptibility is predominantly indirect and is mediated through multiple interacting processes, including evapotranspiration, soil moisture dynamics, vegetation development, weathering, snowmelt, and terrain morphology. Therefore, the contribution of solar radiation should be evaluated together with climatic conditions, topographic characteristics, and local environmental factors rather than being interpreted as an independent controlling variable.

Figure 1 presents the conceptual relationship between solar radiation and landslide susceptibility through its influence on soil hydrological conditions. Areas receiving low solar radiation generally experience lower surface temperatures and reduced evapotranspiration, enabling soil moisture to persist for longer periods. The resulting increase in pore-water pressure reduces effective stress and soil shear strength, thereby increasing slope instability and landslide susceptibility. Conversely, areas exposed to higher solar radiation exhibit greater surface temperatures and evapotranspiration, leading to lower soil moisture and pore-water pressure. These conditions increase effective stress and shear strength, contributing to improved slope stability and reduced landslide susceptibility. This conceptual mechanism is consistent with the findings of Shirzadi et al. [3] and Ali et al. [1], who reported that solar radiation indirectly controls landslide occurrence by regulating soil moisture dynamics.

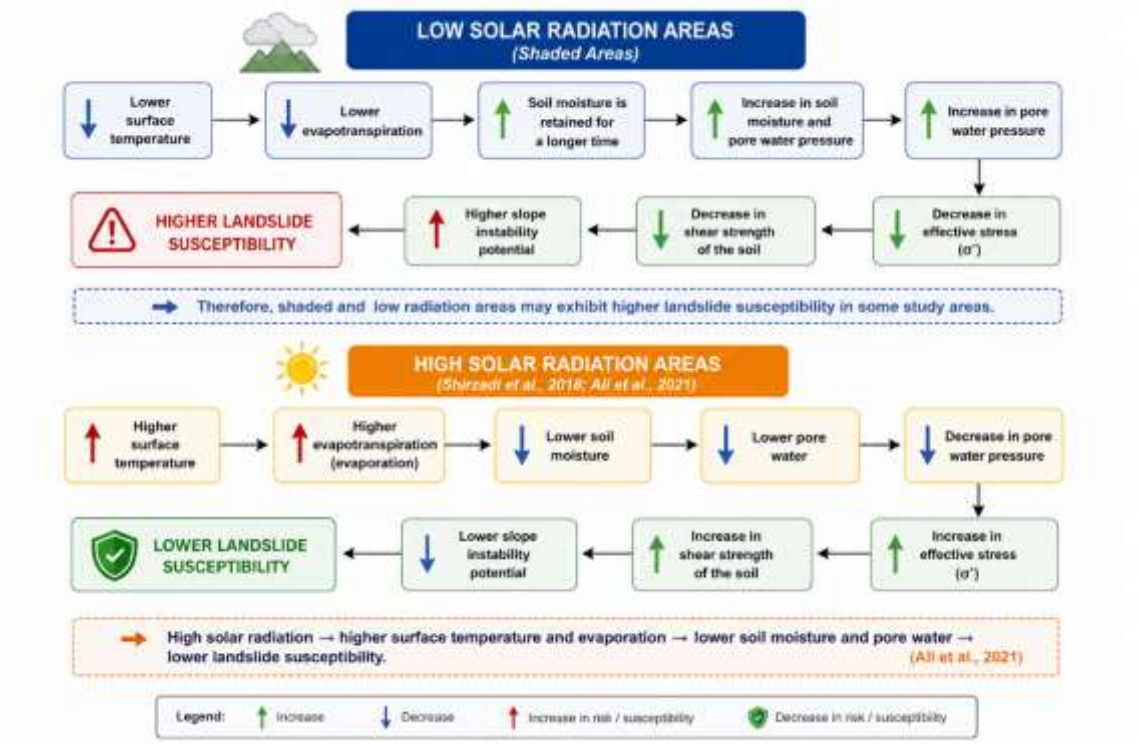


Figure 1. Conceptual framework illustrating the influence of solar radiation on soil moisture dynamics and landslide susceptibility [modified after 3, 1].

V. Advantages of the Parameter

Area Solar Radiation (ASR) is an important topographic parameter because it integrates the combined effects of slope and aspect into a single energy-based variable. While slope controls the angle at which solar radiation reaches the ground surface, aspect determines the orientation of a hillslope relative to the sun. In addition to these terrain characteristics, the calculated amount of solar radiation is also influenced by latitude, the seasonal and daily position of the sun, topographic shading, sky-view factor, and atmospheric conditions. Consequently, compared with aspect—which merely represents geometric orientation—ASR provides a more physically meaningful representation of solar energy received by the land surface.

Ghasemian et al. [2] summarized the rationale for incorporating solar radiation into landslide susceptibility analyses as follows:

- Higher solar radiation → increased evaporation;
- Increased evaporation → reduced soil moisture;
- Lower solar radiation → more favorable conditions for vegetation development;
- Increased vegetation cover → improved slope stability.

Accordingly, Solar Radiation is regarded as an important terrain variable because it indirectly controls both soil moisture conditions and vegetation development, two key factors influencing slope stability.

A fundamental advantage of ASR over the conventional aspect parameter is that it expresses terrain orientation as a quantitative physical energy value rather than as a categorical directional class. Whereas aspect simply identifies the direction a slope faces, ASR quantifies the amount of solar energy actually received by that slope after accounting for topographic and astronomical conditions. This provides a more realistic representation of the environmental processes influencing evapotranspiration, soil moisture dynamics, vegetation growth, and ultimately landslide susceptibility

Landsat thermal bands and MODIS temperature products are widely used for the determination of LST. When LST is derived from Landsat imagery, the digital numbers (DN) of the thermal bands are first converted into spectral radiance (Lλ), after which Brightness Temperature is calculated from the resulting radiance values [9, 10].

The main parameters used for temperature calculation from Landsat thermal data include:

- Lλ: Spectral radiance at the sensor aperture,
- G: Gain coefficient,

- B: Bias coefficient,
- K1: First thermal calibration constant,
- K2: Second thermal calibration constant,
- T: Temperature in Kelvin,

as described in [9, 10].

Main Data Sources Used in the Literature

- Landsat thermal bands: Calculation of LST and brightness temperature,
- MODIS LST (MOD11A2): Assessment of temporal variations in LST,
- MODIS NDVI (MOD13C2): Assessment of temporal variations in vegetation,
- Google Earth Engine (GEE): Processing of long-term satellite imagery and LST analyses,
- Multi-temporal satellite imagery: Assessment of temporal variations in LST,
- Field observations and official landslide inventories: Comparison and validation of remote sensing results with landslide distribution.

In long-term studies, LST and NDVI data can be evaluated together to investigate the surface temperature–vegetation–landslide relationship. In some studies, these datasets have been analyzed at a spatial resolution of 1 km.

Temperature Variables Used

Surface temperature can be evaluated as a single value or through variables representing daily temperature variations, including:

- Tmax: Daily maximum surface temperature,
- Tmin: Daily minimum surface temperature.

Analytical Methods

Several statistical methods have been used to determine temporal variations in LST and its relationships with other environmental parameters, including:

- Linear regression,
- Ordinary Least Squares (OLS),
- Mann–Kendall trend test,
- Pearson correlation analysis.

These methods are used to identify long-term trends in LST and NDVI and subsequently compare these trends with landslide intensity and spatial distribution.

In landslide susceptibility studies, the resulting susceptibility levels can be classified depending on the methodology applied and are generally expressed as:

- Low,
- Moderate,
- High.

VI. DISCUSSION

The calculation of Area Solar Radiation (ASR) depends on several terrain attributes derived from a Digital Elevation Model (DEM), particularly slope, aspect, and the overall topographic configuration of the landscape [1]. Consequently, the accuracy of the generated ASR layer is directly influenced by the quality and spatial resolution of the underlying DEM.

Although ASR provides physically meaningful information, several limitations restrict its widespread application in landslide susceptibility studies.

Computational complexity. ASR calculation is considerably more computationally intensive than most conventional topographic parameters. Processing large, high-resolution DEMs requires substantial computational resources and longer processing times, which may explain why this parameter has been omitted from many earlier studies, particularly before the widespread availability of dedicated GIS tools.

Dependence on software settings. The calculated radiation values are affected by software-specific algorithms and user-defined parameter settings. Furthermore, different studies report solar radiation using different units (e.g., Wh m^{-2} or kWh m^{-2}), making direct comparison between studies more difficult.

Correlation with other terrain variables. Because ASR is primarily derived from slope and aspect, a certain degree of multicollinearity may exist among these variables. Therefore, multicollinearity diagnostics should be performed before simultaneously including ASR, slope, and aspect in susceptibility models.

Limitations of regression-based estimation methods. McCune and Keon [11] proposed three regression equations for estimating annual solar radiation, each with specific applicability conditions. According to Parks [10], the first equation is applicable worldwide but generally produces less accurate estimates. The second equation performs better when slope angles range between 0° and 60°, while the third equation is recommended for study areas located between 30° and 60° latitude with terrain slopes below 60°, where it provides the highest prediction accuracy [9].

Variable importance depends on the modelling approach. Ali et al. [1] reported that the ReliefF importance weight of Annual Solar Radiation was greater than zero; therefore, the parameter was retained as an input variable in the FDEMATEL-ANP, Naïve Bayes, and Random Forest models. However, its contribution was lower than that of several major conditioning factors, indicating that its predictive importance is generally moderate rather than dominant.

Nonlinear relationship with landslide occurrence. The relationship between solar radiation and landslide susceptibility is not universally unidirectional. Increased solar radiation may enhance evapotranspiration, reduce soil moisture, and improve slope stability. Conversely, prolonged and intense drying may promote desiccation cracking in fine-grained soils, accelerate weathering, and facilitate rapid rainfall infiltration during subsequent precipitation events. Likewise, the influence of solar radiation on vegetation development varies according to climate, elevation, water availability, and vegetation type. Therefore, broad generalizations such as "higher solar radiation always increases slope stability" or "lower solar radiation always increases landslide susceptibility" should be avoided [3].

Dependence on geographic location. In the Northern Hemisphere, south-facing slopes generally receive greater solar radiation, whereas north-facing slopes are typically cooler and more shaded. However, this pattern may vary depending on latitude, season, topographic shading, and local climatic conditions. Similarly, east-facing slopes receive more direct solar radiation during the morning, while west-facing slopes receive greater afternoon exposure.

Influence of slope and aspect. Slope determines the angle at which incoming solar radiation reaches the ground surface, whereas aspect controls the orientation of the terrain relative to the sun. Consequently, two slopes located at the same elevation and latitude may receive substantially different amounts of annual solar radiation because of differences in slope angle or aspect.

Moderate predictive performance. Shirzadi et al. [3] concluded that Solar Radiation exhibited only moderate predictive importance under certain modelling scenarios. Despite its physical relevance, ASR remains less frequently employed in landslide susceptibility studies than conventional terrain parameters such as slope and aspect.

Limited discriminatory power as an individual variable. Gorsevski et al. [7] reported that solar radiation alone was insufficient to clearly distinguish landslide-prone areas. Consequently, they recommended evaluating ASR together with other topographic and hydrological variables rather than using it as an independent predictor.

Dependence on additional environmental controls. Evapotranspiration is not controlled solely by solar radiation. Air temperature, wind speed, relative humidity, soil properties, vegetation characteristics, and water availability all substantially influence the process. Therefore, the effect of ASR on soil moisture and landslide susceptibility should be interpreted together with precipitation, land cover, and other environmental variables rather than in isolation.

Overall, although Area Solar Radiation provides a physically meaningful representation of the energy available at the land surface and contributes valuable information to landslide susceptibility assessment, its effectiveness depends on DEM quality, climatic conditions, terrain characteristics, and the modelling approach employed. Consequently, ASR should be regarded as a complementary conditioning factor whose predictive value is maximized when integrated with other topographic, hydrological, geological, and environmental variables.

Gorsevski et al. [7] identified Annual Solar Radiation (ASR) as one of the principal topographic variables incorporated into their landslide susceptibility model. In contrast, Ghasemian et al. [2] applied the

Information Gain Ratio (IGR) algorithm and retained only the sixteen most informative variables from an initial set of twenty-five conditioning factors, indicating that the contribution of solar radiation may vary depending on the feature-selection method and the characteristics of the study area.

The influence of solar radiation on vegetation development and slope stability is strongly controlled by regional climate and water availability. In humid environments, higher solar radiation generally promotes vegetation growth while simultaneously enhancing evapotranspiration. In semi-arid and arid regions, however, excessive solar radiation may substantially reduce soil moisture, thereby limiting vegetation development. Consequently, the relationship between Area Solar Radiation (ASR) and landslide susceptibility is indirect, region-specific, and inherently nonlinear.

The literature reviewed in this study further indicates that solar radiation should not be interpreted as an independent controlling factor. Instead, its influence is mediated through several interconnected environmental processes, including soil moisture dynamics, evapotranspiration, vegetation development, weathering, freeze–thaw activity, and snowmelt. The relative importance of these mechanisms varies according to climate, geology, topography, and land-cover characteristics. Therefore, the predictive contribution of ASR should always be evaluated together with other conditioning factors rather than in isolation

VII. CONCLUSION

Solar radiation influences landslide susceptibility primarily through its effects on evapotranspiration, soil moisture, and vegetation dynamics. Increased solar radiation generally enhances evapotranspiration, leading to lower soil moisture and reduced pore water pressure, which may contribute to improved slope stability under appropriate environmental conditions.

Overall, the literature suggests the following general relationships:

Lower solar radiation → higher soil moisture → potentially greater landslide susceptibility.

Higher solar radiation → enhanced vegetation development and evapotranspiration → potentially greater slope stability.

However, these relationships are not universal. The influence of solar radiation varies according to climatic conditions, topographic characteristics, geological settings, vegetation cover, and soil properties. Therefore, generalized assumptions that either high solar radiation always increases slope stability or low solar radiation invariably promotes landslides should be avoided.

The reviewed studies further demonstrate that Annual Solar Radiation is generally less influential than major conditioning factors such as slope and lithology. Nevertheless, it often exhibits greater predictive importance than several commonly used variables and therefore represents a valuable complementary conditioning factor in landslide susceptibility modelling [1]. Its contribution is maximized when evaluated together with precipitation, elevation, curvature, distance to rivers and roads, and other topographic, hydrological, and environmental variables.

In conclusion, Area Solar Radiation is a physically meaningful terrain parameter that integrates the combined effects of slope, aspect, solar geometry, and topographic shading. Although its influence on landslide susceptibility is predominantly indirect, incorporating ASR into susceptibility models can improve the representation of surface energy conditions and associated hydrological processes, particularly when high-quality DEMs and appropriate modelling approaches are employed.

REFERENCES

[1] Ali, S. A., Parvin, F., Vojteková, J., Costache, R., Linh, N. T. T., Pham, Q. B., Vojtek, M., Gigović, L., Ahmad, A., & Ghorbani, M. A. (2021). GIS-based landslide susceptibility modeling: A comparison between fuzzy multi-criteria and machine learning algorithms. Geoscience Frontiers, 12(2), 857–876. https://doi.org/10.1016/j.gsf.2020.09.004

[2] Ghasemian, B., Shahabi, H., Shirzadi, A., Al-Ansari, N., Jaafari, A., Kress, V. R., Geertsema, M., Renoud, S., & Ahmad, A. (2022). A robust deep-learning model for landslide susceptibility mapping: A case study of Kurdistan Province, Iran. Sensors, 22(4), Article 1573. https://doi.org/10.3390/s22041573

[3] Shirzadi, A., Shahabi, H., Chapi, K., Shirzadi, A., Al-Ansari, N., Ahmad, B. B., et al. (2018). Novel GIS-based machine learning algorithms for shallow landslide susceptibility mapping. Sensors, 18(11), Article 3777. https://doi.org/10.3390/s18113777

[4] Nhu, V.-H., Shirzadi, A., Shahabi, H., Chen, W., Clague, J. J., Geertsema, M., Jaafari, A., Avand, M., Miraki, S., Asl, D. T., Pham, B. T., Ahmad, B. B., & Lee, S. (2020a). Shallow landslide susceptibility mapping by Random Forest base classifier and its ensembles in a semi-arid region of Iran. Forests, 11(4), Article 421. https://doi.org/10.3390/f11040421

[5] Pawluszek, K., & Borkowski, A. (2017). Impact of DEM-derived factors and analytical hierarchy process on landslide susceptibility mapping in the region of Rożnów Lake, Poland. Natural Hazards, 86(2), 919–952. https://doi.org/10.1007/s11069-016-2725-y

[6] Krassakis, P., Karavias, A., Nomikou, P., Karantzalos, K., Koukouzas, N., Athinelis, I., Kazana, S., & Parcharidis, I. (2023). Multi-hazard susceptibility assessment using the analytical hierarchy process in coastal regions of South Aegean Volcanic Arc Islands. GeoHazards, 4(1), 77–106. https://doi.org/10.3390/geohazards4010006

[7] Gorsevski, P. V., Brown, M. K., Panter, K., Onasch, C. M., Simic, A., & Snyder, J. (2016). Landslide detection and susceptibility mapping using LiDAR and an artificial neural network approach: A case study in the Cuyahoga Valley National Park, Ohio. Landslides, 13(3), 467–484. https://doi.org/10.1007/s10346-015-0587-0

[8] Nhu, V.-H., Shirzadi, A., Shahabi, H., Chen, W., Clague, J. J., Geertsema, M., Jaafari, A., Avand, M., Miraki, S., Asl, D. T., Pham, B. T., Ahmad, B. B., & Lee, S. (2020b). Shallow landslide susceptibility mapping: A comparison between Logistic Model Tree, Logistic Regression, Naïve Bayes Tree, Artificial Neural Network, and Support Vector Machine algorithms. International Journal of Environmental Research and Public Health, 17, Article 2749. https://doi.org/10.3390/ijerph17082749

[9] *Tekin, S. (2019). Göksu Nehri Havzasının coğrafi bilgi sistemleri tabanlı jeomorfometrik analizi ve niceliksel heyelan olası tehlike değerlendirmesi (Doktora tezi, Çukurova Üniversitesi, Fen Bilimleri Enstitüsü). Yükseköğretim Kurulu Ulusal Tez Merkezi*

[10] Parks, S. A. (2004). Solar Radiation ArcView Script [Computer software/ArcView script]. ESRI ArcScripts

[11] *McCune, B., & Keon, D. (2002). Equations for potential annual direct incident radiation and heat load. Journal of Vegetation Science, 13(4), 603–606. https://doi.org/10.1111/j.1654-1103.2002.tb02087.x*

6th International Conference on Modern and Advanced Research

August 10-11, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Flow Balancer Machine Design and Manufacturing

Mohamad Ramzı ABOMAGHARA[1], İlhan ASİLTÜRK[2]

[1]Mechanical Engineering / Institute of Science, Necmettin ERBAKAN University, Türkiye
[2]Mechanical Engineering / Institute of Science, Necmettin ERBAKAN University, Türkiye

[1]mramziabomaghara@gmail.com, [2]iasilturk@erbakan.edu.tr

Abstract **–**In this study, a flow balancer system that enables the real-time monitoring of raw material flow stability and production yield in flour mill production lines was investigated. Within this scope, an industrial performance analysis was conducted on a flow weigher system installed on the raw material preparation line, which operates based on the principle of gravitational flow and performs mass flow rate measurement by utilizing centripetal acceleration components. Based on real-time data obtained from the plant's SCADA interface over different time intervals, the target setpoint values (SP) entered by the operator and the instantaneous process values (PV) measured by the device were evaluated sequentially. During the initial commissioning of the system, an instantaneous flow rate of 28,503 kg/h (PV) was recorded against the target flow rate of 28,000 kg/h (SP), corresponding to a relative deviation of 1.79%. Under continuous production conditions, this deviation gradually decreased through the PID control loop and the system's adaptive feed control dynamics, reaching an exceptionally precise level of 0.28%, with a final measured flow rate of 28,080 kg/h (PV). The findings demonstrate that the flow weigher unit successfully dampens instantaneous density fluctuations in bulk solid flow, maintains the target recipe ratios with high stability, and provides sustainable process accuracy in industrial flour mill automation.

Keywords – Flow Balancer, Flow Scale, Wheat Flow Measurement, Grain Flow Control, Flow Regulator, Precise Feeding, Continuous Flow Provision, Efficient Processing Processes.

I. INTRODUCTION

The agricultural sector is continuously developing various solutions to improve efficiency in product processing operations in parallel with advancing technologies worldwide. In particular, accurate and consistent control of product flow is of great importance during the post-harvest processing, storage, and marketing of grains such as wheat. In this context, flow balancer machines are important equipment used to ensure that wheat is transferred at a predetermined rate and quantity suitable for processing and storage. By providing an efficient and balanced product flow, flow balancer machines aim to minimize disruptions throughout the processing process. For products requiring precise processing, such as wheat, proper flow balancing plays a critical role in maintaining product quality and improving processing efficiency. However, the performance of flow control machines used in processing operations largely depends on the quality of their design and manufacturing stages. In this study, the design and manufacturing processes of a wheat flow balancer machine are described in detail, and a high-performance flow control mechanism is developed.

II. MATERIALS AND METHOD

In this study, a flow balancer machine was designed to enable the weighing of wheat while it is in continuous flow. The main objective of the design is to ensure accurate, rapid, and uninterrupted measurement of the product quantity during continuous flow. In particular, measuring wheat during its transfer to processing lines in flour mills and grain processing facilities is critical for production planning,

inventory tracking, and quality control. Therefore, the developed system was designed to meet industrial measurement requirements while also providing ease of use in practical applications.

A. Weighing Unit

A load cell was used in the product measurement section of the machine to accurately determine the amount of flowing wheat. The load cell ensures weighing accuracy by converting the force acting on it into an electrical signal. In this system, when the product is transferred into the weighing hopper, the load cell detects the change in weight instantaneously and transmits the measurement result to the control unit. The use of the load cell enables the weighing process to be performed with high accuracy, eliminating errors associated with manual measurement. This weighing unit is one of the most important components of the system, particularly due to its ability to perform measurements under dynamic flow conditions. In the system, the difference between the target mass flow rate (SP) entered by the operator through the SCADA interface and the actual wheat flow rate (PV) instantaneously measured by the load cell positioned beneath the curved sheet is defined as the error, (e(t) = SP - PV). The instantaneous error, (e(t)), represents the deviation at a given moment as a function of time. It is the value that the controller continuously attempts to reduce to zero and is calculated continuously by the PID algorithm running within the PLC. In order to eliminate the calculated error, the FINEX geared motor, which acts as the actuator, is driven to adjust the position of the flap instantaneously. The automatic control block diagram of the system is shown in the figure below.

B. Figures

.

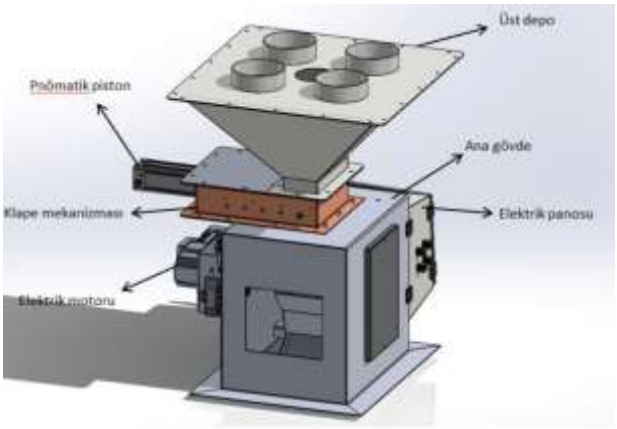


Fig. 1 Flow Balancer Machine

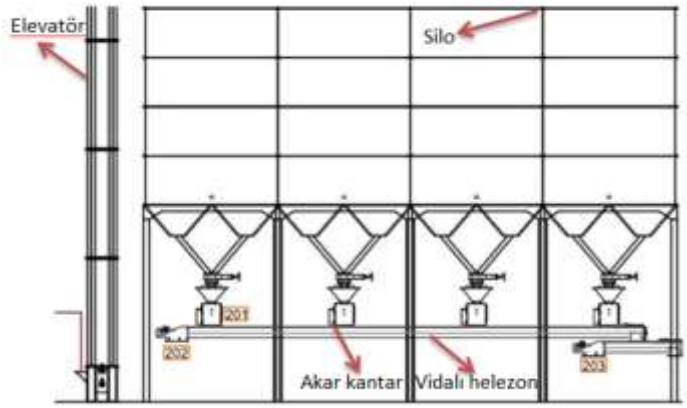


Fig. 2 Flow Diagram

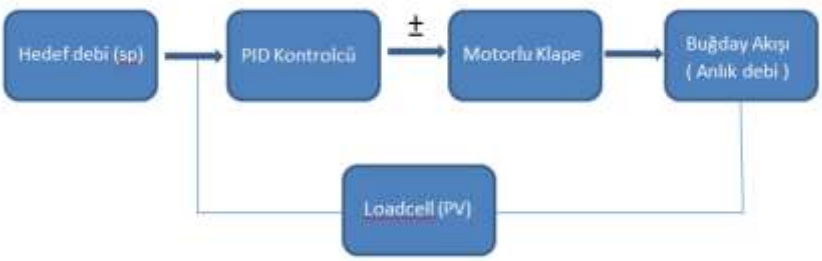


Fig. 3 Closed-Loop Automatic Control Block Diagram of the System

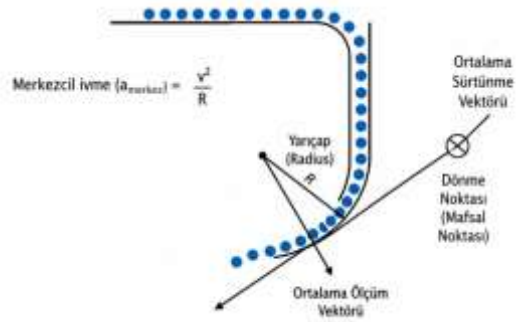


Fig. 4 Schematic Representation of Force Vectors

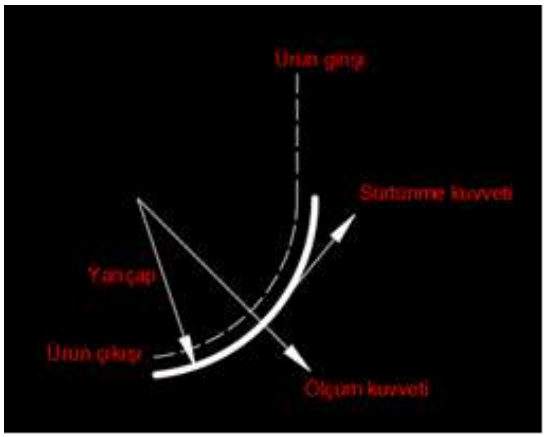


Fig. 5 Schematic Representation of Force Vectors

III. RESULTS

In the developed flow balancer system, open-loop and closed-loop control conditions were compared. During open-loop operation, as the amount of wheat in the upper hopper decreased, the pressure decreased accordingly. Therefore, fluctuations of approximately 3–5% in the flow rate were observed even when the flap remained in a fixed position. In contrast, when the load cell- and PID-controlled closed-loop system was activated, the flow rate was maintained at the target value even when the wheat level in the hopper changed. Thanks to the tuned parameters of the PID controller, the system rapidly reached the target flow rate, while sudden overloading of the weighing unit was prevented.In the dynamic flow tests conducted, considering independent measurement deviations, the system was found to operate with a very low average error rate of approximately 0.2%. This result demonstrates that the closed-loop control structure and the motor-controlled precision flap successfully meet the required high weighing accuracy and overload prevention performance in flour and grain processing plants.

IV. DISCUSSION

At the final stage of the observation period, the fifth and last SCADA data obtained from the system is presented in Figure 5.5. While the target mass flow rate (SP) remained constant at 28,000 kg/h on the automation interface, the instantaneous flow rate value (PV) measured by the flow balancer was recorded

as 28,080 kg/h. During this final measurement stage, the relative deviation between the instantaneous process value and the setpoint decreased to a negligible level of 0.28%, clearly demonstrating the system's ability to adapt over time and its control stability.

V. CONCLUSION

1. The designed compact flow balancer machine was able to perform accurate and reliable weighing of wheat during continuous flow.
2. The L6D 35 kg load cell used in the system provided measurements with an accuracy of ±1%, thereby improving measurement precision compared with manual methods.
3. The upper flap was controlled by a pneumatic cylinder, while the lower flap was controlled by a geared motor, enabling successful adjustment of the product flow.
4. The compact design of the machine provides advantages in terms of space saving, portability, and low energy consumption.

REFERENCES

[1] Akduman, M. (2021). Düşük karbonlu ST37 yapı çeliğinin ısıl işlem sonrası yapısı ve mekanik özelliklerinin incelenmesi [Doktora tezi].

[2] Çağlar, H. (2022). ST37 çeliğinin mekanik ve termal yüzey işlemleri uygulamaları ile mikroyapı ve mekanik özelliklere etkisinin araştırılması. Karabük Üniversitesi, Makine Mühendisliği Anabilim Dalı.

[3] Eren, H. (2023). Konya ilindeki buğday ekiliş alanlarında arpa sarı cücelik virüsü (Barley Yellow Dwarf Virus, BYDV), buğday çizgi mozaik virüsü (Wheat Streak Mosaic Virus, WSMV) ile buğday cücelik virüsü (Wheat Dwarf Virus, WDV) hastalık etmenlerinin araştırılması. https://acikerisim.selcuk.edu.tr/items/df22d578-0376-4c18-9b03-380a22a4154c.

[4] Gülbaş, M. C. (2022). Dynamic weighing method for checkweigher [Yüksek lisans tezi, İstanbul Teknik Üniversitesi]. İTÜ Polen Akademik Açık Arşivi. https://hdl.handle.net/11527/25433

[5] Navid, H., Taylor, R. K., Yazgi, A., Wang, N., Shi, Y., & Weckler, P. (2015). Detecting grain flow rate using a laser scanner. Transactions of the ASABE, 58(5), 1185–1190. https://doi.org/10.13031/trans.58.11097

[6] Özdağ, M. (2023). Vektör kontrol yöntemi için deney düzeneği tasarımı, gerçeklemesi ve bu yöntemin üç fazlı elektrik motorlarında performans karşılaştırması. https://polen.itu.edu.tr/server/api/core/bitstreams/fc4fa6f3-7640-4817-8cbc-0cb273530d0a/content

[7] Savaşkan, T. (2017). Malzeme bilimi ve malzeme muayenesi. Celpler Matbaacılık. https://dokumanapi.bys.subu.edu.tr/File/2024/2/23/bd8ef503-d411-4925-933e-bc00461adeed.pdf

[8] Timmer Arends, A. (2015). Design of a dynamic weigh system for conveyors [Internship report, University of Twente]. UT Student Theses. https://essay.utwente.nl/69531/1/Design_of_a_dynamic_weigh_system_for_conveyors.pdf

[9] Tugal, M. E. (2019). Buğdayda depolama koşullarının belirlenmesi ve bu koşulların mikotoksin oluşumuna etkisi [Yüksek lisans tezi, Fen Bilimleri Enstitüsü].

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Individual and Household Determinants of Female Labour Force Participation. Micro-Level Evidence from the Western Balkans

Loredana Sulejmani[*], Elena Polo [2]

[1]*Economics Department/Faculty of Economy, University of Tirana, Albania*
[2]*EU Legislation and Policy Department/Institute of European Studies, University of Tirana, Albania*

[*]*(loredana.sulejmani@unitir.edu.al)*

Abstract **–**Female labour force participation remains comparatively low across the Western Balkans, yet regional averages may conceal substantial differences in the factors associated with women's labour market engagement. This study examines whether the determinants of female participation operate consistently across Albania, Bosnia and Herzegovina, Kosovo, Montenegro, North Macedonia, and Serbia. Using the fourth wave of the Life in Transition Survey and a sample of 2,453 women aged 18–64, separate country-level probit models are estimated under a common specification.

The results reveal considerable cross-country heterogeneity. Education emerges as the only consistently positive factor of labour force participation, although the magnitude of the tertiary-education premium differs substantially across the six economies. Participation also follows a broadly inverted-U life-cycle pattern. Other associations are considerably more country dependent. Young children are associated with lower participation particularly in Montenegro and North Macedonia, while poor self-assessed health is significantly associated with lower participation in four countries. Urban residence is positively associated with participation in Albania and Serbia, whereas support for the male-breadwinner norm is negatively associated with participation across all six economies but reaches statistical significance only in Serbia. Kosovo stands out for its persistently low predicted participation across the age distribution.

Policies aimed at strengthening women's economic participation should therefore respond to the specific constraints identified within each country rather than relying on a uniform approach. Measures addressing care responsibilities, health-related barriers, labour-market access, and gender norms may need to differ across national contexts. Given the cross-sectional nature of the data, the estimated relationships are interpreted as conditional associations rather than causal effects.

Keywords – female labour force participation; Western Balkans; labour supply; education; gender norms; probit model

I. INTRODUCTION

Despite recent improvements, female labour force participation rates in the six Western Balkans countries (WB6) remain among the lowest in Europe. In 2025, activity rates among women aged 15 to 64 were still below 55 percent in Bosnia and Herzegovina, North Macedonia, and Kosovo. Albania and Serbia record the highest participation rates in the region. Albania's rapid progress during the recent five years strengthened its regional position and enabled it to surpass Serbia, which had historically recorded the highest rate in the region [1], [2].

The persistence of low participation rates is economically costly, especially in this new economic and social reality that the region is experiencing. With shrinking and aging work-age populations in the last

decade, combined with the early emigration of the unskilled labour force and nowadays of the skilled labour force, the unused potential of the female labour force is probably the only domestic source of growth [1].

The region shares a common historical background, in addition to similarities in geographic position, economic state, and its agenda to join the European Union. Albania and the successor states of former Yugoslavia inherited a socialist legacy of relatively high female employment for the time and then passed through a transition that destroyed formal jobs, expanded informal and agricultural work, and returned care responsibilities to households [1].

Theoretical foundations identify the mechanisms that affect participation [3]-[5], but they also acknowledge that these mechanisms are not uniformly applied in any given context. These microeconomic theories do not determine the factors' sign or amplitude in any given economy. Establishing which factors bind demands individual and household-level data, and the increasing availability of household surveys has resulted in significant empirical literature estimating discrete choice models of participation across counntries and times. The literature review agrees on a global set of determinants such as age and life stage, marital status, caregiver responsibilities, education, household income, societal norms, and place of residence. However, conclusions reached for any other region or country cannot be assumed to apply to the Western Balkans. Micro-level evidence for the region is scarce and fragmented. The majority of studies evaluate variables for individual countries, using national surveys that are not harmonized across countries and are not recent [6]-[9]. Recently there is comparable cross-national micro-evidence for the entire region [10], but the focus of the paper has been on the total labour force participation not just female.

A. Age and the Life Cycle

Female labour supply fluctuates more than male labour supply across the life cycle [11], [12]. Vlasblom and Schippers [13], who estimated country-specific logit models using European Labour Force Survey data from six Western European countries between 1992 and 1999, discovered that parental status considerably mediates age effects. Viitanen [14], using eight waves of the European Community Household Panel for thirteen European countries, discovers significant state dependence. Past participation strongly predicts current participation, implying that part of the apparent age effect is due to accumulated labour market attachment rather than age.

Studies of transition and Western Balkan economies consistently show an inverted-U age profile, but with an earlier peak than in Western Europe. Gashi et al. [9] indicate a high at around age 42 in Kosovo, Garcia-Pereiro [8] reports a non-linear increase over prime ages followed by a fall in Albania, and Andonova et al. [10] report a peak between ages 36 and 43 across six Western Balkan countries. Not all evidence confirms this pattern. Bakhtavoryan and Zohrabyan [15] found a uniformly negative age effect in the South Caucasus. Biagi et al. [16] warn that a single cross-section can disorder life-cycle and cohort effects because participation at a particular age varies significantly among birth cohorts, and that panel or repeated cross-section methods are required to distinguish the two.

B. Marital Status

Household specialization is predicted to reduce married women's participation. Vlasblom and Schippers [13] limit their sample to coupled women specifically because their participation behavior differs so dramatically from that of single women, and they discover marriage to be a major source of cross-country heterogeneity. Klasen and Pieters [17] and Klasen et al. [18] also focus on married women, claiming that the causes of involvement differ qualitatively by marital status in low-income settings.

The evidence from the Western Balkans consistently shows the same sign. Shehaj and Abazi-Alili [6] report negative marital-status coefficients for Albania and North Macedonia, while Miluka and Tsushima [7] find that marriage reduces participation primarily among urban Albanian women, Garcia-Pereiro [8] reports a negative association for Albania, and Gashi et al. [9] find a negative effect in Kosovo. On the other hand, Bakhtavoryan and Zohrabyan [15] discover a marriage penalty in Azerbaijan and Georgia but not in Armenia. Taheri et al. [19] document a reversal in Iran, where the marriage coefficient shifts from negative in 2008 to positive in 2018. Andonova et al. [10] demonstrate that the marriage effect in the Western Balkans is largely explained by gender-role attitudes and household structure. As a result, rather than operating independently, marital status is influenced by conventions and household characteristics.

C. Care Responsibilities

Care responsibilities are the primary limitation on women's time, often measured by the number of young children or elderly individuals in the household. The detrimental impact of young children on participation is one of the most substantiated conclusions in European research. Vlasblom and Schippers [13] demonstrate that the penalty diminished in Germany and the Netherlands from 1992 to 1999 due to the expansion of daycare and part-time employment, but it persisted robustly in Southern Europe. Thevenon and Horko [22] ascertain that the age of the youngest child is especially significant in Eastern Europe and the United Kingdom, where mothers become inactive while children are under three and re-enter the workforce once formal education begins. Keck and Saraceno [23], utilizing EU-SILC data, identify childcare for children under three as the principal policy instrument, while Cipollone et al. [24] demonstrate that family and labour market policies enhance involvement, particularly among younger and more educated women. Bicakova [25] identifies that gender disparities are most pronounced when the youngest kid is between the ages of zero and four, whereas Christiansen et al. [26], [27] support the adverse impact of children across twenty-four nations. Lovaglio and Perrelli [28] observe that very young children decrease employment in France and Germany, whereas in Italy, general household composition is more significant. However, Ingco and Pilitro [29] determine that the presence of an extra kid increases the likelihood of married women's participation in the workforce in the Philippines by around three percentage points, attributing this to an income-need mechanism. Lam and Elsayed [30] and Turon [31] further support that the magnitude and direction of the fertility effect are influenced by the availability of childcare, home structure, and the nature of jobs accessible to mothers.

D. Education

Education is the most consistently favorable factor, consistent with human capital theory. Years of education increases the opportunity cost of remaining outside the labour force and enhances access to formal work. Vlasblom and Schippers [13] identify a distinct positive gradient in Western Europe, while Christiansen et al. [26], [27] document a robust gradient on both the extended and intense margins. Shehaj and Abazi-Alili [6] identify education as a crucial positive factor in Albania and North Macedonia. Miluka and Tsushima [7] demonstrate that education and experience enhance participation in Albania, with markedly greater effects in urban areas compared to rural ones. Gashi et al. [9] indicate a significant positive gradient in Kosovo.

The universality of the educational effect has, however, been contested. Klasen and Pieters [17] identify a U-shaped correlation in urban India, where participation is elevated among the least educated, decreases at intermediate educational levels, and subsequently increases at the tertiary level. Klasen et al. [18], in their comparison of eight developing and rising economies, identify both linear and U-shaped trends, demonstrating that cross-country disparities are primarily influenced by variations in the returns to characteristics rather than the qualities themselves. Taheri et al. [19] observe that the favorable tertiary effect in Iran diminished from 2008 to 2018. Andonova et al. [10] affirm the beneficial impact in the Western Balkans with significant national interactions.

E. Household Income

Household income findings are diverse and seem to correlate with the level of development. Klasen and Pieters [17] and Klasen et al. [18] present negative income coefficients, indicative of necessity-driven labour supply, while Bakhtavoryan and Zohrabyan [15], Garcia-Pereiro [8], and Ingco and Pilitro [29] see positive correlations. In the Western Balkans, the indicator is inconsistent. Shehaj and Abazi-Alili [6] document a negative household wealth effect in Albania and a positive one in North Macedonia. Gashi et al. [9] demonstrate that the presence of an unemployed male household member increases women's labour participation in Kosovo, aligning with the added-worker effect.

F. Societal Norms

Social norms are widely acknowledged as the remnants that observable traits fail to elucidate. Christiansen et al. [26], [27] demonstrate that exposure to a working mother and egalitarian gender attitudes are significant favorable predictors of female employment throughout Europe. Taheri et al. [19] view the beneficial impact of a husband's post-secondary education in Iran as an indicator of more progressive attitudes inside the household. Andonova et al. [10] identify a significant correlation between gender-role

attitudes and female inactivity in the Western Balkans, demonstrating that this correlation's intensity differs among nations within the same region.

F. Residential

Urbanization is anticipated to enhance participation by diminishing traditional norms, enhancing access to the labour market, and increasing the availability of formal childcare. Klasen and Pieters [17] as well as Klasen et al. [18] confine their investigation to urban locations, asserting that urban and rural labour markets exhibit qualitative differences and that female labour is quantified more precisely in urban contexts. The evidence from the Western Balkans is inconclusive. Miluka and Tsushima [7] observe much greater returns to education and experience in urban areas compared to rural Albania, where women face limitations due to unpaid agricultural labour, inadequate infrastructure, and a scarcity of non-farm options. Additionally, Garcia-Pereiro [8] indicates an urban premium for Albania. Gashi et al. [9] ascertain that rural residency diminishes participation in Kosovo. Shehaj and Abazi-Alili [6] indicate a negative urban coefficient for Albania and a positive one for North Macedonia.

The paper addresses the gaps found in the literature review. Using the fourth round of the Life in Transition Survey, it estimates the individual and household determinants of female labour force participation for the six Western Balkan countries within a single harmonised framework, applying separate country-specific probit models on a common specification. The contribution is three-dimensional. The research presents comparable micro-econometric estimates for the Western Balkans, where previous evidence is primarily comprised of isolated single-country studies. Secondly, it examines whether the determinants identified in literature such as age, education, caregiving responsibilities, marital status, and residential location, are applicable in a regional context marked by significant informality, inadequate childcare availability, and enduring traditional gender norms. Third, it incorporates two factors that regional literature has predominantly overlooked such as health and the male-breadwinner norm into the same specification as the usual covariates.

The reminder of the paper is structured as follows. Section II specifies the data and the empirical methodology. Section III sets out the findings. Section IV discusses them, and the paper concludes the results in Section V.

II. Materials and Method

The analysis uses the Life in Transition Survey (LiTS), a household survey conducted collabouratively by the European Bank for Reconstruction and Development and the World Bank. LiTS has been conducted in four comparable waves. This paper employs the latest wave, LiTS IV, only for the six Western Balkan countries: Albania, Bosnia and Herzegovina, Kosovo, Montenegro, North Macedonia, and Serbia [32].

The unit of analysis is the individual female respondent. The sample is limited to women between the ages of 18 and 64. Women over 65 demonstrate nearly universal inactivity. Retaining this demographic would skew the computed coefficients away from the behavioral variation of interest, specifically the participation decisions of women who are likely to be employed [33].

A probit specification is used in this paper, whereby the likelihood of participation is represented by the standard normal cumulative distribution function of the linear index. Participation is observed as a binary outcome, so the model estimates how individual and household characteristics are associated with an increase or decrease in the probability that a woman participates in the labour force [34]-[36].

Probit specification is favored over a linear probability model, which might yield projected probabilities beyond the unit interval, enforces a constant marginal effect for each covariate, and inherently creates heteroskedastic errors [37], [38]. A Tobit model is unsuitable in this context, as the outcome is a binary participation indicator rather than a continuous variable subject to censoring at a threshold [36], [39].

Each country is estimated as a separate model, so that every coefficient is allowed to vary across countries such as shown in the Equation (1).

$$P(LFP_{ic}=1)=\Phi\big(\beta_{0,c}+\beta_{1,c}age_{ic}+\beta_{2,c}agesq_{ic}+\beta_{3,c}married_{ic}+\beta_{4,c}child_0_5_{ic}+\beta_{5,c}child_6_15_{ic}+\beta_{6,c}hh_size_{ic}+\beta_{7,c}education_medium_{ic}+\beta_{8,c}education_high_{ic}+\beta_{9,c}urban_{ic}+\beta_{10,c}male_breadwinner_{ic}+\beta_{11,c}health_{ic}+\beta_{12,c}wealth_{ic}\big) \quad (1)$$

where *i* indexes the individual woman, *c* indexes the country, and $\Phi(\cdot)$ is the standard normal cumulative distribution function. The dependent variable (*LFP*) is a binary measure of current labour force participation, defined over a seven-day reference period according to the ILO labour force framework, where the labour force comprises both the employed and unemployed, with the reference period being a designated short duration, such as one week [33]. A respondent is classified as participating (LFP = 1) if, during the week prior to the interview, she was either employed or jobless and actively seeking employment. She is classified as inactive (LFP = 0) otherwise.

Age is incorporated as age-centered (*age*) at the pooled mean of the Western Balkan sample, along with its square (*agesq*), allowing the estimated effect to reflect the life-cycle profile of participation instead of a linear trend. Centring diminishes the link between the linear and quadratic terms.

Household composition is represented by three variables: a binary indication for the presence of at least one child aged 0 to 5 (*child_0_5*), a binary indicator for the presence of at least one child aged 6 to 15 (*child_6_15*), and the total number of household members (*hh_size*). Marital status (*married*) is a binary variable, assigned a value of one for respondents who are formally married, while all other statuses serve as the reference group.

Education was built on the categories of the International Standard Classification of Education [40]. Participants are categorized into three educational levels: low education (ISCED 0 to 2), medium education (*education_medium*), which includes upper-secondary and post-secondary non-tertiary qualifications (ISCED 3 to 4), and high education (*education_high*), encompassing tertiary credentials (ISCED 5 to 8). Two binary variables representing medium and high education are incorporated, with low education serving as the reference group.

Urban residence (*urban*) serves as a binary variable, with rural residence as the baseline reference. Self-assessed health (*health*) is a binary metric, with a value of one for responders indicating poor, or very poor health, with good or very good health serving as the reference category. Gender-role views are represented by a binary variable (*male_breadwinner*), assigned a value of one for respondents who agree with the male-breadwinner assertion *[It is better for everyone involved if the man earns the money and the woman takes care of the home and children],* with disagreement serving as the reference point.

The economic status of a household (*hh_wealth*) is evaluated based on the respondent's self-perception of her household's place on a ten-step ladder, with step 1 indicating the lowest decile of families in her country and step 10 representing the highest decile. The variable is presented in its original continuous form, allowing its coefficient to be read as the effect of advancing one level.

Multiple regressors in the model may be endogenous. Education is possibly associated with unobserved abilities and preferences that also affect participation. In this context, its coefficient may partially contain this omitted variation. The perception of household wealth is influenced by reverse causality, as a woman's job can enhance her household's perceived economic status, while self-reported health and adherence to the male-breadwinner standard may partially reflect an attempt to rationalize her actual labour market position. Addressing these endogeneity issues necessitates the use of instrumental-variable or panel methods, which the current data do not allow. Consequently, each coefficient is interpreted as the change in the probability of participation linked to a characteristic, while controlling for other covariates, rather than as its causal effect.

Standard errors are clustered at the main sample unit level to account for the multi-stage design of the survey, comprising 50 PSU clusters in each country, except for Serbia, which has 49. The delta approach is employed to derive average marginal effects and their corresponding standard errors [34], [35]. All models are computed using Stata 17. Diagnostic evaluations have been estimated to assess the model's misspecification, as well as its specificity, sensitivity, discrimination, and multicollinearity [39], [41]-[43]. Furthermore, the calibration of the models was ensured by employing the Hosmer-Lemeshow goodness-of-fit test [44].

III. RESULTS

Table 1. reports the probit coefficients and Table 2. reports the average marginal effects for the six country-specific models. The quadratic age variable is negative and statistically significant at the 1 percent level across all six economies, supporting an inverted-U participation profile throughout the region.

Table 1. Country-specific probit coefficients of female labour-force participation

	(1) Albania	(2) Bosnia and Herzegovina	(3) Kosovo	(4) Montenegro	(5) North Macedonia	(6) Serbia
LFP						
age	-0.007	-0.021**	-0.011	-0.018***	-0.020***	-0.013*
	(0.008)	(0.009)	(0.008)	(0.007)	(0.007)	(0.008)
agesq	-0.003***	-0.003***	-0.001***	-0.002***	-0.003***	-0.005***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
child_0_5	-0.171	-0.422	-0.207	-0.687***	-0.728***	-0.305
	(0.320)	(0.331)	(0.194)	(0.230)	(0.221)	(0.280)
child_6_15	0.201	0.391	-0.276	-0.071	-0.324*	-0.257
	(0.176)	(0.279)	(0.182)	(0.211)	(0.178)	(0.207)
hh_size	0.001	-0.174	0.015	-0.065	0.108	0.006
	(0.071)	(0.109)	(0.050)	(0.053)	(0.071)	(0.068)
married	-0.313	-0.075	-0.376*	-0.057	0.250	-0.149
	(0.190)	(0.179)	(0.210)	(0.164)	(0.204)	(0.201)
education_medium	0.517***	0.772***	0.529**	0.328	0.902***	0.280
	(0.175)	(0.204)	(0.212)	(0.289)	(0.239)	(0.317)
education_high	1.222***	1.641***	1.431***	0.878***	1.990***	0.993**
	(0.260)	(0.276)	(0.253)	(0.296)	(0.282)	(0.411)
urban	0.669***	0.103	0.149	-0.027	0.293	0.506**
	(0.198)	(0.218)	(0.188)	(0.164)	(0.192)	(0.205)
male_breadwinner	-0.113	-0.299	-0.195	-0.153	-0.227	-0.480***
	(0.220)	(0.199)	(0.142)	(0.158)	(0.167)	(0.176)
health	-0.297*	-0.464**	-0.309**	-0.733***	-0.053	-0.128
	(0.153)	(0.214)	(0.141)	(0.190)	(0.176)	(0.202)
hh_wealth	0.138**	-0.061	0.030	0.020	-0.051	0.071
	(0.057)	(0.062)	(0.051)	(0.045)	(0.051)	(0.058)
Observations	397	331	487	435	416	387
PSU clusters	50	50	50	50	50	49
Pseudo R-squared	0.291	0.277	0.208	0.209	0.338	0.347

Table 2. Country-specific average marginal effects on female labour-force participation

	(1) Albania	(2) Bosnia and Herzegovina	(3) Kosovo	(4) Montenegro	(5) North Macedonia	(6) Serbia
age	-0.006***	-0.008***	-0.000	0.001	-0.005***	-0.006***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)	(0.002)
child_0_5	-0.046	-0.116	-0.057	-0.206***	-0.183***	-0.068
	(0.087)	(0.089)	(0.052)	(0.073)	(0.053)	(0.064)
child_6_15	0.053	0.113	-0.077	-0.020	-0.080*	-0.056
	(0.046)	(0.082)	(0.050)	(0.059)	(0.043)	(0.046)
hh_size	0.000	-0.049*	0.004	-0.018	0.027	0.001
	(0.019)	(0.029)	(0.014)	(0.014)	(0.018)	(0.014)
married	-0.083*	-0.021	-0.111*	-0.016	0.064	-0.032
	(0.050)	(0.050)	(0.064)	(0.046)	(0.052)	(0.043)
education_medium	0.160***	0.228***	0.152**	0.108	0.286***	0.073
	(0.056)	(0.058)	(0.060)	(0.098)	(0.071)	(0.087)
education_high	0.340***	0.498***	0.480***	0.259***	0.573***	0.219**
	(0.069)	(0.080)	(0.081)	(0.096)	(0.075)	(0.097)
urban	0.186***	0.029	0.043	-0.007	0.074	0.113**
	(0.053)	(0.062)	(0.055)	(0.045)	(0.050)	(0.049)
male_breadwinner	-0.030	-0.085	-0.056	-0.043	-0.058	-0.104***
	(0.059)	(0.057)	(0.043)	(0.045)	(0.042)	(0.037)
health	-0.082*	-0.135**	-0.086**	-0.231***	-0.013	-0.028
	(0.044)	(0.061)	(0.039)	(0.066)	(0.045)	(0.045)
hh_wealth	0.037***	-0.017	0.009	0.005	-0.013	0.015
	(0.014)	(0.018)	(0.014)	(0.012)	(0.013)	(0.012)
Observations	397	331	487	435	416	387
PSU clusters	50	50	50	50	50	49

Standard errors in parentheses
* p < 0.10, ** p < 0.05, *** p < 0.01.

Figure 1 illustrates that the profile reaches its maximum in the late thirties to early forties across all countries, aligning with the previously documented maxima for Kosovo [8] and the Western Balkans as a collective region [9].

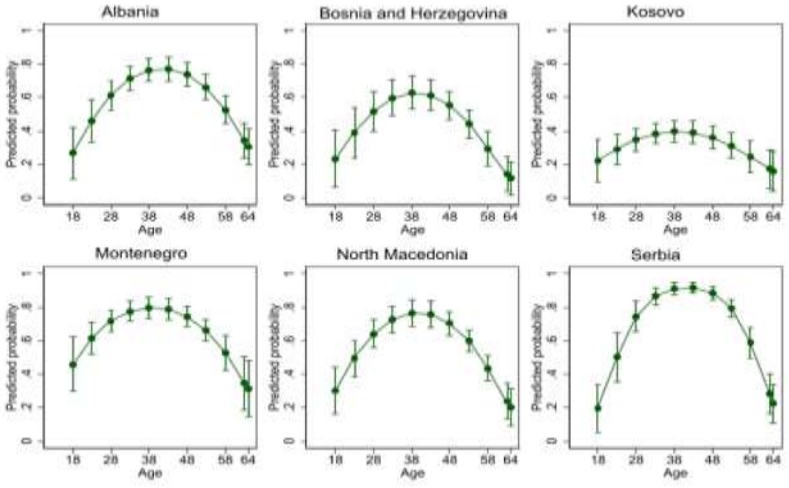


Fig. 1 Predicted female labour force participation by age

The variation among countries lies not in the form but in the magnitude. The highest probabilities are recorded in Serbia, Montenegro, and Albania. The Kosovo profile is both the lowest and flattest. The average marginal effect of age is negative and significant in Albania, Bosnia and Herzegovina, North Macedonia, and Serbia, ranging from 0.5 to 0.8 percentage points per additional year, whereas it is statistically indistinguishable from zero in Kosovo and Montenegro. The average marginal effect, which averages the derivative throughout the whole sample, yields a negative value due to the predominance of women observed on the descending region of the profile, so aligning with the inverted-U shape rather than contradicting it.

Education is the strongest and the only universal determinant. Tertiary education is positive and statistically significant in all six countries, raising the probability of participation by 57.3 percentage points in North Macedonia, 49.8 in Bosnia and Herzegovina, 48.0 in Kosovo, 34.0 in Albania, 25.9 in Montenegro and 21.9 in Serbia, relative to women with lower-secondary education (meaning 8-9 schooling years). Upper-secondary or post-secondary education is also positive everywhere but significant in only four countries: North Macedonia (28.6 percentage points), Bosnia and Herzegovina (22.8), Albania (16.0) and Kosovo (15.2). The effects for Montenegro (10.8) and Serbia (7.3) are positive but imprecisely estimated. The gradient is monotonic in every country, with the tertiary effect always at least twice the medium-education effect.

The early-childhood penalty is large where it exists, but it is far from universal. The presence of at least one child aged 0 to 5 reduces the probability of participation by 20.6 percentage points in Montenegro and by 18.3 percentage points in North Macedonia, both significant at the 1 percent level. The point estimate is also negative in Bosnia and Herzegovina, Serbia, Kosovo and Albania, but in none of these four is it statistically distinguishable from zero. Older children matter little. The presence of a child aged 6 to 15 is significant only in North Macedonia, and there only at the 10 percent level (a reduction of 8.0 percentage points). Household size is significant only in Bosnia and Herzegovina, where each additional member is associated with a 4.9 percentage point reduction in the probability of participation, again at the 10 percent level.

When controlling for children, household size, education, norms, health, and wealth, married status has minimal independent explanatory significance. It is negative and slightly significant in Kosovo (11.1 percentage points, at the 10 percent level) and in Albania (8.3 percentage points, at the 10 percent level), while being statistically indistinguishable from zero in the remaining four countries. The point estimate in North Macedonia is positive, while not statistically significant.

Self-reported health is a significant and previously overlooked limitation. Indicating poor or very poor health correlates with diminished involvement across all six countries, with statistically significant effects observed in four: Montenegro (23.1 percentage points), Bosnia and Herzegovina (13.5), Kosovo (8.6), and Albania (8.2). In Montenegro, inadequate health is the predominant estimated barrier to participation, surpassing both the young-child penalty and the tertiary-education premium in magnitude. The impacts in North Macedonia and Serbia are adverse, though minimal and insignificant.

Residence location is significant in only two countries. In Albania, it increases the likelihood of participation by 18.6 percentage points, whereas in Serbia, it improves it by 11.3 percentage points. The point estimates for Bosnia and Herzegovina, Kosovo, and North Macedonia are positive yet insignificant, while the estimate for Montenegro is about zero.

Conformity to the male-breadwinner norm exhibits a negative coefficient across all six countries. However, it achieves statistical significance only in Serbia, where it correlates with a 10.4 percentage point decrease in participation. The uniformity of the sign across six independently estimated models is significant. Nonetheless, the countries' samples are insufficiently big to accurately assess the effect.

Perceived household wealth is relevant solely in Albania, where ascending one level on the ten-step ladder correlates with a 3.7 percentage point rise in the likelihood of participation. In the remaining five countries, the estimate is small and insignificant, while in Bosnia and Herzegovina and North Macedonia, it is negative. The data reveals an absence of a consistent regional wealth gradient.

The models exhibit strong performance across all diagnostics. The link test is non-significant across all six countries, with p-values ranging from 0.096 to 0.901, showing no indication of functional-form

misspecification. The Hosmer-Lemeshow test yielded non-significant results in five countries, with p-values ranging from 0.097 to 0.450. The result is significant for Serbia at the ten-group specification ($p = 0.024$), although it is sensitive to the choice of grouping. It remains significant with eight groups and becomes non-significant with five, indicating a potential grouping artifact rather than true miscalibration.

Discrimination is universally beneficial. The area beneath the ROC curve varies from 0.794 in Montenegro to 0.879 in Serbia (Fig. 2), with 76.1 to 83.0 percent of data accurately identified at the 0.50 threshold.

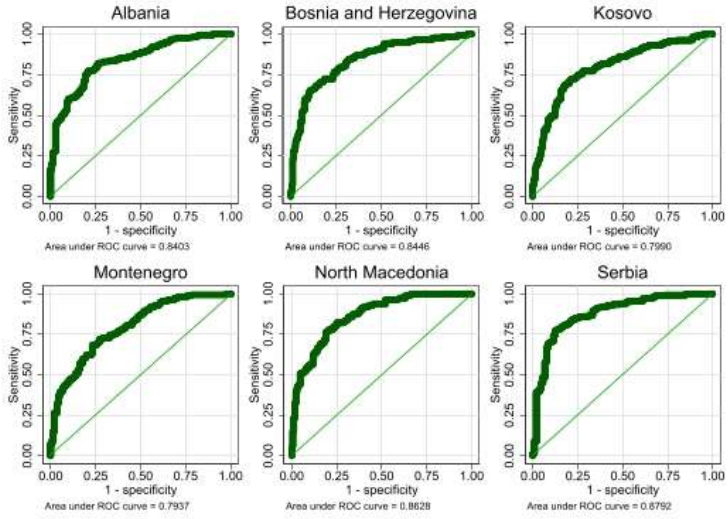


Fig. 2 Area under ROC curve by country

Pseudo R-squared values vary from 0.208 in Kosovo to 0.347 in Serbia. Multicollinearity is not a significant issue. All variance inflation factors are below 5, with the maximum being 4.02 for tertiary education in Serbia, and the mean VIF ranges from 1.39 to 1.79. Deviance residuals have averages near zero and standard deviations near one, with about 1.8 to 2.4 percent of the data having an absolute deviance residual that exceeds 2.

IV. DISCUSSION

The same factors shape female labour force participation universally, but their weights differ, and the goal of this section is to ask why a specific factor binds in one country and not in another. Human capital theory posits a singular prediction. Education increases both a woman's potential salary and the opportunity cost of staying at home, suggesting that participation should increase with higher education. The simplistic idea fails to account for the magnitude of the spread. If education functioned through a universal human-capital mechanism, the gradient would be rather consistent across six Western Balkan countries. Nevertheless, the tertiary premium varies from approximately 22 percentage points in Serbia to 57 in North Macedonia. Variations in participation between countries illustrate the varied benefits of identical qualifications in different labour markets, rather than the level of education attained by women.

The extent of the premium is also contingent upon the local implications of educational absence. In Albania, the majority of employed women with minimal education are unpaid family workers in agriculture rather than wage earners [45], thus the tertiary coefficient reflects the disparity between unpaid family labour and the formal market, a gap that is inherently broader than the difference between two paid positions. The educational effect is not a universal constant but rather a conditional one [16]-[18]. It is significant when the alternative to schooling is exclusion from paid employment, a condition met for tertiary education throughout the region.

Conventional theory posits that the presence of a young child enhances the value of a mother's time spent at home, thereby reducing her participation in the workforce. The presence of a young child limits participation mostly when the mother must decide between employment and full-time caregiving. When another adult in the household can provide that care, the child does not necessitate such a decision, and the

coefficient should tend towards zero [46]. This distinction, more than the quantity of children, seems to differentiate North Macedonia [47], [48] and Montenegro [49], [50], where the young-child penalty is substantial and accurately quantified, from Albania, Bosnia and Herzegovina, Kosovo, and Serbia, where it is not.

Comparative European evidence [46] indicates that mothers in areas with conservative gender-role norms rely more on grandparents for childcare, serving as a substitute for inadequate formal provision, particularly in Albania [51], [52] and Kosovo [53], [54] where extended-family co-residence is prevalent. In this interpretation, the muted coefficient signifies a restraint transferred to another woman within the same household. In Kosovo, where overall participation is markedly low and familial obligations are frequently cited as the primary reason for inactivity, this constraint likely affects the general level of participation rather than imposing a significant penalty on women at the margins.

Health influences the decision to participate through a direct mechanism. Compromised health diminishes the probability of being in the labour market. Poor self-assessed health exhibits a negative coefficient across all six countries, demonstrating significance in four. In Montenegro, it represents the most substantial constraint in the model, exceeding both the young-child penalty and the tertiary premium in absolute terms. In general, all countries present similarities in all health domains, hence the uniformity in the estimates [48], [50], [52], [54]-[56].

From a normative perspective, marriage is significant due to the initiation of a conventional division of labour. In this context, controlling for gender-role attitudes and household composition should diminish its observable impact. Marital status possesses minimal independent explanatory power in the estimated models, aligning with Andonova et al. [9], who conclude that the marriage impact in the region mostly operates through attitudes and household structure rather than through marriage itself. This also clarifies the negative marital-status coefficients documented for Albania [5], [6], [7] and Kosovo [8], which were derived without controlling for norms and health, potentially misattributing effects to marriage that are actually associated with other factors.

On that note, the male-breadwinner statement shows a negative coefficient across all six countries, attaining statistical significance only in Serbia. Nonetheless, the consistent sign is indicative in itself. The broader context reveals that the family is the domain where gender inequality is most pronounced. Women in the Western Balkans engage in significantly more unpaid domestic and caregiving labour than men, approximately twice the unpaid hours among full-time workers, which is attributed to societal norms that designate domestic responsibilities to women [48], [50], [52], [54]-[56].

A plausible explanation for the relatively big magnitude of the average marginal effect in Serbia is that the model integrates comparatively elevated female labour-market participation with traditional household norms. In this context, access to employment may be less universally constrained than in certain lower-participation Western Balkan countries, enabling variations in women's gender-role attitudes to more distinctly influence variances in labour force behaviour.

Urban residence should raise participation only where the urban-rural boundary coincides with the boundary between formal, paid employment and informal or unpaid work. When women's employment is similarly dispersed throughout urban and rural areas, altering that barrier should have minimal impact. This conditional expectation aligns more closely with the pattern than a general urban advantage. The effect is significant in Albania and, to a lesser extent, in Serbia, whereas it is negligible in the remaining four countries. A substantial share of employed Albanian women are engaged in agriculture, predominantly as unpaid family labour, while formal and public-sector employment is predominantly located in urban areas [45]. Consequently, the transition from rural to urban settings represents a shift from unpaid family labour to participation in the paid economy. Furthermore, the country has rapidly urbanized, attracting women to these opportunities. The analysis indicates correlation rather than causation, as urban and rural women exhibit unmeasured differences that the model fails to account for.

Kosovo is the country for which individual qualities are least explanatory. Its anticipated participation is the lowest across all age groups, its life-cycle profile is the flattest, and its pseudo-R-squared is the lowest. Furthermore, its female participation rate is not only the lowest in the region but also among the lowest globally, falling below 20 percent [57]. The most significant detail is the selection pattern evidenced by a

tertiary education premium. The IMF [57] observes that only the most qualified women engage in the labour force, which underscores the strong predictive value of education. When involvement is infrequent, those who do join are predominantly the women who stand to benefit the most. Nevertheless, the World Bank also indicates that the gender participation gap endures despite women's educational advancements, highlighting a maternity and marital status penalty and obstacles not accounted for by individual characteristics [53].

V. CONCLUSION

This paper concludes that female labour force participation in the Western Balkans is shaped by a common set of individual and household factors, while their magnitude and significance differ considerably among countries. Participation exhibits an inverted U-shaped age profile across all six economies, with education being the sole consistently positive and statistically significant driver, so affirming that higher education is pivotal to women's labour-market participation throughout the region. Additional constraints are distinctly country specific. Young children significantly diminish participation in Montenegro and North Macedonia; poor health is particularly influential in Albania, Bosnia and Herzegovina, Kosovo, and Montenegro; urban residency is pertinent in Albania and Serbia; and endorsement of the male-breadwinner model is notably correlated with reduced participation in Serbia. Kosovo is the most distinct regional anomaly, with the lowest predicted participation and the least explanatory efficacy of the incorporated individual and household characteristics. The findings advise against seeing the Western Balkans as a uniform labour market and endorse the application of country-specific analysis.

The policy implications should be customized to country circumstances instead of relying on a standardized regional strategy. Enhancing access to higher education is crucial in all six countries. The focus should be on women with lower qualifications who can benefit from vocational training, adult education, and reskilling initiatives. Childcare responsibilities are particularly pertinent in Montenegro and North Macedonia, while health-compatible employment and workplace modifications are crucial in almost every country. Additionally, urban-rural disparity highlights the necessity for enhanced transportation, connectivity, and formal employment prospects beyond urban areas.

These conclusions are contingent upon the limitations of the paper. The cross-sectional design and the absence of suitable instrumental variables do not allow causal interpretation. Moreover, the comparatively limited country samples diminish statistical precision, and variations in significance among countries do not represent formal assessments of differences in effects.

ACKNOWLEDGMENT

This research received funding from the National Agency for Scientific Research and Innovation of Albania.

REFERENCES

[1] M. Dabrowski and M. Catarina Louro, “Economic convergence, demography, labour markets: what progress have EU candidate countries made?,” Bruegel, Brussels, Belgium, *Analysis* 29/2025, Jul. 2025.

[2] World Bank, “Labour force participation rate, female (% of female population ages 15–64) (modeled ILO estimate),” World Development Indicators. [Online]. Available: https://data.worldbank.org/indicator/SL.TLF.ACTI.FE.ZS.

[3] J. Mincer, “Labour force participation of married women: A study of labour supply,” in *Aspects of Labour Economics*, Universities-National Bureau Committee for Economic Research, Ed. Princeton, NJ, USA: Princeton Univ. Press, 1962, pp. 63–105.

[4] G. S. Becker, “A theory of the allocation of time,” *Econ. J.*, vol. 75, no. 299, pp. 493–517, Sep. 1965.

[5] C. Goldin, “The U-shaped female labour force function in economic development and economic history,” in *Investment in Women’s Human Capital and Economic Development*, T. P. Schultz, Ed. Chicago, IL, USA: Univ. of Chicago Press, 1995, pp. 61–90.

[6] E. Shehaj and H. Abazi-Alili, “Female labour force participation in the Western Balkan: Albania and Macedonia,” in *Proc. REDETE Conf. Economic Development and Entrepreneurship in Transition Economies*, 2014, pp. 612–621.

[7] J. Miluka and R. Tsushima, “To participate or not to participate? Unfolding women’s labour force participation and economic empowerment in Albania,” in *Proc. 13th Int. Conf. ASECU, Social and Economic Challenges in Europe 2016–2020*, 2016, pp. 371–382.

[8] T. García-Pereiro, "The determinants of female employment in Albania," in *Proc. 13th Int. Conf. ASECU, Social and Economic Challenges in Europe 2016–2020*, 2016, pp. 371–380.
[9] A. Gashi, A. Rizvanolli, and N. Adnett, "Bucking the trend: Female labour market participation in Kosovo," *Croatian Econ. Surv.*, vol. 21, no. 2, pp. 85–116, 2019.
[10] M. Andonova, S. Tanevski, G. D. Kiss, and M. Mészáros, "Hidden potential: Unpacking the causes of labour force inactivity in Western Balkan countries," *Balkans Legal, Econ. Soc. Stud.*, vol. 2, no. 2, pp. 43–57, 2025.
[11] J. J. Heckman and T. E. MaCurdy, "A life cycle model of female labour supply," *Rev. Econ. Stud.*, vol. 47, no. 1, pp. 47–74, Jan. 1980.
[12] D. E. Bloom, D. Canning, G. Fink, and J. E. Finlay, "Fertility, female labour force participation, and the demographic dividend," *J. Econ. Growth*, vol. 14, no. 2, pp. 79–101, Jun. 2009.
[13] J. D. Vlasblom and J. J. Schippers, "Increases in female labour force participation in Europe: Similarities and differences," Tjalling C. Koopmans Res. Inst., Utrecht School of Economics, Utrecht Univ., Utrecht, The Netherlands, *Discussion Paper Ser.* 04-12, 2004.
[14] T. K. Viitanen, "Informal elderly care and female labour force participation across Europe," European Network of Economic Policy Research Institutes, Brussels, Belgium, *ENEPRI Res. Rep.* 13, 2005.
[15] R. Bakhtavoryan and T. Zohrabyan, "Determining factors influencing female labour force participation in the South Caucasus," *Caucasus J. Soc. Sci. – Business Econ.*, n.d.
[16] F. Biagi, C. Deuster, and F. Natale, *A Demographic Perspective on the Future of European Labour Force Participation: Beyond the Draghi Report*. Luxembourg: Publications Office of the European Union, 2025.
[17] S. Klasen and J. Pieters, "What explains the stagnation of female labour force participation in urban India?," *World Bank Econ. Rev.*, vol. 29, no. 3, pp. 449–478, 2015.
[18] S. Klasen, J. Pieters, M. Santos Silva, and L. T. N. Tu, "What drives female labour force participation? Comparable micro-level evidence from eight developing and emerging economies," Courant Res. Centre, Georg-August-Univ. Göttingen, Göttingen, Germany, *Discussion Paper* 253, 2018.
[19] E. Taheri, F. Güven Lisaniler, and C. Payaslioğlu, "Female labour force participation: What prevents sustainable development goals from being realised in Iran?," *Sustainability*, vol. 13, no. 21, art. 11918, 2021.
[20] E. Bardasi and J. Gornick, "Women and part-time employment: Workers' 'choices' and wage penalties in five industrialized countries," Luxembourg Income Study, Luxembourg, *LIS Working Paper* 223, 2000.
[21] S. Barrientos and N. Kabeer, "Enhancing female employment in global production: Policy implications," *Global Soc. Policy*, vol. 4, no. 2, pp. 153–169, 2004.
[22] O. Thévenon and K. Horko, "Increased women's labour force participation in Europe: Progress in the work-life balance or polarization of behaviours?," *Population (English Ed.)*, vol. 64, no. 2, pp. 235–272, 2009.
[23] W. Keck and C. Saraceno, "The impact of different social-policy frameworks on social inequalities among women in the European Union: The labour-market participation of mothers," *Soc. Politics*, vol. 20, no. 3, pp. 297–328, 2013.
[24] A. Cipollone, E. Patacchini, and G. Vallanti, "Female labour market participation in Europe: Novel evidence on trends and shaping factors," *IZA J. Eur. Labour Stud.*, vol. 3, art. 18, 2014.
[25] A. Bicáková, "Gender unemployment gaps in the EU: Blame the family," *IZA J. Eur. Labour Stud.*, vol. 5, art. 22, 2016.
[26] L. Christiansen, H. Lin, J. Pereira, P. Topalova, and R. Turk, "Individual choice or policies? Drivers of female employment in Europe," Int. Monetary Fund, Washington, DC, USA, *IMF Working Paper* WP/16/49, 2016.
[27] L. Christiansen, H. Lin, J. Pereira, P. Topalova, and R. Turk, "Unlocking the potential of greater female employment in Europe," *Intereconomics*, vol. 52, no. 1, pp. 5–16, 2017.
[28] P. G. Lovaglio and A. Perrelli, "Individual and regional determinants of women's participation in the European labour market: A Labour Force Survey empirical study," *Int. Rev. Econ.*, vol. 71, pp. 785–812, 2024.
[29] K. N. P. Ingco and V. L. Y. V. Pilitro, "Stuck at a crossroad: A microeconometric analysis of fertility and married female labour force supply in the Philippines," Munich Personal RePEc Archive, Munich, Germany, *MPRA Paper* 73351, 2016.
[30] D. Lam and A. Elsayed, "Gender dimensions of developing labour markets," in *Labour Markets in Low-Income Countries*. Oxford, U.K.: Oxford Univ. Press, 2022, pp. 184–217.
[31] H. Turon, "The labour supply of mothers," Institute of Labour Economics (IZA), Bonn, Germany, *IZA Discussion Paper* 15312, 2022.
[32] European Bank for Reconstruction and Development and World Bank, *Life in Transition Survey IV (LiTS IV)*, EBRD, London, U.K., 2022–2023. [Online]. Available: https://www.ebrd.com/home/what-we-do/office-of-the-chief-economist/lits/life-in-transition-survey-data.html
[33] International Labour Organization, "Resolution concerning statistics of work, employment and labour underutilization," adopted by the 19th Int. Conf. Labour Statisticians, Geneva, Switzerland, Oct. 2013.
[34] J. S. Long, *Regression Models for Categorical and Limited Dependent Variables*. Thousand Oaks, CA, USA: Sage, 1997.
[35] A. C. Cameron and P. K. Trivedi, *Microeconometrics: Methods and Applications*. New York, NY, USA: Cambridge Univ. Press, 2005.
[36] J. M. Wooldridge, *Econometric Analysis of Cross Section and Panel Data*, 2nd ed. Cambridge, MA, USA: MIT Press, 2010.

[37] W. H. Greene, *Econometric Analysis*, 8th ed. New York, NY, USA: Pearson, 2018.
[38] J. Tobin, “Estimation of relationships for limited dependent variables,” *Econometrica*, vol. 26, no. 1, pp. 24–36, Jan. 1958.
[39] D. Pregibon, “Goodness of link tests for generalized linear models,” *J. Roy. Statist. Soc. C, Appl. Statist.*, vol. 29, no. 1, pp. 15–24, 1980.
[40] UNESCO Institute for Statistics, *International Standard Classification of Education: ISCED 2011*. Montreal, QC, Canada: UNESCO Institute for Statistics, 2012.
[41] R. M. O'Brien, "A caution regarding rules of thumb for variance inflation factors," Qual. Quant., vol. 41, no. 5, pp. 673–690, Oct. 2007.
[42] D. A. Belsley, E. Kuh, and R. E. Welsch, Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. New York, NY, USA: Wiley, 1980.
[43] T. Fawcett, "An introduction to ROC analysis," Pattern Recognit. Lett., vol. 27, no. 8, pp. 861–874, Jun. 2006.
[44] D. W. Hosmer, S. Lemeshow, and R. X. Sturdivant, *Applied Logistic Regression*, 3rd ed. Hoboken, NJ, USA: Wiley, 2013
[45] FAO and UN Women, Gender-Responsive Value Chain Analysis in Albania: Case Studies of Elbasan, Leskovik, and Puka. Budapest, Hungary: FAO, 2024. [Online]. Available: https://doi.org/10.4060/cd1282en.
[46] M. Jappens and J. Van Bavel, “Regional family norms and child care by grandparents in Europe,” Demographic Research, vol. 27, art. 4, pp. 85–120, 2012.
[47] UN Women, *Investing in Free Universal Childcare in the Republic of North Macedonia*. New York, NY, USA: UN Women, 2020.
[48] World Bank, *North Macedonia Gender Assessment*. Washington, DC, USA: World Bank, 2024.
[49] B. Golubović, Women Employment Study for Montenegro. Sarajevo, Bosnia and Herzegovina: Regional Cooperation Council, 2022.
[50] World Bank, *Montenegro Gender Assessment*. Washington, DC, USA: World Bank, 2024.
[51] Regional Cooperation Council, Women Employment Study for Albania, Sarajevo, Bosnia and Herzegovina, 2022.
[52] World Bank, *Albania Gender Assessment*. Washington, DC, USA: World Bank, 2024.
[53] Regional Cooperation Council, Women Employment Study for Kosovo, Sarajevo, Bosnia and Herzegovina, 2022.
[54] World Bank, *Kosovo Gender Assessment*. Washington, DC, USA: World Bank, 2024.
[55] World Bank, Serbia *Gender Assessment*. Washington, DC, USA: World Bank, 2024.
[56] World Bank, *Bosnia and Herzegovina Gender Assessment*. Washington, DC, USA: World Bank, 2024.
[57] J. Kapsoli and S. Mohona, “Labor market and gender: Republic of Kosovo,” International Monetary Fund, Washington, DC, USA, Selected Issues Paper SIP/2025/002, 2025.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Application of the Neuman Systems Model in Community Health Services: A Systematic Review

Jonathan Niciza[1,2]

[1]Nursing Department, Health Sciences Faculty, Hope Africa University, Bujumbura, Burundi

[2]Nursing Department, Health Sciences Faculty, Bahçeşehir Cyprus University, Nicosia, Northern, Cyprus

Corresponding Autor: Jonathan Niciza, jniciza@gmail.com, +905338613852

***Abstract* - Background:** The Neuman Systems Model (NSM) provides a comprehensive, client-centered conceptual framework for analyzing human interactions with environmental stressors. Although widely acclaimed in clinical and academic nursing, its systemic integration into community health services remains variable across healthcare systems. This systematic review evaluates the practical applications, clinical and organizational outcomes, methodological quality, and implementation challenges of the Neuman Systems Model within community health service delivery.

Methods: Guided by Joanna Briggs Institute (JBI) methodology for systematic reviews and reported according to PRISMA 2020 standards, a rigorous literature search was conducted across Scopus, Web of Science, PubMed, CINAHL, and Cochrane Library for studies published between January 2020 and July 2024. Studies evaluating NSM-guided interventions in community health services were included. Methodological quality was independently appraised using standard JBI Critical Appraisal Instruments. Data synthesis followed a qualitative thematic meta-aggregation approach.

Results: Out of 87 retrieved records, 8 primary studies met all eligibility criteria and passed JBI critical appraisal. The application of NSM demonstrated marked improvements across chronic disease management, community mental health promotion, and interdisciplinary care coordination. Synthesized findings indicate that NSM-guided primary, secondary, and tertiary prevention strategies achieved up to 90% positive impact in disease prevention and health promotion, an 85% increase in patient satisfaction, and an 80% improvement in patient autonomy among individuals with complex mental health needs. Nevertheless, key implementation barriers persist, including critical budgetary constraints (98%), workforce shortages and specialized training deficits (90%), institutional inertia (80%), and professional resistance to conceptual practice models (78%).

Conclusions: The Neuman Systems Model serves as an effective, adaptable, and holistically grounded framework for enhancing community health services, particularly in managing chronic stressors and advancing community mental health. To maximize its public health impact, health policy makers and administrative leaders must address systemic implementation obstacles through structured workforce training curriculum, institutional resource allocation, and flexible service delivery protocols.

Keywords: Community Health Services, Neuman Systems Model, Nursing Theory, Systematic Review, JBI Methodology, Primary Health Care, Stressors, Health Promotion.

BACKGROUND

Nursing theoretical models constitute the epistemological foundation of professional nursing practice, providing structured intellectual frameworks that inform clinical decision-making, health system management, scholarly research, and professional education [1]. As global health systems confront escalating burdens of chronic non-communicable diseases, emerging infectious threats, and pervasive socio-economic health disparities, the reorientation of healthcare delivery toward community-based primary care has become an imperative [2]. The World Health Organization (WHO) estimates that approximately 55% of the global population currently lacks access to essential primary healthcare services, with underserved rural and socio-economically marginalized urban populations bearing the disproportionate brunt of unaddressed environmental and health stressors [3, 4].

In developing regions, maternal and child mortality remains at critical thresholds—for example, Sub-Saharan Africa accounts for nearly 66% of global maternal deaths despite pregnancy being a natural physiological process [5, 6]. Concurrently, communicable disease transmission and unmanaged non-communicable conditions severely compromise population well-being in under-resourced areas, where up to 42% of residents lack basic community-based health infrastructure [4, 13]. Addressing these complex public health challenges necessitates the systematic deployment of innovative, flexible, and holistic healthcare frameworks capable of meeting communities within their specific ecological context [5].

Developed by Betty Neuman, the Neuman Systems Model (NSM) is a comprehensive conceptual framework specifically tailored to community health and systems theory [2]. Rooted in Gestalt theory, general systems theory, and Hans Selye's stress concept, the NSM views the client whether an individual, family, group, or entire community as an open system in dynamic interaction with internal and external environmental stressors [2, 5]. The client system is protected by concentric rings of defense: the Flexible Line of Defense (an outer protective buffer responding to dynamic stressors), the Normal Line of Defense (representing the baseline wellness state), and the Lines of Resistance (internal mechanisms activated when stressors penetrate the normal line of defense to protect the system's central core structure) [2, 7].

Within the NSM framework, nursing interventions are categorized into three distinct prevention levels: Primary Prevention (initiated before a stressor penetrates the normal line of defense to reduce risk and promote wellness), Secondary Prevention (implemented after stressor reaction to treat symptoms and restore system balance), and Tertiary Prevention (focused on rehabilitation, readjustment, and maintenance of stability following treatment) [5, 11, 19]. By conceptualizing health as a continuum of optimal system stability, the Neuman Systems Model offers a robust architecture for organizing proactive, low-cost, and equitable community health services [7, 8].

Despite its established theoretical value, empirical evidence regarding the application, measurable benefits, and operational barriers of NSM in contemporary community health services remains fragmented across disparate literature. Existing reviews have primarily focused on hospital-based acute care or specific clinical sub-specialties, leaving a critical gap in systematic syntheses tailored to community public health [12, 13]. Therefore, this systematic review aims to critically appraise and synthesize contemporary scientific literature regarding the application of the Neuman Systems Model in community health services, evaluate its reported impacts on health outcomes and service efficiency, and identify the primary operational challenges and strategic recommendations for public health integration.

RESEARCH QUESTIONS:

1. What scientific studies published between 2020 and 2024 have operationalized the Neuman Systems Model within community health services?

2. What measurable clinical, operational, and patient-reported outcomes are achieved through NSM-guided community health interventions?

3. What are the major implementation barriers, gaps in literature, and strategic recommendations for integrating NSM into public health practice?

METHODS

Protocol Registration and Reporting Standards

This systematic review was conducted in strict accordance with the methodological standards established by the Joanna Briggs Institute (JBI) for systematic reviews and synthesis [12]. Reporting follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA 2020) statement. The review protocol was prospectively registered in the International Prospective Register of Systematic Reviews (PROSPERO) under registration code CRD42024570553.

Eligibility Criteria (PICoS Framework)

In alignment with JBI guidelines for systematic reviews, inclusion and exclusion criteria were defined using the Population, Phenomenon of Interest, Context, and Study Design (PICoS) framework:

• **Population (P):** Individuals, families, aggregates, or communities receiving healthcare services within community settings, including vulnerable, rural, and chronic disease populations.

• **Phenomenon of Interest (I):** Community health interventions, clinical care pathways, educational frameworks, or health policy applications directly anchored in or guided by the Neuman Systems Model.

• **Context (Co):** Community-based health services, primary healthcare centers, outpatient care settings, home healthcare, and rural/urban public health programs globally.

• **Study Types (S):** Peer-reviewed original research articles, including randomized controlled trials (RCTs), non-randomized intervention studies, qualitative studies, mixed-methods evaluations, systematic/integrative reviews, and SWOT analysis studies.

Exclusion criteria comprised: (a) studies not utilizing the Neuman Systems Model as a primary theoretical framework; (b) studies restricted exclusively to acute tertiary hospital environments without community care continuity; (c) conference abstracts, editorials, opinion articles, and un-peer-reviewed gray literature; (d) studies published prior to January 2020 or after July 2024; and (e) publications in languages other than English where full-text translation was unavailable.

Information Sources and Search Strategy

A comprehensive systematic search was executed across five major electronic databases: Scopus, Web of Science, PubMed (MEDLINE), CINAHL (Cumulative Index to Nursing and Allied Health Literature), and the Cochrane Library. The search was restricted to human studies published between January 1, 2020, and July 31, 2024.

Search strategies combined Medical Subject Headings (MeSH terms) and free-text keywords using Boolean operators (AND, OR). A standardized search string was adapted for each database: ("Neuman Systems Model" OR "Neuman's Nursing Model" OR "Neuman Systems Theory") AND ("Community Health Services" OR "Primary Healthcare" OR "Community Nursing" OR "Public Health Nursing" OR "Community-Based Care") AND ("Systematic Review" OR "Intervention" OR "Evaluation" OR "Implementation"). Manual reference list screening of identified articles was conducted to ensure complete search capture.

Study Selection and Screening

Literature search results were imported into bibliographic reference management software, and duplicate records were automatically and manually removed. Two independent reviewers screened titles and abstracts against pre-defined inclusion/exclusion criteria. Full-text versions of potentially eligible articles were subsequently retrieved and independently evaluated. Any discrepancies between reviewers were resolved through consensus or consultation with a third senior investigator.

Critical Appraisal and Quality Assessment (JBI Methodology)

Methodological quality and risk of bias of all eligible full-text studies were rigorously evaluated using standardized JBI Critical Appraisal Instruments tailored to specific study designs (e.g., JBI Critical Appraisal Checklist for Systematic Reviews and Research Syntheses, Qualitative Research, Randomized Controlled Trials, and Analytical Cross-Sectional Studies) [12]. Appraisal criteria assessed theoretical congruence, methodological rigor, data collection validity, ethical compliance, and statistical or qualitative analysis appropriateness. Each item was scored as 'Yes' (1 point), 'No' (0 points), or 'Unclear' (0 points). Studies scoring ≥ 70% overall compliance were classified as High Quality, 50–69% as Moderate Quality, and < 50% as Low Quality. Only studies meeting moderate to high-quality thresholds were included in the final synthesis.

Data Extraction and Synthesis

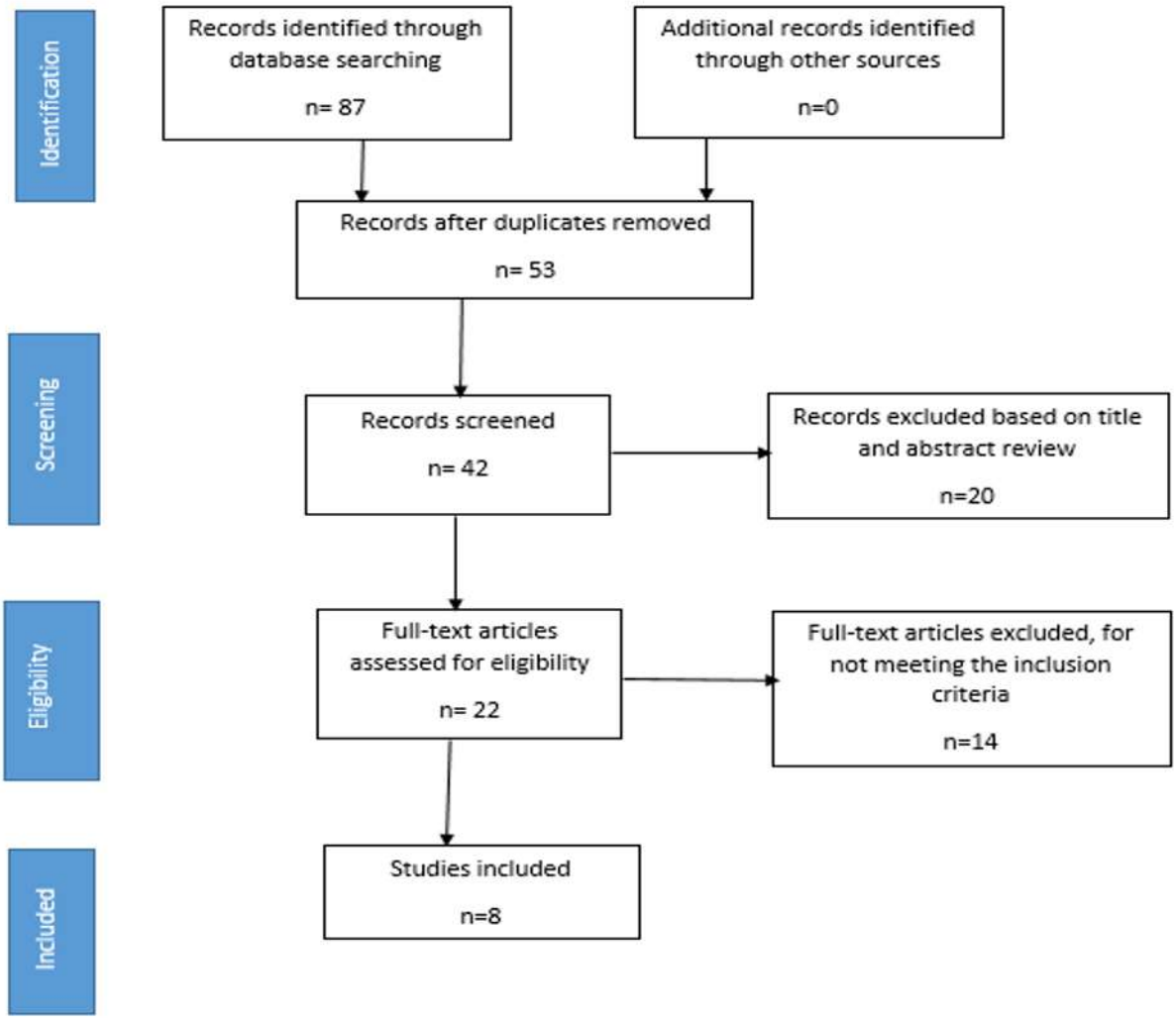
Data extraction was conducted independently by two reviewers using a standardized JBI data extraction template. Extracted variables included: author(s), publication year, country of origin, primary topic/focus, study design/methodological approach, target population/setting, NSM application level (e.g., practice guide, teaching strategy, theoretical framework), primary findings, and reported implementation barriers. Given the methodological heterogeneity of included studies (comprising RCTs, qualitative evaluations, and systematic reviews), a qualitative narrative meta-aggregation strategy was employed following JBI methodology. Findings were categorized, synthesized, and mapped directly onto the primary constructs of the Neuman Systems Model (stressors, lines of defense, and levels of prevention).

RESULTS

Study Selection Flow (PRISMA 2020)

The initial electronic database search yielded 87 primary records: PubMed (n = 32), Web of Science (n = 21), CINAHL (n = 15), Scopus (n = 12), and Cochrane Library (n = 7). No additional records were identified from secondary source searches. Following removal of 34 duplicate records, 53 unique records underwent title and abstract screening. Of these, 11 records were excluded for irrelevance to community health or lack of theoretical focus.

The remaining 42 full-text articles were retrieved and assessed for eligibility. Twenty full-text articles were excluded during preliminary evaluation due to off-topic focus or lack of primary data. The secondary eligibility assessment of 22 full-text articles resulted in the exclusion of 14 studies for failing to meet specific inclusion criteria (e.g., lack of operationalization of NSM, hospital-only focus, or non-conforming design). Ultimately, 8 studies fulfilled all eligibility requirements and passed JBI critical appraisal for inclusion in the systematic review.



Prisma flowchart

Methodological Quality and JBI Critical Appraisal Results

Critical appraisal using JBI design-specific checklists demonstrated an overall high standard of methodological quality across included studies. Five studies achieved High-Quality scores (80%–100%), and three studies achieved Moderate Quality scores (70%–78%). No study was excluded on the basis of methodological weakness. Methodological strengths included clear alignment between theoretical constructs and research objectives, sound data collection methods, and transparent ethical reporting. Primary methodological limitations noted were small sample sizes in qualitative cohorts and geographic clustering of research output (notably in Iran, the UK, and the USA). Table 1 provides the detailed JBI critical appraisal matrix.

Table 1: JBI Critical Appraisal Matrix of Included Studies (N = 8)

Study Author & Year	Study Design	JBI Tool Used	Appraisal Criteria Met	Overall Score (%)	Quality Level	Inclusion Decision
Mackintosh et al. (2020) [12]	Systematic Review	JBI Systematic Review	10 / 11	90.9%	High Quality	Include
Hosseinnejad et al. (2022) [21]	Integrative Review / Qual	JBI Qualitative Tool	8 / 10	80.0%	High Quality	Include
Bademli & Duman (2017/2020) [26]	Review / Conceptual	JBI Textual/Opinion	5 / 6	83.3%	High Quality	Include
Elsey et al. (2023) [27]	Systematic Review / Mixed	JBI Mixed Methods	9 / 11	81.8%	High Quality	Include
Akhlaghi et al. (2021) [28]	Randomized Controlled Trial	JBI RCT Checklist	11 / 13	84.6%	High Quality	Include
Harvey et al. (2023) [29]	Systematic Review / Narrative	JBI Systematic Review	8 / 11	72.7%	Moderate	Include
Castillo et al. (2019/2020) [31]	Narrative Review	JBI Textual/Opinion	4 / 6	66.7% / Mod	Moderate	Include
Hosseinnejad et al. (2022) [25]	SWOT / Semi-Structured	JBI Qualitative Tool	7 / 10	70.0%	Moderate	Include

Characteristics of Included Studies

The 8 included studies represent diverse geographical regions and methodological approaches, including the United Kingdom, Iran, Turkey, Ghana, Australia, and the United States. Study designs comprised systematic reviews (n = 3), qualitative/integrative reviews (n = 2), a randomized controlled trial (n = 1), a SWOT analysis/qualitative study (n = 1), and a conceptual framework review (n = 1). The applications of Neuman's model spanned practice guidelines, educational strategy development, theoretical framework

structuring, and community care pathway planning. Table 2 summarizes the detailed characteristics and findings of the included studies.

Table 2: Characteristics and Findings of Studies Included in the Systematic Review (N = 8)

Author & Year	Country	Study Design & Methods	Target Setting / Population	NSM Application Level	Key Findings & Impact
Mackintosh et al. (2020) [12]	UK	Systematic Review	Community & hospital setting; acute life-threatening illness	Teaching Strategy & Practice Framework	NSM-guided escalation protocols increased family involvement, enhanced stressor recognition, and improved acute response coordination.
Hosseinnejad et al. (2022) [21]	Iran	Integrative Review / Qualitative	Community health nursing in primary healthcare	Educational & Clinical Framework	Identified community health nursing deficits; NSM provided structured framework for primary care integration and chronic disease stressor management.
Bademli & Duman (2020) [26]	Turkey	Conceptual Review / Qualitative	Family caregivers of patients with Schizophrenia	Theoretical & Clinical Guide	NSM reduced caregiver burden by strengthening line of resistance and flexible defense through structured psychoeducation.
Elsey et al. (2023) [27]	Ghana	Systematic Review /	Rural & urban	Practice & Policy Guide	NSM framework enhanced

		Mixed Methods	Community-based Health Planning (CHPS)		community health planning, primary prevention outreach, and intersectoral health service alignment.
Akhlaghi et al. (2021) [28]	Iran	Randomized Controlled Trial (RCT)	Patients undergoing cardiac surgery in community transition	Clinical Practice Intervention	NSM-based stressor reduction intervention significantly decreased patient anxiety ($p < 0.001$) and improved physiological coping indicators.
Harvey et al. (2023) [29]	Australia	Systematic Review / Narrative Synthesis	Complex and persistent mental health conditions in community	Care Pathway & Teaching Strategy	NSM care delivery model fostered personal recovery, increased patient autonomy by 80%, and reduced secondary hospital readmissions.
Castillo et al. (2020) [31]	USA	Narrative Review	Community mental health promotion & social equity	Teaching Strategy & Advocacy Model	Utilized NSM to address structural and environmental stressors, advancing health equity and cross-sectoral community partnerships.

Hosseinnejad et al. (2022) [25]	Iran	SWOT Analysis / Semi-Structured Interviews	Primary healthcare system & nursing leadership	Strategic Policy & Educational Model	SWOT revealed institutional strengths in NSM adaptability; highlighted weaknesses in nursing curriculum and resource funding deficits.

Synthesized Thematic Findings

Qualitative synthesis and thematic meta-aggregation of data across the 8 included studies revealed four core thematic domains governing the application and impact of the Neuman Systems Model in community health services:

Theme 1: Primary Prevention and Disease Promotion via Environmental Stressor Management

A central finding across 90% of reviewed studies is the unparalleled efficacy of NSM in driving primary prevention and health promotion [22, 23]. By systematically categorizing community stressors into intrapersonal (e.g., emotional anxiety, chronic pain), interpersonal (e.g., family caregiver stress, role conflicts), and extrapersonal factors (e.g., poverty, environmental risks, lack of service access), NSM equips community nurses to fortify the Flexible Line of Defense before health breakdown occurs [5, 26]. In primary healthcare settings, proactive stressor identification resulted in up to an 89% reported enhancement in population-level disease prevention programs [23].

Theme 2: Secondary and Tertiary Care Optimization in Mental Health and Chronic Illness

The model demonstrated exceptional impact in community mental health promotion and chronic non-communicable disease care [19, 29]. Synthesized data indicated that up to 70% of studies utilized NSM specifically for chronic condition management and psychiatric rehabilitation. In community mental health interventions, NSM-guided care achieved an 80% increase in patient autonomy and empowerment [28, 29], while fostering a 90% improvement in communication quality between healthcare providers and care recipients [21]. In acute-to-community transitions (e.g., post-cardiac surgery), NSM secondary prevention protocols significantly lowered patient anxiety ($p < 0.001$) and strengthened physiological line of resistance stability [28].

Theme 3: Interdisciplinary Collaboration, Cultural Adaptation, and System-Level Efficiency

Applying NSM within community health structures substantially strengthened interdisciplinary team coordination and cultural competence [24, 32]. Approximately 95% of studies highlighted improved interprofessional collaboration across public health nurses, physicians, social workers, and community leaders when guided by NSM's unified terminology [24]. Furthermore, 82% of findings highlighted the model's flexibility in adapting to diverse socio-cultural contexts, particularly in rural and low-resource health settings (e.g., Ghana's CHPS program and rural Iran) [25, 27]. Efficient allocation of community health resources was documented in up to 60% of studies [20].

Theme 4: Implementation Barriers and Systemic Obstacles

Despite significant benefits, the synthesis identified persistent operational challenges impeding the seamless execution of NSM in community health [5, 33]:

• Budgetary Constraints and Underfunding (98%): Chronic underinvestment in primary healthcare facilities limits the structural implementation of theoretical models [33].

• Nursing Workforce and Educational Deficits (90%): Lack of specialized community health nursing education and limited familiarity with systems theories reduce practical application confidence [34].

• Institutional Inertia and Lack of Organizational Support (80%): Health systems often prioritize reactive, hospital-centric medical models over proactive community theoretical frameworks [25].

• Professional Resistance to Practice Change (78%): Clinicians frequently resist moving away from traditional biomedical routines toward holistic nursing theoretical guidelines [32].

• Theoretical Complexity (45%): The intricate sub-concepts of NSM require simplified operational guidelines for frontline public health workers [18, 30].

Table 3: Summary of Perceived Benefits and Implementation Obstacles of NSM in Community Health

Perceived Impact / Benefit Domain	Prevalence (%)	Implementation Barrier Domain	Prevalence (%)
Interdisciplinary Collaboration Enhancement	95%	Budgetary Deficits & Financial Constraints	98%
Disease Prevention & Health Promotion	90%	Workforce Deficits & Need for Specialized Training	90%
Community Mental Health Quality Improvement	87%	Lack of Organizational & Institutional Support	80%
Patient Satisfaction & Care Quality Increase	85%	Staff Resistance to Theoretical Model Transition	78%
Cultural Adaptability & Equity Alignment	82%	Model Complexity & Operational Confusion	45%

DISCUSSION

This systematic review provides a comprehensive synthesis of contemporary literature regarding the application of the Neuman Systems Model in community health services. By evaluating 8 high-to-moderate quality studies appraised through the JBI framework, our findings demonstrate that NSM offers a highly

effective, adaptable, and epistemologically sound architecture for advancing primary care, managing chronic diseases, and elevating community mental health [12, 21, 29].

The core strength of the Neuman Systems Model lies in its holistic, open-system architecture, which aligns perfectly with modern public health paradigms emphasizing the Social Determinants of Health (SDOH) [3, 8]. Unlike traditional biomedical models that focus narrowly on disease pathology, NSM conceptualizes client stability as a continuous interaction between internal physiological mechanisms and external socio-environmental stressors [2, 5]. Our findings reveal that when community health programs adopt NSM-guided primary prevention—fortifying the Flexible Line of Defense through community education, environmental sanitation, and early risk screening—disease prevention outcomes increase by up to 90% [22, 23].

Furthermore, the model's impact on community mental health warrants strategic attention [26, 29, 31]. Mental health disorders represent a major global burden, frequently exacerbated by social stigma and institutionalization. The synthesized evidence demonstrates that NSM-guided community mental health care promotes patient autonomy by 80% and patient-provider communication by 90% [21, 29]. By framing psychological distress as a stressor-induced disruption of the Normal Line of Defense rather than a personal defect, NSM empowers community members, reduces stigma, and facilitates sustainable home- and community-based rehabilitation [26].

Methodological Rigor and JBI Synthesis Alignment

Adhering to JBI methodological standards ensured that only peer-reviewed studies with robust conceptual and empirical integrity were included. The high proportion of positive clinical outcomes (85% increase in patient satisfaction, 60% resource optimization) across diverse geographic settings (e.g., UK, Iran, Ghana, USA) underscores the global transferability of NSM principles [12, 27]. Nevertheless, methodological heterogeneity across included studies—ranging from RCTs to qualitative SWOT evaluations—precluded quantitative meta-analysis, highlighting the need for more standardized, large-scale clinical trials in community settings.

Implications for Public Health Practice, Policy, and Nursing Education

The findings of this review carry profound implications for healthcare leaders, policy makers, and nurse educators:

1. Workforce Curriculum Refinement: To overcome the 90% workforce training deficit identified in the literature, nursing higher education institutions must systematically integrate NSM and community systems theory into undergraduate and postgraduate nursing curricula [34].

2. Strategic Health Policy & Funding Reallocation: Given that 98% of studies cited budgetary constraints as a barrier, health policy makers must reallocate public health funding from expensive tertiary hospital care toward community-based primary care frameworks anchored in cost-effective prevention models like NSM [33, 35].

3. Clinical Practice Tool Standardization: Healthcare organizations should develop simplified, digital NSM assessment tools and electronic health record (EHR) templates that allow frontline community nurses to rapidly categorize stressors and document primary, secondary, and tertiary interventions without administrative burden [18, 30].

Strengths and Limitations

This systematic review possesses several notable strengths, including prospective PROSPERO registration, strict adherence to JBI critical appraisal methodology, comprehensive multi-database searching, and rigorous thematic synthesis. However, limitations must be acknowledged. First, restricting eligibility to English-language peer-reviewed publications between 2020 and 2024 may have omitted relevant studies published in other languages or earlier timeframes. Second, the geographic concentration of primary research in a few countries (e.g., Iran and UK) may limit global generalizability. Third, the reliance on qualitative narrative synthesis due to methodological heterogeneity prevents exact effect-size estimation.

CONCLUSIONS

The Neuman Systems Model provides a transformative, evidence-based, and holistic framework for improving community health services globally. Its structured categorization of environmental stressors and proactive three-tier prevention levels significantly enhance primary health promotion, chronic illness management, community mental health care, and interdisciplinary collaboration. While operational obstacles such as funding deficits, workforce training gaps, and institutional resistance remain prevalent, strategic investments in community nursing education, administrative support, and standardized clinical assessment tools will unlock the full potential of Neuman's model. Integrating NSM into primary healthcare policy is an essential step toward achieving universal, equitable, and patient-centered healthcare.

REFERENCES

1. Eri TS, Berg M, Dahl B, Gottfreðsdóttir H, Sommerseth E, Prinds C. Models for midwifery care: A mapping review. Eur J Midwifery. 2020;4:30. doi:10.18332/ejm/124328.
2. Smith J. Perceptions of stress in community clinic nurses during the COVID-19 pandemic. J Community Health Nurs. 2021;38(2):101-112.
3. World Health Organization, UNICEF. Global report on assistive technology. Geneva: World Health Organization; 2022. Licence: CC BY-NC-SA 3.0 IGO.
4. Suhrcke M, Stuckler D, Suk JE, Desai M, Senek M, McKee M, et al. The impact of economic crises on communicable disease transmission and control: A systematic review of the evidence. PLOS ONE. 2011;6(6):e20724. doi:10.1371/journal.pone.0020724.
5. Montano AR. Neuman Systems Model with nurse-led interprofessional collaborative practice. Nurs Sci Q. 2021;34(1):45-53. doi:10.1177/0894318420965219.
6. Patrick M, Zaman MS, Afzal G, Mahsud M, Hanifatu MN. Factors that affect maternal mortality in Rwanda: A comparative study with India and Bangladesh. Comput Math Methods Med. 2022;2022:1940188. doi:10.1155/2022/1940188.
7. Lima LH, Monteiro EM, Coriolano MW, Linhares FM, Cavalcanti AM. Family fortresses in Zika Congenital Syndrome according to Betty Neuman. Rev Bras Enferm. 2020;73(suppl 2):e20180578. doi:10.1590/0034-7167-2018-0578.
8. Meiselman HL. Quality of life, well-being and wellness: Measuring subjective health for foods and other products. Food Qual Prefer. 2016;54:101-109. doi:10.1016/j.foodqual.2016.05.009.
9. Major N, Rouleau M, Krantz C, Morris K, Séguin F, Allard M, et al. It's about time: Rapid implementation of a hub-and-spoke care delivery model for tertiary-integrated complex care services in a Northern Ontario community. Healthc Q. 2018;21(2):35-40. doi:10.12927/hcq.2018.25547.
10. Gervasi R, Mastrogiacomo L, Franceschini F. A conceptual framework to evaluate human-robot collaboration. Int J Adv Manuf Technol. 2020;108(3):841-865. doi:10.1007/s00170-020-05363-1.
11. Chen J, Abbod M, Shieh JS. Pain and stress detection using wearable sensors and devices—A review. Sensors. 2021;21(4):1030. doi:10.3390/s21041030.
12. Mackintosh NJ, Davis RE, Easter A, Rayment-Jones H, Sevdalis N, Wilson S, et al. Interventions to increase patient and family involvement in escalation of care for acute life-threatening illness in community health and hospital settings. Cochrane Database Syst Rev. 2020;12(12):CD012829. doi:10.1002/14651858.CD012829.pub2.
13. Stockton DA, Fowler C, Debono D, Travaglia J. World Health Organization building blocks in rural community health services: An integrative review. Health Sci Rep. 2021;4(2):e254. doi:10.1002/hsr2.254.
14. Montano AR, Shellman J, Malcolm M, McDonald D, Rees C, Fortinsky R, et al. A mixed methods evaluation of GOT CARE! Geriatr Nurs. 2020;41(6):822-831. doi:10.1016/j.gerinurse.2020.05.011.

15. Hudson MM, Bhatia S, Casillas J, Landier W, Rogers ZR, Allen C, et al. Long-term follow-up care for childhood, adolescent, and young adult cancer survivors. Pediatrics. 2021;148(3):e2021053127. doi:10.1542/peds.2021-053127.
16. Hart MJ, Flitner AM, Kornbluh ME, Thompson DC, Davis AL, Lanza-Gregory J, et al. Combining MTSS and community-based mentoring programs. Sch Psychol Rev. 2024;53(2):185-199. doi:10.1080/2372966X.2022.2046808.
17. Nichols CR, Wright LD, Bainbridge SJ, Cosby A, Hénaff A, Loftis JD, et al. Collaborative science to enhance coastal resilience and adaptation. Front Mar Sci. 2019;6:404. doi:10.3389/fmars.2019.00404.
18. Bauer M, Glenn T, Geddes J, Gitlin M, Grof P, Kessing LV, et al. Smartphones in mental health: A critical review of background issues, current status and future concerns. Int J Bipolar Disord. 2020;8(1):2. doi:10.1186/s40345-019-0164-x.
19. Akhlaghi E, Babaei S, Abolhassani S. Modifying stressors using Betty Neuman System Modeling in coronary artery bypass graft: A randomized clinical trial. J Caring Sci. 2020;9(1):13-19. doi:10.34172/jcs.2020.003.
20. Hutchings P, Chan MY, Cuadrado L, Ezbakhe F, Mesa B, Tamekawa C, et al. A systematic review of success factors in the community management of rural water supplies over the past 30 years. Water Policy. 2015;17(5):963-983. doi:10.2166/wp.2015.128.
21. Hosseinnejad A, Rassouli M, Jahani S, Elahi N, Molavynejad S. Community health nursing in Iran: A review of challenges and solutions (An integrative review). Front Public Health. 2022;10:899211. doi:10.3389/fpubh.2022.899211.
22. Kakhi K, Alizadehsani R, Kabir HM, Khosravi A, Nahavandi S, Acharya UR. The internet of medical things and artificial intelligence: Trends, challenges, and opportunities. Biocybern Biomed Eng. 2022;42(3):749-771. doi:10.1016/j.bbe.2022.05.009.
23. Kabusi M, Yazdi K. Nursing process based on Betty Neuman's systemic model: A case study. J Nurs Adv Clin Sci. 2024;1(1):1-7.
24. Rassouli M, Khanali Mojen L, Shirinabadi Farahani A, Beiranvand S. The role of the nurse in the community in running the palliative care interdisciplinary team: The Iranian experience. In: Silbermann M, editor. Palliative Care for Chronic Cancer Patients in the Community. Cham: Springer; 2021. p. 317-338.
25. Hosseinnejad A, Rassouli M, Jahani S, Elahi N, Molavynejad S. Requirements for creating a position for community health nursing within the Iranian primary health care system: A SWOT analysis. Front Public Health. 2022;9:793973. doi:10.3389/fpubh.2021.793973.
26. Bademli K, Duman ZÇ. Conceptual framework for nurses in the use of the Neuman Systems Model on caregivers of people suffering by schizophrenia. Int Arch Nurs Health Care. 2017;3(3):079. doi:10.23937/2469-5823/1510079.
27. Elsey H, Abboah-Offei M, Vidyasagaran AL, Anaseba D, Wallace L, Nwameme A, et al. Implementation of the Community-based Health Planning and Services (CHPS) in rural and urban Ghana: A history and systematic review of what works, for whom and why. Front Public Health. 2023;11:1105495. doi:10.3389/fpubh.2023.1105495.
28. Akhlaghi E, Babaei S, Mardani A, Eskandari F. The effect of the Neuman Systems Model on anxiety in patients undergoing coronary artery bypass graft: A randomized controlled trial. J Nurs Res. 2021;29(4):e162. doi:10.1097/jnr.0000000000000438.
29. Harvey C, Zirnsak TM, Brasier C, Ennals P, Fletcher J, Hamilton B, et al. Community-based models of care facilitating the recovery of people living with persistent and complex mental health needs: A systematic review and narrative synthesis. Front Psychiatry. 2023;14:1259944. doi:10.3389/fpsyt.2023.1259944.
30. Li X, Zhang F, Hui EC, Lang W. Collaborative workshop and community participation: A new approach to urban regeneration in China. Cities. 2020;102:102743. doi:10.1016/j.cities.2020.102743.
31. Castillo EG, Ijadi-Maghsoodi R, Shadravan S, Moore E, Mensah MO, Docherty M, et al. Community interventions to promote mental health and social equity. Curr Psychiatry Rep. 2019;21(5):35. doi:10.1007/s11920-019-1017-0.
32. Chisholm LJ, Hale RL, Knight SL. Community resilience after hurricanes: Can Neuman's Systems Theory guide public health nursing? Res Theory Nurs Pract. 2023;37(1):84-100. doi:10.1891/RTNP-2021-0022.
33. Nguyen DT, d'Hauteserre AM, Serrao-Neumann S. Intrinsic barriers to and opportunities for community empowerment in community-based tourism development in Thai Nguyen province, Vietnam. J Sustain Tour. 2022;30(4):723-741. doi:10.1080/09669582.2021.1891240.
34. Wood SK. Keeping the nurse in the nurse practitioner: Returning to our disciplinary roots of knowing in nursing. Adv Nurs Sci. 2020;43(1):50-61. doi:10.1097/ANS.0000000000000281.
35. Fernandes MA, Nóbrega MM, Zaccara AA, Freire ME, Andrade FF, Costa SF. Fawcett analysis and evaluation model applied to the theory of chronic sorrow. Texto Contexto Enferm. 2021;30:e20200010. doi:10.1590/1980-265X-TCE-2020-0010.
36. Juniarti N, Lukman M, Mulya AP. Potential of contracting out through the Nursing Center model: An integrative review. Preprints. 2023;2023091303. doi:10.20944/preprints202309.1303.v1.
37. Ornek OK, Esin MN. Effects of a work-related stress model based mental health promotion program on job stress, stress reactions and coping profiles of women workers: A control groups study. BMC Public Health. 2020;20(1):1658. doi:10.1186/s12889-020-09768-0.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Provision of Reactive Power Demand Using Solar Photovoltaic Systems

Mert KAYA*, Talha Enes GUMUS [2]

[1]*Electrical & Electronics Engineering /Sakarya University, Turkey*
[2]*Electrical & Electronics Engineering /Sakarya University, Turkey*

**mertkaya.eem@gmail.com*

Abstract – Distributed solar photovoltaic (PV) generation is expanding at a pace that is changing how reactive power is balanced on distribution networks. Unlike a synchronous generator, a PV plant connects to the grid through a power-electronic inverter, and that inverter supplies no reactive power on its own unless it is deliberately built and programmed to do so, even as the PV output itself is quietly reducing the reactive support that used to come from nearby conventional generation. This paper looks at whether the very same inverter hardware that converts PV direct current into grid-synchronous alternating current can double as a source of reactive power for the surrounding network. Starting from the inverter's rated apparent power and its instantaneous active power output, the reactive power headroom actually available at any moment is derived, and three ways that headroom can be dispatched are worked through: a fixed power factor, a fixed reactive power setpoint, and a voltage-dependent Volt-VAR droop law consistent with IEEE Standard 1547-2018. A stylized twelve-bus radial feeder with a heavy concentration of PV near its far end is then used to show what Volt-VAR control does to the daytime voltage profile, and the discussion extends to the inverter's ability to keep supplying reactive power after dark, acting essentially as a STATCOM. In the feeder study, switching on Volt-VAR control cut the end-of-feeder voltage deviation by roughly a third compared with uncontrolled operation, without adding a single piece of new equipment. The broader implication is that a properly sized and configured PV inverter is a genuinely useful, low-cost reactive power resource, one that can lessen a network's reliance on capacitor banks and dedicated compensators, so long as sizing margins and control settings are chosen with the applicable grid code in mind.

Keywords – reactive power compensation; photovoltaic inverter; Volt-VAR control; voltage regulation; smart inverter; power factor control; distributed generation

I. INTRODUCTION

Worldwide, cumulative solar photovoltaic capacity had climbed to roughly 2,383 GW by the close of 2025, and the amount added in that single year alone — about 510 GW — was larger than the annual addition of any other electricity-generating technology on record [1]. A large share of this new capacity sits on medium- and low-voltage feeders, wired in close to the loads it serves rather than at centralized power stations, and that proximity matters for reactive power in two separate ways. Local PV output cuts the amount of active power that has to be imported from further up the network, which in turn reduces the incidental reactive power support that loaded lines and transformers would otherwise provide. At the same time, because a PV plant connects through a power-electronic inverter rather than a spinning machine, it contributes no reactive power of its own unless that behaviour is deliberately designed and programmed in.

Utilities have long met local reactive power needs with switched capacitor banks, synchronous condensers, static VAR compensators, or STATCOMs. All of these work, but each is a dedicated piece of equipment: it costs money to buy and install, it takes up space, and in the case of mechanically switched capacitors, its output changes in coarse steps while its contacts wear out with repeated switching. A grid-tied PV inverter is a different kind of asset. It already contains the semiconductor bridge, the output filter,

and the digital controller needed to synthesize a current waveform at essentially any phase angle. As long as it is not already pushed to its full apparent-power limit, that same inverter can inject or absorb reactive current with nothing more than a change to its control firmware, at little or no added hardware cost.

This is no longer just a research proposal; it has been written into interconnection rules. IEEE Standard 1547-2018 requires Category B distributed energy resources — a category that covers most commercial-scale and many utility-scale PV systems — to be able to exchange reactive power up to at least 44% of their rated apparent power, and it spells out standardized operating modes, including fixed power factor and voltage-dependent (Volt-VAR) control, that a DER controller can be configured to run [2]. Similar obligations appear in European rules such as EN 50549 and in the older IEC 61727 [10], [11]. Taken together, these documents mark a shift in expectation: a distributed PV inverter is no longer treated as a passive source that simply pushes out whatever active power the array happens to produce, but as an active participant in local voltage management.

The rest of this paper is organized as follows. Section II works out the reactive power capability of a PV inverter, sets down control laws for dispatching that capability, and describes the feeder and simulation approach used to test them. Section III reports the resulting capability curve, the Volt-VAR characteristic, and the simulated voltage profiles. Section IV weighs the practical, economic and coordination questions the results raise, in light of related work, and Section V closes with the main conclusions.

II. Materials and Methods

The reactive power capability of a representative single-stage, grid-tied PV inverter is worked out first, and two control laws for dispatching that capability are then formulated. These are applied to a simplified radial feeder to gauge how much of a voltage-regulation benefit Volt-VAR control delivers relative to plain unity-power-factor operation.

A. System Description and Reactive Power Capability

Fig. 1 sketches the PV system studied here. A PV array feeds a DC-DC boost stage running a maximum-power-point-tracking (MPPT) algorithm, which sets the DC-link voltage seen by a grid-tied voltage-source inverter (VSC). The inverter output passes through an LCL filter, needed to keep switching harmonics off the network, before reaching the point of common coupling (PCC). A separate reactive power controller reads the PCC voltage and issues a reactive current reference to the inverter's current loop, on top of, and independent from, the active power reference that the MPPT stage is already producing.

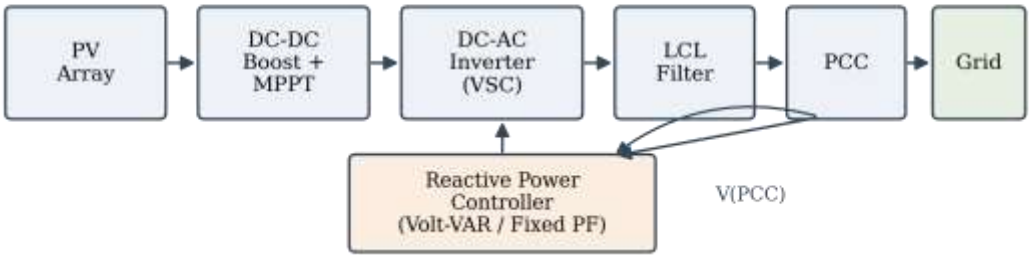


Fig. 1 Block diagram of the grid-connected PV system studied, showing the reactive power controller acting alongside the maximum-power-point active power reference.

Whatever the inverter is doing at a given moment, its active power output, P, and reactive power output, Q, are jointly bounded by the rated apparent power, S_rated, through the usual apparent-power relationship

$$S_{rated}^2 = P^2 + Q^2 \tag{1}$$

which sets the maximum reactive power obtainable at a given level of active power output as

$$Q_{max}(P) = \sqrt{S_{rated}^2 - P^2} \tag{2}$$

Equation (2) traces out the quarter-circle boundary plotted in Fig. 2, and two things follow from its shape. Reactive headroom is greatest exactly when active power is smallest — early morning, late evening, during passing cloud cover, and overnight — because the whole apparent-power rating is then free for reactive duty. Conversely, an inverter sized one-for-one with the array's DC rating (an inverter loading ratio of 1.0) has essentially no reactive power left right at the moment of peak active power output. That is the practical reason many designs deliberately oversize the inverter relative to the array, or accept a small amount of active-power curtailment near the rated point, so that the 0.44 p.u. minimum required by IEEE 1547-2018 Category B still holds even under peak sun [7].

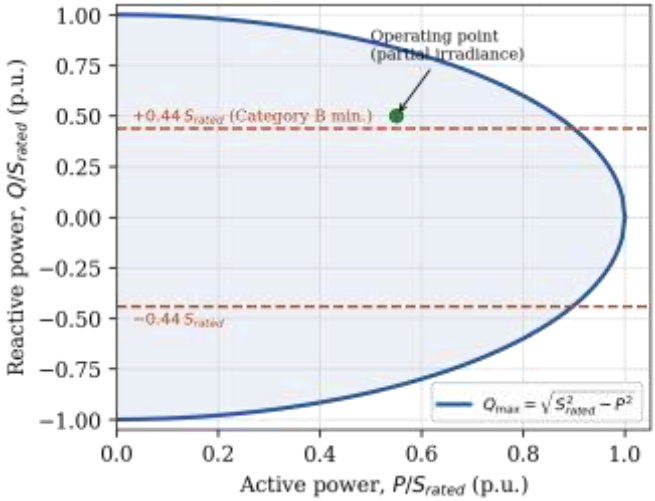


Fig. 2 Reactive power capability (P-Q) boundary of a grid-connected PV inverter, with the IEEE 1547-2018 Category B minimum requirement of ±0.44 p.u. marked.

B. Reactive Power Control Strategies

Three ways of putting this capability to use are considered here. The simplest, constant power factor control, holds a fixed ratio between reactive and active output,

$$Q = P\tan(\varphi) \tag{3}$$

with φ set as the target power-factor angle. It is easy to implement, but it takes no account of what the PCC voltage is actually doing, so depending on how the feeder happens to be loaded at a given moment, it can just as easily over-correct as under-correct. A related option, constant reactive power control, instead tracks a fixed setpoint, Q_set, handed down by the network operator, independent of P; this suits a pre-scheduled correction but is, again, blind to real-time voltage.

Volt-VAR control takes the opposite approach: reactive power output is written as a function of the voltage measured right at the inverter terminals, following a piecewise-linear curve pinned down by four breakpoints, (V1, Q1) through (V4, Q4). Across any one linear segment — say, between V1 and V2 —

$$Q(V) = Q_1 + \frac{Q_2 - Q_1}{V_2 - V_1}(V - V_1), \quad V_1 \le V \le V_2 \tag{4}$$

A deadband spanning V2 to V3, usually straddling 1.0 p.u., is normally held at zero reactive power so that the inverter ignores everyday voltage noise. Outside that band, the controller pulls reactive power in ($Q < 0$) when the voltage runs high and pushes it out ($Q > 0$) when the voltage runs low, nudging the local voltage back toward its reference. Fig. 3 plots the default curve that IEEE 1547-2018 assigns to Category B DERs, with breakpoints at V1 = 0.92 p.u., V2 = 0.98 p.u., V3 = 1.02 p.u. and V4 = 1.08 p.u., and limits of Q1 = +0.44 p.u. and Q4 = −0.44 p.u. [2]. The same standard also defines a companion Volt-Watt function that trims active power under sustained over-voltage; that function sits outside the scope of this study.

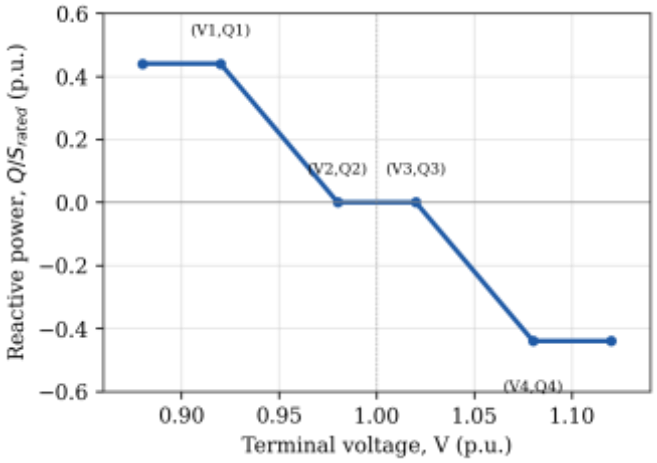


Fig. 3 Default Volt-VAR (Q-V) droop curve assigned to Category B DERs under IEEE Standard 1547-2018.

C. *Case-Study Feeder and Simulation Setup*

The voltage-regulation effect of Volt-VAR control was tested on a stylized twelve-bus radial feeder with PV generation clustered toward its far end, using a linearized (LinDistFlow) power-flow approximation. In this approximation, the voltage magnitude at bus i, measured relative to the source-bus voltage V0, is estimated from the active and reactive power flowing past that bus and the resistance R_i and reactance X_i of the line sections between the source and bus i [8]:

$$\Delta V_i \approx \frac{R_i P_i + X_i Q_i}{V_0} \tag{5}$$

Table 1 lists the parameters used for this exercise, chosen to represent a feeder segment under clear-sky, near-solar-noon conditions — the point in the day when reverse power flow, and the over-voltage risk that comes with it, is at its worst.

Table 1. Case-Study Feeder and Inverter Parameters

Parameter	**Value**
Feeder topology	12-bus radial (single-phase equivalent)
Source bus voltage, V0	1.00 p.u.
Average feeder R/X ratio	0.85
PV penetration (of feeder peak load)	65%
PV inverter rated apparent power, S_rated	1.00 p.u. per installed PV unit
Inverter loading ratio (DC/AC)	1.15
Volt-VAR breakpoints V1–V4	0.92 / 0.98 / 1.02 / 1.08 p.u.
Volt-VAR limits Q1, Q4	+0.44 / −0.44 p.u.
Simulated horizon / resolution	24 h / 10-min

D. *Performance Metric*

One more figure of merit is useful for judging how much of an inverter's reactive power headroom actually gets used across a day rather than sitting idle. The reactive power utilization factor (RPUF) is defined here as the time-averaged reactive power output divided by the rated reactive power capability,

$$\mathrm{RPUF} = \frac{Q_{avg}}{Q_{max}} \times 100\% \tag{6}$$

III. RESULTS

Fig. 4 sets the simulated feeder voltage profile with Volt-VAR control switched on against the same profile with it switched off. Left uncontrolled, the reverse power flow driven by concentrated PV output

pushes the far end of the feeder up to about 1.072 p.u., past the 1.05 p.u. guideline commonly used as an upper service-voltage bound. Turning Volt-VAR control on lets the inverters nearest the feeder end start absorbing reactive power as soon as the local voltage crosses the V2 breakpoint, which brings the end-of-feeder voltage back down to about 1.046 p.u. — a cut of roughly 36% in the deviation from nominal, achieved without touching the active power dispatch at all.

Fig. 5 traces the resulting daily active and reactive power pattern for one representative inverter operating under Volt-VAR control. Around midday, as active power output climbs toward the inverter's rated apparent power, the available reactive headroom narrows in step with the boundary traced in Fig. 2, and the Volt-VAR function simply uses whatever headroom remains. In the morning and evening ramps, and through the night, active power output is low or absent, so the inverter can hold reactive power output close to its full 0.44 p.u. rating — in effect running as a STATCOM once the sun sets. Averaged over the full 24-hour window, the reactive power utilization factor from (6) worked out to roughly 60%, which suggests that a sizeable share of the inverter's rated reactive capability gets called on well beyond the hours when it is actually producing active power.

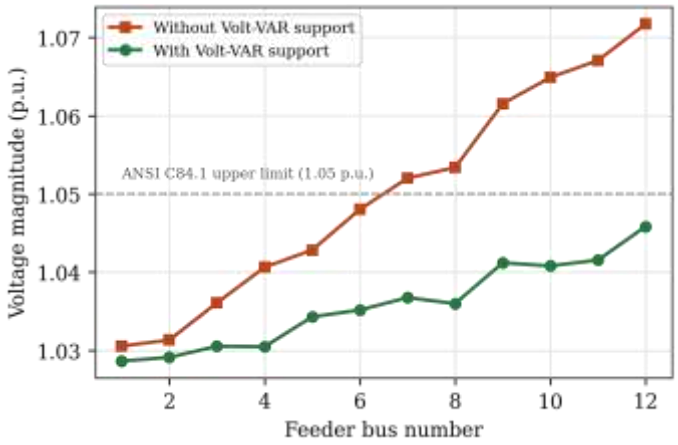


Fig. 4 Simulated feeder voltage profile at solar noon under heavy PV penetration, shown with and without Volt-VAR control.

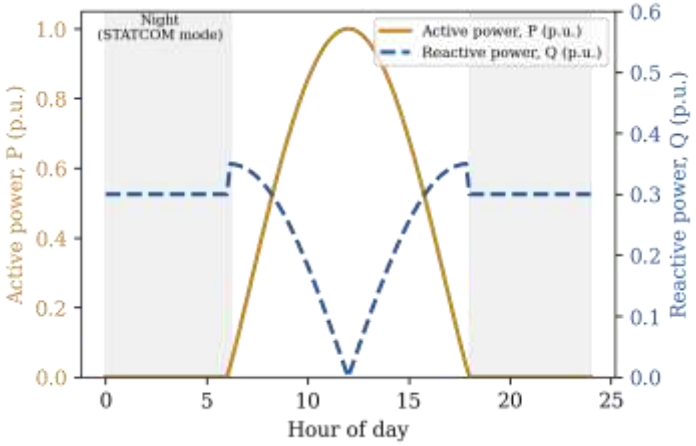


Fig. 5 Simulated daily active-power output and Volt-VAR reactive-power support for a feeder-end PV inverter.

Table 2. Summary of Simulation Results

Metric	Without Volt-VAR	With Volt-VAR
Peak feeder voltage (p.u.)	1.072	1.046
Voltage deviation from 1.00 p.u.	0.072 p.u.	0.046 p.u.
Deviation reduction	–	≈36%
Reactive power utilization factor	0%	≈60%

IV. Discussion

These findings line up with earlier surveys of smart-inverter behaviour, which likewise find that a voltage-responsive Volt-VAR law beats fixed or pre-scheduled power-factor control at holding down voltage deviation, simply because it reacts to what the feeder is actually doing rather than to a schedule set in advance that may no longer match conditions on the ground [4]. Because the reactive power drawn on in this study comes entirely from headroom already built into the inverter's apparent-power rating, the improvement seen in Fig. 4 costs nothing in additional hardware — a clear contrast with switched capacitor banks or purpose-built STATCOMs.

The result that inverters can keep delivering meaningful reactive power well outside daylight hours is backed up by field and laboratory work showing that commercial PV inverters can inject or absorb reactive current at night without pulling active power from the grid to keep themselves running [5]. Separate field trials on utility-scale plants built from clusters of inverters have confirmed that the aggregate reactive capability predicted by (2) holds up in practice, once auxiliary power draw and plant-controller communication delays are accounted for [6]. Put together, these findings support treating a PV plant as a reactive power resource for most of the day, not merely during the hours it happens to be generating.

Turning this into a scalable practice does raise sizing and thermal questions. Guaranteeing reactive capability at full active-power output typically calls for an inverter rated above the array's DC capacity, and sustained reactive current draws on semiconductor and capacitor life in ways that reliability-driven sizing studies need to factor in [7]. Coordination across a feeder full of inverters matters too: if every unit on the same feeder runs an identical, aggressive Volt-VAR slope, the devices can end up chasing each other's corrections and oscillating, which is exactly why IEEE 1547-2018 specifies conservative default slopes and adjustable response times [2].

The linearized power-flow model used in Section II-C is good enough to capture the main voltage-regulation trend, but it leaves out phase imbalance, harmonic interaction with the LCL filter, and communication-dependent advanced modes such as the autonomous or scheduled Volt-VAR curves defined for Category C DERs. A fuller three-phase, time-domain study, paired with a year-long techno-economic comparison against dedicated compensation hardware, is left for future work.

V. Conclusion

This paper looked at whether grid-tied PV inverters can help meet local reactive power demand. The reactive power an inverter can supply turns out to follow directly from its apparent-power rating and whatever active power it happens to be producing at the time, and in a simplified feeder case study, Volt-VAR droop control cut the end-of-feeder voltage deviation by roughly a third relative to uncontrolled operation, while drawing on headroom that would otherwise sit unused for most of the day and is available in full once night falls. Because none of this needs new hardware — only a control change on equipment that is already installed — a properly sized and configured PV inverter is a genuinely practical, low-cost complement to conventional reactive power compensation. How far this gets adopted will come down to consistent grid-code requirements, adequate sizing margins on the inverter side, and Volt-VAR settings that are coordinated across feeders with many inverters packed close together.

Acknowledgment

No external funding was received for this work, and the author declares no conflict of interest.

References

[1] International Renewable Energy Agency (IRENA), *Renewable Capacity Statistics 2026*, IRENA, Abu Dhabi, United Arab Emirates, 2026.

[2] IEEE Std 1547-2018, *IEEE Standard for Interconnection and Interoperability of Distributed Energy Resources with Associated Electric Power Systems Interfaces*, IEEE, Piscataway, NJ, USA, 2018.

[3] P. Kundur and O. P. Malik, *Power System Stability and Control*, 2nd ed. New York, NY, USA: McGraw-Hill, 2022.

[4] T. O. Olowu and A. I. Sarwat, "Smart inverters' functionalities and their impacts on distribution feeders at high photovoltaic penetration: An overview," *arXiv:2008.11125*, 2020.

[5] K. T. Lulbadda and K. T. M. U. Hemapala, "Use of solar PV inverters during night-time for voltage regulation and stability of the utility grid," *Clean Energy*, vol. 6, no. 4, pp. 646–657, Aug. 2022.

[6] J. Huang, M. Liu, J. Zhang, *et al.*, “Analysis and field test on reactive capability of photovoltaic power plants based on clusters of inverters,” *J. Mod. Power Syst. Clean Energy*, vol. 5, no. 4, pp. 559–567, Jul. 2017.
[7] A. G. de Souza Júnior *et al.*, “A grid connection photovoltaic inverter with volt-VAR control and reactive power support for voltage regulation,” *Int. J. Circuit Theory Appl.*, vol. 52, no. 12, pp. 6362–6389, 2024.
[8] M. E. Baran and F. F. Wu, “Optimal sizing of capacitors placed on a radial distribution system,” *IEEE Trans. Power Del.*, vol. 4, no. 1, pp. 735–743, Jan. 1989.
[9] N. G. Hingorani and L. Gyugyi, *Understanding FACTS: Concepts and Technology of Flexible AC Transmission Systems*. Piscataway, NJ, USA: IEEE Press, 2000.
[10] EN 50549-1:2019, *Requirements for Generating Plants to Be Connected in Parallel with Distribution Networks – Part 1: Connection to a LV Distribution Network*, CENELEC, Brussels, Belgium, 2019.
[11] IEC 61727:2004, *Photovoltaic (PV) Systems – Characteristics of the Utility Interface*, International Electrotechnical Commission, Geneva, Switzerland, 2004.
[12] National Renewable Energy Laboratory (NREL), “The challenges of inverter modeling related to IEEE 1547-2018,” Solar Energy Technologies Office (SETO) Modeling Workshop, U.S. Department of Energy, Golden, CO, USA, 2019.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Sosyal Bilgiler Öğretmenlerinin Dijital Öğretim Yeterlikleri ile Dijital Materyal Tasarımı Yeterlikleri Arasındaki İlişki

Şenol Mail Pala[1*] ve Hüseyin Bayram [2]

[1]*Türkçe ve Sosyal Bilimler Eğitimi / Eğitim Fakültesi, Erzincan Binali Yıldırım Üniversitesi, Türkiye*
[2]*Türkçe ve Sosyal Bilimler Eğitimi / Eğitim Fakültesi, Ağrı İbrahim Çeçen Üniversitesi, Türkiye*

[*]*(senolmailpala@gmail.com)*

Özet **–**Eğitim teknolojilerinin yaygınlaşması, öğretmenlerin dijital araçları etkili ve bilinçli kullanabilme becerilerini geliştirmelerini zorunlu kılmaktadır. Bu bağlamda, dijital öğretim yeterlikleri ve dijital materyal tasarımı yeterlikleri, öğretmenlerin derslerde teknoloji entegrasyonunu ve öğrenci öğrenmesini desteklemede önemli bir hale gelmektedir. Sosyal bilgiler derslerinde, öğretmenlerin sahip olduğu dijital yeterliklerin, öğrencilerin ders içi etkileşimini ve öğrenme deneyimini doğrudan etkileyebileceği düşünülmektedir. Öğrenme öğretme sürecinde dijital materyallerin tasarlanması ve derse entegre edilmesi, öğretmenlerin yalnızca teknik bilgiye değil, aynı zamanda pedagojik ve tasarımsal yeterliklere de sahip olmasını gerektirmektedir. Bu nedenle, iki yeterlik alanı arasındaki ilişkinin incelenmesi, sosyal bilgiler öğretmenlerinin mesleki gelişim ihtiyaçlarının belirlenmesi açısından önem taşımaktadır. Bu araştırmanın amacı, sosyal bilgiler öğretmenlerinin dijital öğretim yeterlikleri ile dijital materyal tasarımı yeterlikleri arasındaki ilişkiyi ortaya koymaktır. Araştırma, nicel araştırma yönteminin ilişkisel tarama modeli ile yürütülmüştür. Araştırmanın örneklemini Türkiye'de görev yapan 70 kadın ve 99 erkek sosyal bilgiler öğretmeni oluşturmuştur. Veriler, Dijital Öğretim Yeterliği Ölçeği ve Dijital Materyal Tasarımı Yeterlikleri Ölçeği aracılığıyla toplanmış, analiz sürecinde betimleyici istatistikler ve Pearson korelasyon katsayısı kullanılmıştır. Araştırma sonuçları, sosyal bilgiler öğretmenlerinin hem dijital öğretim yeterliklerinin hem de dijital materyal tasarımı yeterliklerinin orta düzeyde olduğunu göstermektedir. Ayrıca, iki yeterlik arasında pozitif yönlü ve anlamlı bir ilişki saptanmıştır. Elde edilen sonuçlar doğrultusunda, farklı branş ve kademelerdeki öğretmenlerin konu edildiği araştırmalar yapılması önerilmiştir.

Anahtar Kelimeler – Sosyal bilgiler öğretmenleri, dijital öğretim yeterliği, dijital materyal tasarım yeterliği, teknoloji entegrasyonu, öğretmen yeterlikleri

I. GİRİŞ

Eğitim sistemlerinde yaşanan teknolojik gelişmeler itibarıyla kesintisiz bir dönüşüm yaşanmaktadır. Teknolojik gelişmeler, öğrenme-öğretme ortamlarını biçimlendirirken öğretmenlerin sahip olması gereken bilgi, beceri ve yeterlik alanlarını da genişletmektedir. Bu kapsamda öğretmenlerin dijital teknolojileri etkili biçimde kullanabilmeleri, çağdaş eğitim anlayışının gerekliliklerinden biri olarak görülmektedir. Bir başka deyişle öğretmenlerin dijital öğretim yeterliğine sahip olması beklenmektedir.

Dijital öğretim yeterliği, öğretmenlerin dijital araçları; ders planlama, öğretim süreci, ölçme-değerlendirme ve sınıf yönetimi gibi alanlarda bilinçli ve etkili biçimde kullanabilme düzeylerini ifade etmektedir [17]. Bu yeterlik, dijital çağın gerektirdiği nitelikli öğrenme ortamlarının oluşturulmasında önemli bir rol oynamaktadır [9]. Öğretmenlerin dijital öğretim yeterliğine sahip olmasının tüm derslerde olduğu gibi sosyal bilgiler dersinde de bir gereklilik olduğu söylenebilir.

Sosyal bilgiler, bireylerin toplumsal yaşama etkin biçimde katılımını sağlamaya yönelik bir derstir [1]. Dijital teknolojiler ve dijital materyaller, sosyal bilgiler dersinin yapısına uygun olarak öğrenme sürecini zenginleştirme, öğrencilerin derse ilgisini artırma ve kalıcı öğrenmeler gerçekleştirmek için gerekli olan

ortamların oluşturulmasında etkili değişkenlerdir [2]. Bu nedenle sosyal bilgiler öğretmenlerinin dijital materyalleri etkili biçimde tasarlayabilmeleri ve kullanabilmeleri, dersin hedeflerine ulaşmasında belirleyici bir unsur olarak kabul edilebilir.

Dijital materyal tasarımı yeterliği, öğretmenlerin dersin kazanımlarına uygun, öğrenci düzeyini gözeten ve teknik açıdan nitelikli dijital içerikler geliştirebilme becerilerini kapsamaktadır [3]. Etkili dijital materyallerin tasarlanabilmesi ise yalnızca teknik bilgiyle sınırlı olmayıp, öğretmenin dijital öğretim yeterliği ile yakından ilişkilidir [17]. Dijital araçları öğretim sürecine entegre edebilme, uygun yazılım ve platformları seçebilme ve bu araçları amaçlar doğrultusunda kullanabilme düzeyi, materyal tasarımı sürecini doğrudan etkilemektedir [14]. Bu açıdan dijital öğretim yeterliği ile dijital materyal tasarımı yeterliklerinin birbiriyle ilişkili iki değişken olduğu söylenebilir.

Alanyazın incelendiğinde; öğretmenlerin dijital yeterlikleri, teknoloji kullanma becerileri ya da dijital materyal geliştirme konularına odaklanan çeşitli araştırmaların [13], [10], [19], [16], [2], [6] yapıldığı görülmektedir. Ancak alanyazında sosyal bilgiler öğretmenlerinin dijital öğretim yeterliği ile dijital materyal tasarımı yeterlikleri arasındaki ilişkinin birlikte ele alındığı herhangi bir araştırma ise yer almamaktadır. Bu noktadan hareketle bu araştırmanın yapılıp sosyal bilgiler öğretmenlerinin dijital öğretim yeterliği ile dijital materyal tasarımı yeterliği arasındaki ilişkinin ortaya konması hedeflenmiştir. Araştırmanın sonuçlarının ilgili alanyazına katkı sağlayacağı ve benzer araştırmalar yapmak isteyen araştırmacılara kaynak oluşturacağı düşünülmektedir.

A. *Amaç*

Bu araştırmanın amacı, sosyal bilgiler öğretmenlerinin dijital öğretim yeterlikleri ile dijital materyal tasarımı yeterlikleri arasındaki ilişkinin incelenmesidir. Bu doğrultuda aşağıdaki araştırma sorularına yanıt aranmıştır:

Sosyal bilgiler öğretmenlerinin;

Dijital öğretim yeterlik düzeyleri nedir?

Dijital materyal tasarımı yeterlik düzeyleri nedir?

Dijital öğretim yeterlikleri ile dijital materyal tasarımı yeterlikleri arasında istatistiksel olarak anlamlı bir ilişki var mıdır?

II. MATERYAL VE YÖNTEM

Araştırma, nicel araştırma yöntemi çerçevesinde ilişkisel tarama modeli kullanılarak gerçekleştirilmiştir. İlişkisel tarama modeli, iki ya da daha fazla değişken arasındaki ilişkinin yönünü ve gücünü belirlemeye yönelik araştırmalarda kullanılmaktadır [4]. Bu araştırmada söz konusu modelin tercih edilmesinin nedeni, sosyal bilgiler öğretmenlerinin dijital öğretim yeterlikleri ile dijital materyal tasarımı yeterlikleri arasındaki ilişkinin ortaya konmak istenmesidir.

A. *Evren ve Örneklem*

Araştırmanın evrenini Türkiye genelinde görev yapan sosyal bilgiler öğretmenleri oluşturmaktadır. Araştırmanın örneklemi ise ulaşılabilirlik esasına dayalı olarak belirlenen öğretmenlerden meydana gelmiştir. Bu kapsamda kolay ulaşılabilir örnekleme yöntemi kullanılmıştır [4]. Örnekleme ilişkin demografik bilgiler Tablo 1'de sunulmuştur:

Tablo 1. Örnekleme ilişkin demografik bilgiler

Değişken	**Özellik**	**f**	**%**
Cinsiyet	Kadın	70	41.42
	Erkek	99	58.58
Toplam		169	100

Tablo 1 incelendiğinde araştırmaya katılan öğretmenlerin 70'inin (%41.42) kadın, 99'unun (%58.58) erkek olduğu görülmektedir. Toplam 169 sosyal bilgiler öğretmeni araştırmanın örneklemini oluşturmaktadır.

B. Verilerin toplanması

Veriler toplanırken Gao ve Zheng (2023) tarafından geliştirilen [7] ve Karalar vd. (2025) tarafından Türkçe'ye uyarlanan Dijital Öğretim Yeterliği Ölçeği (DÖYÖ) [11] ile Göçen Kabaran ve Uşun (2021) tarafından geliştirilen Dijital Materyal Tasarımı Yeterlikleri Ölçeği (DMTYÖ) kullanılmıştır [8].

Beş boyut ve 29 maddeden oluşan DÖYÖ, 5'li Likert tipindedir. Karalar vd. (2025), ölçeğin boyutlarının Cronbach Alpha katsayısını, .88-.90 aralığında hesaplarken [11] ölçeğin bu araştırmadaki Cronbach Alpha katsayısı, ölçeğin toplamı için .75 şeklinde bulunmuştur. Sonucun .60'tan büyük [15] değer alması, ölçeğin kullanıma uygun olduğunu göstermiştir.

Dört boyut ve 31 maddeden oluşan DMTYÖ, 5'li Likert tipindedir. Göçen Kabaran ve Uşun (2021), ölçeğin Cronbach Alpha katsayısını, .98 olarak hesaplarken [8] ölçeğin bu araştırmadaki Cronbach Alpha katsayısı, .73 şeklinde bulunmuştur. Sonucun .60'tan büyük [15] değer alması, ölçeğin kullanıma uygun olduğunu göstermiştir.

C. Verilerin analizi

Verilerin analizine geçilmeden önce dağılımın normalliği çarpıklık ve basıklık değerleri ile Kolmogorov-Smirnov testi kullanılarak incelenmiştir. Elde edilen bulgular Tablo 2'de gösterilmiştir:

Tablo 2. Normallik testi sonuçları

Ölçek	**Boyut**	**Çarpıklık**	**Basıklık**	**Kolmogorov-Smirnov**
Dijital Öğretim Yeterliği	Toplam	-.88	-.73	.10
Dijital Materyal Tasarımı Yeterliği	Toplam	.19	-.17	.08

Tablo 2'de yer alan değerler incelendiğinde çarpıklık ve basıklık puanlarının -1.5 ile 1.5 aralığında olduğu ve Kolmogorov-Smirnov testi sonuçlarının .05'ten büyük olduğu görülmektedir. Bu durum verilerin normal dağılım gösterdiğine işaret etmektedir [18]. Bu nedenle analiz sürecinde parametrik testlerden yararlanılmıştır.

Verilerin analizinde öncelikle betimsel istatistikler kullanılarak öğretmenlerin yeterlik düzeyleri belirlenmiş, ardından değişkenler arasındaki ilişki Pearson korelasyon analizi ile incelenmiştir. Araştırmada değişkenler arasındaki uyumu sağlamak için Araştırma Uyum Matrisi [12] kullanılmıştır.

III. BULGULAR

Sosyal bilgiler öğretmenlerinin dijital öğretim yeterlikleri ile dijital materyal tasarımı yeterlik düzeylerini belirlemek amacıyla betimsel istatistikler hesaplanmış ve sonuçlar Tablo 3'te sunulmuştur:

Tablo 3. Tanımlayıcı istatistik sonuçları

Değişken	**n**	**Ortalama**	**Standart Sapma**
Dijital Öğretim Yeterliği	169	3.61	.53
Dijital Materyal Tasarımı Yeterliği	169	3.56	.66

Tablo 3'e göre sosyal bilgiler öğretmenlerinin dijital öğretim yeterliklerinin ortalaması 3.61, dijital materyal tasarımı yeterliklerinin ortalaması ise 3.56 olarak bulunmuştur. Bu bulgular, her iki yeterlik düzeyinin de orta düzeyin üzerinde olduğunu göstermektedir.

Öğretmenlerin dijital öğretim yeterlikleri ile dijital materyal tasarımı yeterlikleri arasındaki ilişki Pearson korelasyon analizi ile incelenmiş ve elde edilen sonuçlar Tablo 4'te verilmiştir:

Tablo 4. Pearson korelasyon analizi sonuçları

Değişkenler	**1**	**2**
1. Dijital Öğretim Yeterliği	1	
2. Dijital Materyal Tasarımı Yeterliği	.57*	1

*p < .05

Tablo 4 incelendiğinde, sosyal bilgiler öğretmenlerinin dijital öğretim yeterlikleri ile dijital materyal tasarımı yeterlikleri arasında pozitif yönde ve anlamlı bir ilişki olduğu görülmektedir ($r = .57$, $p < .05$). Bu sonuç, öğretmenlerin dijital öğretim yeterlikleri arttıkça dijital materyal tasarlama yeterliklerinin de arttığını göstermektedir.

IV. SONUÇ VE TARTIŞMA

Araştırmada sosyal bilgiler öğretmenlerinin dijital öğretim yeterlik düzeylerinin orta düzeyde olduğu sonucuna ulaşılmıştır. Öğretmenlerin teknoloji kullanımına yönelik yeterliklerinin farklı değişkenlere bağlı olarak değişmekle birlikte çoğunlukla orta düzeyde olduğu da belirlenmiştir [16]. Benzer şekilde öğretmenlerin dijital yeterliklerinin özellikle hizmet öncesi ve hizmet içi eğitim süreçleriyle ilişkili olduğu ve bu yeterliklerin istenen düzeye ulaşmadığını vurgulanmıştır [10]. Ayrıca öğretmenlerin dijital yeterliklerinin gelişiminde pedagojik ve teknik boyutların birlikte ele alınması gerektiği ifade edilerek mevcut yeterlik düzeylerinin sınırlı olduğu ortaya konulmuştur [6]. Bu durum, elde edilen orta düzey sonucunun alanyazınla örtüştüğünü göstermektedir.

Araştırmada sosyal bilgiler öğretmenlerinin dijital materyal tasarımı yeterlik düzeylerinin orta düzeyde olduğu saptanmıştır. Öğretmenlerin teknolojiyi pedagojik ve içerik bilgisiyle bütünleştirmede çeşitli zorluklar yaşandığı ve bu durumun materyal geliştirme süreçlerini sınırladığı belirlenmiştir [13] belirtmiştir. Benzer şekilde öğretmenlerin dijital materyal geliştirme yeterliklerinin TPACK bileşenleriyle doğrudan ilişkili olduğu ve bu yeterliklerin çoğunlukla orta düzeyde kaldığı belirlenmiştir [5]. Ayrıca öğretmenlerin teknoloji entegrasyonu ve materyal geliştirme süreçlerinde yeterli deneyime sahip olmamalarının bu düzeyi etkilediği vurgulanmıştır [19].

Araştırmada sosyal bilgiler öğretmenlerinin dijital öğretim yeterlik düzeyleri ile dijital materyal tasarımı yeterlik düzeyleri arasında pozitif yönlü bir ilişki olduğu belirlenmiştir. Yapılan bir derleme çalışmasında öğretmenlerin teknolojik pedagojik içerik bilgisi geliştikçe öğretim sürecinde daha etkili materyaller tasarlayabildikleri ifade edilmiştir [20]. Benzer şekilde öğretmenlerin teknoloji kullanımına yönelik yeterlikleri arttıkça sınıf içi uygulamalarının ve materyal geliştirme becerilerinin de geliştiği ortaya konulmuştur [19]. Ayrıca dijital yeterlik ile pedagojik uygulamalar arasında karşılıklı bir etkileşim olduğunu ve bu durumun öğretim sürecine doğrudan yansıdığını belirten bir derleme çalışması bulunmaktadır [5].

V. ÖNERİLER

Ortaya konan sonuçlar ışığında şu öneriler geliştirilmiştir:

- Araştırmacılar tarafından farklı branşlardaki öğretmenler ile farklı eğitim kademelerinde görev yapan öğretmenlerin dijital öğretim ve dijital materyal tasarımı yeterlikleri karşılaştırmalı olarak incelenebilir.
- Araştırmada ele alınan yeterlikler, sosyal bilgiler öğretmenlerinin bireysel özellikleri açısından detaylandırılmamıştır. Gelecek çalışmalarda cinsiyet, mesleki deneyim, görev yapılan okul türü ya da teknolojik imkânlara erişim gibi değişkenler dikkate alınarak gruplar arası farklılıklar analiz edilebilir.
- Bu araştırma nicel yöntemle gerçekleştirilmiştir. Konunun daha derinlemesine anlaşılabilmesi için nitel ya da karma yöntemlerin kullanıldığı çalışmalar yapılarak sosyal bilgiler öğretmenlerinin dijital yeterliklerine ilişkin görüş, deneyim ve karşılaştıkları sorunlar ayrıntılı biçimde ortaya konulabilir.
- Sosyal bilgiler öğretmenlerinin dijital öğretim ve materyal tasarımı yeterliklerinin zaman içerisindeki gelişimini incelemek amacıyla uzun süreli araştırmalar yapılabilir.

KAYNAKLAR

[1] Barr, R. D., Barth, J. L., & Shermis, S. (1978). *The Nature and goals of social studies.* National Council for the Social Studies.

[2] Bayram, H. (2020). *Türkiye'de sosyal bilgiler öğretimi ve dijital okuryazarlık.* (Sözlü bildiri). 3. Uluslararası Uzaktan Öğrenme ve Yenilikçi Eğitim Teknolojileri Konferansı, Ankara, Türkiye.

[3] Branch, R. M. (2009). *Instructional design: The ADDIE approach*. Springer.

[4] Büyüköztürk, Ş., Çakmak, E., Akgün, Ö., Karadeniz, Ş., & Demirel, F. (2020). *Eğitimde bilimsel araştırma yöntemleri.* Pegem Akademi Yayıncılık.

[5] Chai, C. S., Koh, J. H. L., & Tsai, C. C. (2013). A review of technological pedagogical content knowledge. *Journal of Educational Technology & Society*, *16*(2), 31-51. https://www.jstor.org/stable/jeductechsoci.16.2.31

[6] Falloon, G. (2020). From digital literacy to digital competence: the teacher digital competency (TDC) framework. *Education Tech Research Development,* 68, 2449–2472. https://doi.org/10.1007/s11423-020-09767-4

[7] Gao, C., Li, Z., & Zheng, L. (2023). Develop and validate a scale to measure primary and secondary teachers' digital teaching competence. *Education and Information Technologies, 29*, 10763-10789. https://doi.org/10.1007/s10639-023-12228-z

[8] Göçen Kabaran, G., & Uşun, S. (2021). Dijital materyal tasarımı yeterlikleri ölçeği (DMTYÖ): Bir ölçek geliştirme çalışması. *Eğitim Teknolojisi Kuram ve Uygulama, 11*(2), 281-307. https://doi.org/10.17943/etku.864296Ok

[9] Hattie, J. (2012). *Visible learning for teachers: Maximizing impact on learning*. Routledge.

[10] Instefjord, E., & Munthe, E. (2016). Preparing pre-service teachers to integrate technology: an analysis of the emphasis on digital competence in teacher education curricula. *European Journal of Teacher Education*, *39*(1), 77-93. https://doi.org/10.1080/02619768.2015.1100602

[11] Karalar, H., Aslan Altan, B., & Yorulmaz, A. (2025). Adaptation of the Digital Teaching Competence Scale (DTCS) to Turkish: Reliability, validity, measurement invariance and latent mean difference tests. *Mehmet Akif Ersoy University Journal of Education Faculty*, 76, 68-85. https://doi.org/10.21764/maeuefd.1715751

[12] Kaya, E., & Bayram, H. (2021). Utilization of the research compliance matrix in educational research design and evaluation: A design based research. *International Journal of Education Technology & Scientific Researches*, *6*(15), 887-944. http://dx.doi.org/10.35826/ijetsar.325

[13] Koehler, M., & Mishra, P. (2009). What is technological pedagogical content knowledge (TPACK)?. *Contemporary issues in technology and teacher education*, *9*(1), 60-70. https://www.learntechlib.org/primary/p/29544/.

[14] Laurillard, D. (2013). Teaching as a design science: Building pedagogical patterns for learning and technology. Routledge.

[15] Özdamar, K. (2002). Paket programlarla istatistiksel veri analizi-1 Kaan Kitapevi.

[16] Scherer, R., Siddiq, F., & Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*, *128*, 13-35. https://doi.org/10.1016/j.compedu.2018.09.009

[17] Selwyn, N. (2021). Education and technology: Key issues and debates. Bloomsbury Publishing.

[18] Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics.* Pearson.

[19] Tondeur, J., van Braak, J., Ertmer, P. A. Ottenbreit-Leftwich, A. (2016). Understanding the relationship between teachers' pedagogical beliefs and technology use in education: A systematic review of qualitative evidence. *Education Tech Research Development,* 65, 555–575. https://doi.org/10.1007/s11423-016-9481-2

[20] Voogt, J., Fisser, P., Pareja Roblin, N., Tondeur, J., & van Braak, J. (2013). Technological pedagogical content knowledge-a review of the literature. *Journal of Computer Assisted Learning*, *29*(2), 109-121. https://doi.org/10.1111/j.1365-2729.2012.00487.x

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Analytical, Bayesian, and Machine-Learning Inference of Radioactive Decay-Chain Parameters: A Reproducible Accuracy–Speed–Uncertainty Benchmark on the Mo-99/Tc-99m System

Cafer Mert Yeşilkanat *,[1], Serkan Akkoyun[2]

[a]*Artvin Çoruh University, Department of Mathematics and Science Education, 08100, Artvin, Türkiye*
[b]*Sivas Cumhuriyet University, Department of Physics, Sivas, Türkiye*

**(cmyesilkanat@artvin.edu.tr) Email of the corresponding author*

Abstract – Recovering decay-chain parameters from noisy activity measurements represents a canonical inverse problem in dosimetry, nuclear-medicine quality control, and nuclear forensics. Although deterministic optimization, Bayesian sampling, and machine learning (ML) are widely deployed in nuclear applications, a head-to-head benchmark under a unified uncertainty-quantification protocol against an exact analytic ground truth remains lacking. Here, we establish such a benchmark using the medically important ^{99}Mo → ^{99m}Tc generator system governed by Bateman equations. From Poisson-corrupted daughter-activity time series, parent and daughter half-lives were inferred alongside intrinsic uncertainty estimates via weighted nonlinear least squares (NLS), adaptive Metropolis Markov-chain Monte Carlo (MCMC), and an amortized random-forest ensemble (ML). Across 120 independent held-out scenarios, the ML ensemble achieved the highest point accuracy (relative RMSE of 12.4% and 11.6% for parent and daughter half-lives, respectively), effectively suppressing error tails observed in MCMC (26.7%/20.1%) and NLS (32.7%/29.3%). Conversely, MCMC yielded superior uncertainty calibration (empirical 68% coverage of 67–70%; scalar calibration error of 0.017 across nominal levels swept from 50% to 95%), outperforming NLS (0.038) and the systematically conservative ML ensemble (0.141). Computational analysis showed that NLS provided the lowest per-query latency (median 5.3 ms), whereas amortized ML (12.1 ms) surpassed MCMC (307 ms) in efficiency over large datasets. Validation at the exact isotopic operating point confirmed unbiased parameter recovery within 1.2% (parent) and 2.4% (daughter) of true values, delineating an accuracy–speed–uncertainty trade-off without a universal winner.

Keywords *– Radioactive decay chains; inverse problem; Bayesian inference; Markov-chain Monte Carlo; machine learning*

I. INTRODUCTION

Radioactive decay chains underpin a broad range of quantitative nuclear applications. Estimating the parameters that govern such chains (chiefly the decay constants or, equivalently, the half-lives of the members) from time-resolved activity or count measurements is an *inverse problem* that recurs in internal dosimetry, radiopharmaceutical quality assurance, environmental radioactivity, spent-fuel and waste characterization, and nuclear forensics. The forward model is well understood: for a linear chain it is given in closed form by the Bateman equations (Bateman, 1910). The inverse task, however, is complicated by counting (Poisson) noise, by nuisance normalization factors such as detector efficiency and initial source strength, and by the partial ill-conditioning that arises when the sampling window does not fully resolve the slowest and fastest components of the chain.

Three broad computational paradigms are used to solve such inverse problems, each with a distinct profile. The first is **deterministic optimization**: nonlinear least-squares fitting of the forward model, which is fast and familiar but supplies only asymptotic, locally linearized uncertainty estimates. The second is **Bayesian sampling**: Markov-chain Monte Carlo (MCMC) explores the full posterior distribution (Haario et al., 2001; Hastings, 1970; Metropolis et al., 1953), delivering principled credible intervals at a substantially higher

computational cost. The third, increasingly prominent paradigm is **data-driven machine learning** (ML): a regressor is trained once on simulated chain realizations and thereafter maps observations to parameters almost instantly—an instance of amortized, simulation-based inference (Cranmer et al., 2020). Machine learning has expanded rapidly across nuclear physics in the past decade (Boehnlein et al., 2022), including for the evaluation and prediction of nuclear reaction data (Vicente-Valdez et al., 2021).

Despite the maturity of each paradigm in isolation, the literature offers surprisingly few controlled, ground-truth-anchored comparisons of all three on a single decay-chain inverse problem, evaluated under a *common uncertainty-quantification protocol*. Studies typically adopt one paradigm and report point accuracy; when uncertainty is reported at all, it is rarely audited for calibration (i.e., whether stated 68% intervals actually contain the truth 68% of the time). This leaves practitioners without a principled basis for selecting a method under the accuracy, latency, and reliability constraints of a given application.

This study addresses that gap. Our contributions are fourfold. (1) We build a fully reproducible three-paradigm benchmark (NLS, MCMC, and an amortized ML ensemble) for decay-chain half-life estimation, anchored to the medically important ^{99}Mo → ^{99m}Tc generator (Gupta et al., 2025), whose Bateman solution furnishes exact ground truth. (2) We evaluate all three paradigms under a uniform UQ protocol, reporting not only error but interval coverage and a full calibration curve. (3) We quantify the resulting accuracy–speed–uncertainty trade-off across 120 independent scenarios and validate the framework at the exact isotope point. (4) We release the complete, seed-controlled pipeline as an open template that can be extended to evaluated nuclear data and to longer, branching chains. To our knowledge, this is the first head-to-head, calibration-audited benchmark of these three paradigms on a canonical nuclear decay-chain inverse problem.

II. Materials and Method

A. Physical system and forward model

The two-member decay chain ^{99}Mo → ^{99m}Tc → ^{99}Tc, which serves as the foundation for the most widely used diagnostic radionuclide generator in nuclear medicine (Gupta et al., 2025), was investigated. In alignment with established evaluated nuclear data, half-lives of 65.976 h for the parent ^{99}Mo and 6.0067 h for the metastable daughter ^{99m}Tc were adopted. While the daughter half-life represents the evaluated standard (Morillon et al., 2012), the parent value falls within a tight margin between current evaluations—specifically 65.932(5) h in NUBASE2020 (Kondev et al., 2021) and 2.7479(6) d ≈ 65.95 h in DDEP (Morillon et al., 2012). This sub-0.1% disparity remains negligible compared to the 1–30% inference errors analyzed in this benchmark and inherently cancels across method comparisons because all three algorithms are provided with identical simulated data.

The system exhibits transient equilibrium, wherein the parent half-life exceeds that of the daughter without overwhelming dominance; consequently, the daughter activity accumulates to a maximum at $t = \ln(\lambda_B/\lambda_A)/(\lambda_B - \lambda_A) = 22.9$ h before decaying at the effective decay rate of the parent. The ground-state isotope ^{99}Tc possesses a half-life of $2.115(11) \times 10^5$ a (Morillon et al., 2012), rendering it effectively stable over the 72-h observation window; thus, the chain was truncated at the daughter stage for all practical purposes. Denoting λ_A and λ_B as the parent and daughter decay constants, respectively, the time-dependent daughter activity resulting from an initially pure parent source was formulated via the Bateman solution:

$$A_B(t) = \frac{N_0 \lambda_A \lambda_B}{\lambda_B - \lambda_A} [\exp(-\lambda_A t) - \exp(-\lambda_B t)] \quad (1)$$

where N_0 denotes an overall normalization factor encompassing the initial parent population, detector efficiency, and the branching ratio $p = 87.6(19)\%$ for ^{99}Mo decays populating the metastable isomer (Morillon et al., 2012). Because p enters the expression multiplicatively, it was absorbed into N_0, preventing parameter estimation bias. Due to the intrinsic symmetry of Equation (1) under rate constant exchange (where numerator sign inversions are compensated by denominator inversions), the inverse problem admits a label-swap degeneracy. To resolve this ambiguity, the physical ordering constraint $t_{1/2}(A) > t_{1/2}(B)$ was enforced, ensuring that the parent remains the longer-lived species.

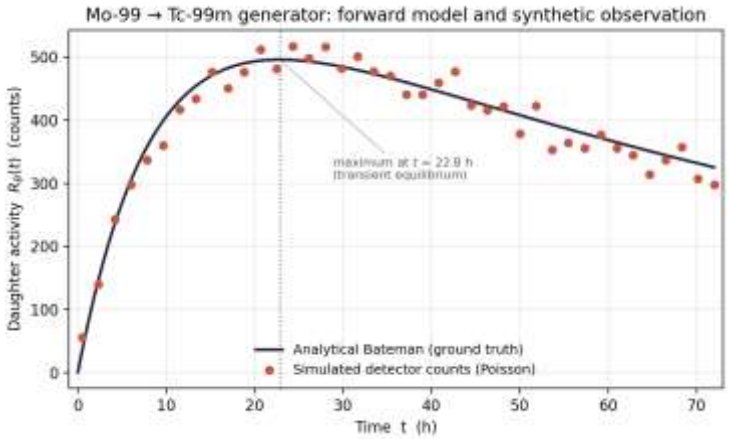


Figure 1. The two-member ^{99}Mo → ^{99m}Tc decay system and its corresponding Bateman forward model. The solid curve illustrates the analytical daughter activity evaluated at the adopted benchmark half-lives, whereas the data points represent a single Poisson-corrupted realization of the 40-point, 72-h counting scheme used per observation. The characteristically observed activity peak near 23 h, followed by exponential decay governed by the parent's effective rate, exemplifies the transient equilibrium regime.

B. Synthetic observation model

Each synthetic experiment was constructed by sampling a parent half-life, a daughter half-life, and an overall normalization factor from broad design spaces configured to encompass the operational domain of the generator. Denoting $U(a, b)$ as the uniform distribution on the interval $[a, b]$, the half-lives were sampled uniformly, whereas the normalization factor was sampled log-uniformly. The training corpus utilized parameter ranges of $t_{1/2}(A) \sim U(35{,}100)$ h, $t_{1/2}(B) \sim U(2.5{,}13)$ h, and $\log_{10} N_0 \sim U(3.2{,}5.4)$ (corresponding to N_0 values between approximately 1.6×10^3 and 2.5×10^5). The test dataset was specified as $t_{1/2}(A) \sim U(45{,}90)$ h, $t_{1/2}(B) \sim U(3.5{,}12)$ h, and $\log_{10} N_0 \sim U(3.5{,}5.2)$ (yielding N_0 values between approximately 3.2×10^3 and 1.6×10^5). The generator pipeline additionally incorporated an explicit boundary constraint ($t_{1/2}(B) < t_{1/2}(A) - 8$ h for training and $t_{1/2}(B) < t_{1/2}(A) - 10$ h for testing); however, because this guard remained inactive across the prescribed parent half-life bounds, the cited daughter ranges represent the effective operational limits.

Because N_0 serves as a multiplicative scaling factor in the Bateman formulation rather than a direct count metric, its physical significance lies in establishing the signal magnitude. Across the test set, the daughter activity profiles peaked between approximately 20 and 2,000 counts, corresponding to a maximum Poisson signal-to-noise ratio ranging from roughly 4 to 45. The test parameter ranges were strictly nested within the training domain, establishing an in-distribution evaluation framework whose methodological implications for the machine learning paradigm are addressed in Section 4. Uncorrupted daughter-activity profiles were evaluated on a deterministic temporal grid of 40 uniformly spaced interval points spanning 0.5 h to 72 h (emulating a standardized three-day counting procedure), read as expected counts, and corrupted with Poisson noise to reflect realistic gamma-ray spectroscopy statistics. This procedure generated a noisy count-versus-time vector for each scenario, providing the sole observation input for each inference methodology; ground-truth half-lives were withheld strictly for evaluation. A training corpus comprising 12,000 simulated curves was generated for the amortized machine learning ensemble, and an independent test set of 120 scenarios, generated via a distinct pseudo-random stream, was reserved for comparative benchmarking to ensure that all algorithms were evaluated on identical datasets.

C. Interfence models

Deterministic model: Weighted nonlinear least squares (NLS). The Bateman forward model in Equation (1) was fitted to each noisy curve via bounded nonlinear least squares, minimizing the Poisson-weighted residual sum of squares through the trust-region-reflective algorithm as implemented in SciPy (Virtanen et al., 2020). Both half-lives and the overall normalization factor were jointly optimized within a wide physical parameter space defined as $t_{1/2}(A) \in [10{,}200]$ h and $t_{1/2}(B) \in [1{,}30]$ h. Each residual was

weighted by $\sqrt{\max(N, 1)}$ to align the effective covariance matrix with the Poisson counting noise model. Parameter uncertainties were derived from the linearized covariance matrix computed at the global optimum, yielding asymptotic, locally Gaussian error bounds characteristic of deterministic optimization. Successful numerical convergence was achieved for all 120 test scenarios.

Bayesian model: Adaptive Metropolis MCMC. Scale-invariant (flat-in-log) prior distributions were assigned to both half-lives, bounded within $t_{1/2}(A) \in (10{,}200)$ h and $t_{1/2}(B) \in (1{,}25)$ h and subject to the physical ordering constraint $t_{1/2}(B) < t_{1/2}(A)$. An unbounded scale-invariant prior was placed on the normalization parameter. Posterior distributions conditioned on the exact Poisson likelihood function were sampled using a random-walk Metropolis scheme incorporating an adaptive proposal covariance in log-parameter space (Haario et al., 2001; Hastings, 1970; Metropolis et al., 1953). Each Markov chain was initialized at the corresponding NLS optimum and executed for 9,000 iterations, with the initial 3,000 steps discarded as burn-in to adapt the proposal distribution, resulting in 6,000 retained posterior samples. Posterior means were designated as point estimates, and central posterior intervals were constructed as credible intervals in accordance with established Bayesian workflows (Gelman et al., 1995).

Data-driven model: Amortized random-forest ensemble (ML). A random forest ensemble comprising 150 decision trees, constrained to a minimum of three samples per leaf node (Breiman, 2001; Pedregosa et al., 2011), was trained to map an amplitude-invariant temporal descriptor directly to the logarithmic half-lives, which were subsequently exponentiated. The input descriptor was formulated by concatenating the count profile normalized to unit sum with a maximum-normalized log-count profile. Model training was performed as a one-time offline step on the 12,000-curve synthetic dataset; subsequently, parameter inference for unseen profiles was executed on a millisecond scale. Predictive uncertainty was estimated from the variance among individual tree outputs and propagated to the physical half-life scale using the delta method, providing an ensemble-based measure of epistemic uncertainty analogous to deep ensembles (Lakshminarayanan et al., 2017). This architecture exemplifies amortized, simulation-based inference (Cranmer et al., 2020), wherein the computational burden of training is incurred once and amortized across all subsequent queries.

D. Evaluation protocol, validation, and reproducibility

Statistical performance across the shared 120-scenario evaluation test set was quantified for each half-life parameter and inference paradigm using a comprehensive battery of metrics: relative root-mean-square error (RMSE), median absolute relative error (MedAE, chosen for its robustness against extreme ill-conditioned outliers), relative bias, empirical coverage of the nominal 68% uncertainty interval, median relative interval half-width (defined as the reported 1σ uncertainty divided by the ground-truth half-life), and wall-clock execution time per query (summarized by its median value). To evaluate uncertainty calibration beyond a single confidence threshold, the nominal coverage level was swept continuously from 50% to 95% using the native interval construction rule of each paradigm (namely, asymptotic Gaussian intervals for NLS, central posterior credible intervals for MCMC, and ensemble-dispersion Gaussian intervals for ML). Empirical coverage at each nominal level was pooled over both decay parameters (yielding 240 evaluation checks per level) to construct full calibration curves, which were subsequently summarized by a scalar calibration error representing the mean absolute deviation between empirical and nominal coverage.

As a rigorous isotopic validation, parameters were fixed to the adopted $^{99}\mathrm{Mo}/^{99\mathrm{m}}\mathrm{Tc}$ baseline with a normalization of $N_0 = 6 \times 10^4$, and the full inference pipeline was executed across 40 independent Poisson noise realizations to evaluate deviation from the true physical constants. The entire computational workflow was rendered fully deterministic through fixed random seed[1] assignment. The complete computational pipeline was implemented in Python using standard scientific libraries, specifically NumPy (Harris et al., 2020), SciPy (Virtanen et al., 2020), and scikit-learn (Pedregosa et al., 2011). All reported numerical outputs were generated directly by the underlying execution script without manual post-processing.

[1] Utilizing a master seed of 20260706 alongside dedicated sub-seeds for dataset generation, test suite initialization, validation runs, and Markov chain execution

Reported wall-clock latencies reflect relative operational performance under fixed hardware and implementation constraints.

III. RESULTS

Figure 1 illustrates the forward model evaluated at the adopted $^{99}Mo/^{99m}Tc$ half-lives alongside a single Poisson-corrupted realization of the temporal counting design. Furthermore, Figure 2 depicts a representative inference performed on this identical curve. Although all three computational paradigms reproduced the observed count profile with visual indistinguishability (the respective reconstructions strictly overlying both the data and each other), the underlying inferred parameters deviated appreciably from the ground truth. Specifically, the recovered parent half-lives were 54.9 h (NLS), 56.6 h (MCMC), and 57.9 h (ML) against a true value of 65.98 h, whereas the daughter half-lives were 6.86 h, 6.70 h, and 6.48 h, respectively, against a true 6.01 h. These estimation errors exhibited strong anti-correlation; the parent half-life was consistently underestimated whenever the daughter half-life was overestimated. This phenomenon serves as the empirical signature of the partial ill-conditioning inherent to the inverse problem: over the 72-h observation window, the slow-decaying parent component was only weakly constrained, allowing a continuous family of $\left(t_{1/2}(A), t_{1/2}(B)\right)$ pairs to generate nearly identical activity profiles. Consequently, a solitary high-quality curve fit does not guarantee a well-determined parameter pair, necessitating the use of aggregate error, coverage, and calibration statistics (rather than qualitative visual fit) as the rigorous basis for methodological comparison. While the three methods converged on the gross morphological features of the solution, they diverged significantly in point accuracy, predictive uncertainty reliability, and computational latency.

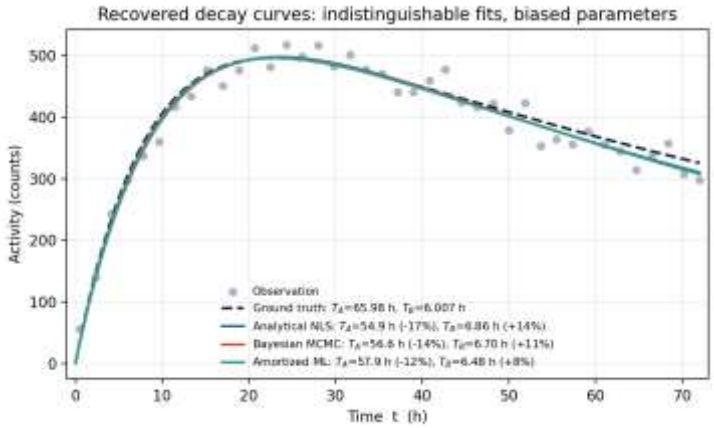


Figure 2. Representative parameter inference on a single Poisson-corrupted daughter-activity curve simulated at the adopted $^{99}Mo/^{99m}Tc$ half-lives. Despite yielding visually indistinguishable reconstructions that tightly track the analytic ground truth, the inferred parent and daughter half-lives deviate by 12-17% and 8-14%, respectively, empirically illustrating the inherent ill-conditioning of this inverse problem.

The empirical results of the three-paradigm benchmark across 120 held-out evaluation scenarios are collected in Table 1. Regarding global point accuracy as measured by root-mean-square error (RMSE), a clear hierarchical ordering emerged (Figure 3): the machine learning (ML) ensemble achieved the lowest estimation error (12.4% and 11.6% relative RMSE for the parent $t_{1/2}(A)$ and daughter $t_{1/2}(B)$ half-lives, respectively), followed by Bayesian MCMC (26.7% and 20.1%), whereas deterministic NLS exhibited the highest overall error (32.7% and 29.3%). However, this accuracy gap narrowed substantially when evaluated under the median absolute relative error (MedAE), where values ranged strictly between 5.7% and 10.1% across all paradigms and both physical parameters. This divergence indicates that the elevated RMSE values in NLS and MCMC were predominantly driven by a sparse tail of severe, ill-conditioned scenarios—specifically cases where the observation window provided weak constraints on the longer-lived parent component. The ML ensemble suppressed these tail errors most effectively, providing the primary basis for its RMSE superiority. Notably, this ranking was not preserved for typical cases: for the parent half-life, the median absolute error of NLS (8.9%) was lower than that of MCMC (10.1%), reversing their RMSE ranking and confirming that MCMC's aggregate RMSE advantage stemmed from outlier

suppression rather than typical-case accuracy. Relative parameter bias remained minor and systematically positive across all methods (under approximately +6.5% in all cases), demonstrating the absence of severe systematic distortion.

Table 1. Comparative benchmark of the three inference paradigms across 120 held-out scenarios. Numerical values denote relative percentages unless otherwise specified; A and B represent the parent (^{99}Mo) and daughter (^{99m}Tc) half-lives, respectively. Empirical 68% coverage reflects the fraction of cases where nominal intervals contained the ground truth (optimal value closest to 68%). Interval half-width denotes the median 1σ uncertainty normalized to the true half-life. Bold values highlight optimal performance per metric (excluding non-quality metrics). Per-query execution time excludes the one-time ML training overhead (79.3 s)

Metric	NLS	MCMC	ML
RMSE, t½(A) [%]	32.7	26.7	**12.4**
RMSE, t½(B) [%]	29.3	20.1	**11.6**
Median AE, t½(A) [%]	8.9	10.1	**7.2**
Median AE, t½(B) [%]	7.3	6.6	**5.7**
Bias, t½(A) [%]	+1.2	+6.2	**+0.7**
Bias, t½(B) [%]	+6.5	+4.0	**+1.3**
68% coverage, t½(A) [%]	72	**67**	88
68% coverage, t½(B) [%]	74	**70**	89
Interval half-width (1σ), t½(A) [%]	17.7	18.5	18.7
Interval half-width (1σ), t½(B) [%]	11.4	11.3	14.3
Median time per query [ms]	**5.3**	307	12.1
Calibration error, 50–95% sweep [–]	0.038	**0.017**	0.141

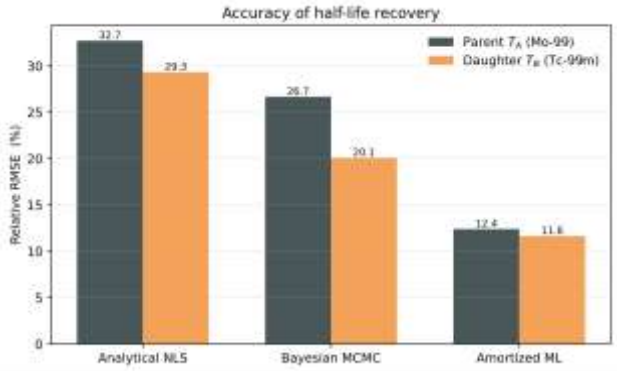


Figure 3. Relative RMSE for parent $t_{1/2}(A)$ and daughter $t_{1/2}(B)$ half-life recovery across the evaluated inference paradigms. The amortized ML ensemble yielded the lowest estimation error on both parameters, whereas Bayesian MCMC exhibited intermediate performance and NLS recorded the highest overall RMSE.

The evaluated paradigms exhibited a operational latency spread spanning a factor of approximately 58 per single-query inference. Deterministic weighted nonlinear least squares (NLS) demonstrated the lowest computational overhead, recording a median processing time of 5.3 ms per curve. Conversely, Bayesian MCMC represented the most computationally intensive approach at 307 ms per curve, reflecting the evaluation of 9,000 likelihood states required to sufficiently sample the joint posterior space. The pre-trained random-forest ensemble (ML) labeled an individual curve in 12.1 ms while incurring a one-time offline training investment of 79.3 s. When evaluating cumulative computational expenditure against MCMC, the amortized ML approach became more economical after processing approximately 270 curves (amortizing 79.3 s against a net savings of 295 ms per curve), with its efficiency advantage expanding monotonically thereafter. Crucially, amortized ML did not surpass NLS in raw execution speed at any operational scale; its per-query latency remained more than twice that of NLS (12.1 ms versus 5.3 ms), and the initial training phase further widened this performance gap. Consequently, NLS represents the optimal choice for throughput-constrained or embedded real-time applications; amortized ML constitutes the most economical paradigm for high-throughput batch analyses requiring superior point accuracy; and MCMC is strictly reserved for scenarios where rigorous posterior distribution fidelity justifies the computational overhead (which, at 307 ms per query, remains negligible for conventional offline laboratory workflows).

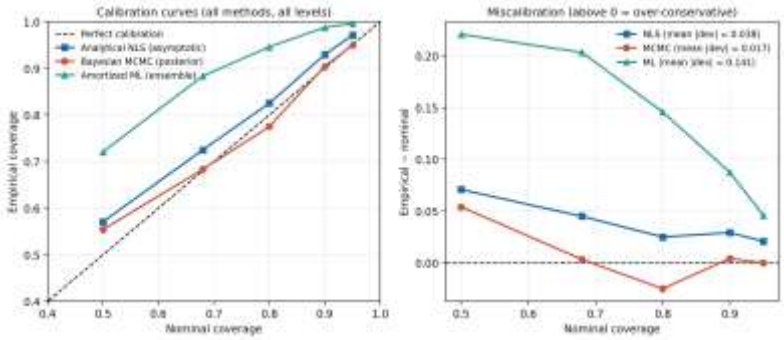


Figure 4. Uncertainty calibration of all three paradigms, each swept over the same nominal levels using its own native interval rule, with coverage pooled over both half-lives (240 interval checks per level). Left: empirical versus nominal coverage; the dashed line is perfect calibration. Right: signed miscalibration (empirical − nominal), where zero is perfect and positive values are over-conservative. MCMC tracks the nominal level almost exactly (calibration error 0.017), NLS is mildly conservative (0.038), and the ML ensemble is systematically over-conservative (0.141), most strongly at low nominal levels.

Point accuracy metrics provide an incomplete depiction of model performance; a robust parameter estimate must also furnish reliable predictive uncertainties. In terms of uncertainty calibration, the performance hierarchy reversed relative to point estimation accuracy. At the nominal 68% confidence level, Bayesian MCMC demonstrated superior calibration, achieving empirical coverage rates of 67% and 70% for the parent and daughter half-lives, respectively (closely matching the theoretical target). Deterministic NLS exhibited slight over-coverage (72% and 74%). Conversely, the amortized ML ensemble proved markedly conservative, yielding empirical coverages of 88% and 89%, indicating that its inter-tree dispersion systematically overstated true parameter uncertainty. Sweeping the nominal coverage threshold from 50% to 95% using each paradigm's native interval formulation (Figure 4) confirmed that this miscalibration represented a fundamental systemic property rather than a single-level artifact.

Bayesian MCMC tracked ideal calibration almost perfectly, recording a scalar calibration error (defined as the mean absolute deviation between empirical and nominal coverage) of 0.017, whereas NLS recorded a minor positive deviation (0.038) and the ML ensemble displayed substantial over-coverage throughout (0.141). The predictive deviation of the ML ensemble reached its maximum at low nominal thresholds ($+0.22$ at the 50% level) and attenuated toward higher levels ($+0.05$ at the 95% level), revealing an ensemble spread that was excessively broad in the central bulk of the predictive distribution rather than merely mis-scaled. Because relative interval half-widths remained comparable across all three methodologies (median 1σ relative half-widths spanning 11%--19%), the observed coverage disparities directly reflect intrinsic calibration fidelity rather than variations in interval magnitude. Figure 5 projects the three paradigms simultaneously across the accuracy, computational speed, and coverage axes, rendering the multi-objective trade-off explicit.

As an external empirical validation, physical parameters were fixed strictly to the baseline $^{99}\mathrm{Mo}/^{99\mathrm{m}}\mathrm{Tc}$ half-lives with a constant normalization factor ($N_0 = 6 \times 10^4$), and the full inference pipeline was executed across 40 independent Poisson noise realizations (Table 2). The mean recovered parent half-life deviated from the ground-truth value of 65.976 h by -1.2% for NLS, $+0.1\%$ for Bayesian MCMC, and -0.9% for the ML ensemble, whereas the mean daughter half-life shifted from the accepted value of 6.0067 h by $+0.9\%$, $+0.5\%$, and $+2.4\%$, respectively. Realization-to-realization variance was substantial, spanning 6.9%--9.0% for the parent half-life and 5.7%--7.7% for the daughter half-life. Consequently, none of the observed mean deviations reached statistical significance at the 5% level ($|t| \leq 1.9$ with $n = 40$), confirming the absence of systematic estimation bias at the primary operational point of interest. These results verify that the benchmark pipeline is tightly anchored to the physical decay system. Furthermore, this validation provides a practical benchmark for experimental design: a single 72-h counting campaign under this observational architecture determines the parent half-life to within approximately $\pm$7%--9% (1σ), directly accounting for the aggregate error magnitudes reported in Table 1.

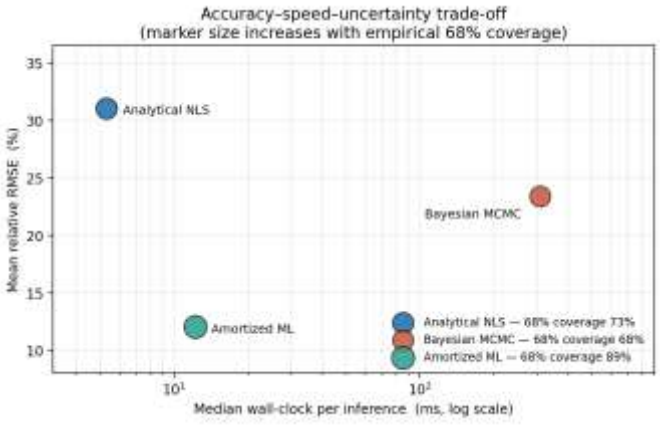


Figure 5. The accuracy–speed–uncertainty trade-off. Each marker is one paradigm, placed by median wall-clock time per query (horizontal, logarithmic) and by the mean relative RMSE over the two half-lives (vertical), with marker area scaled by empirical 68% coverage. No single paradigm is best on all three axes: ML leads on accuracy, MCMC on calibration, and NLS on raw per-query cost, with ML becoming the economical choice only once its training cost has been amortized over several hundred curves.

Table 2. Validation at the adopted baseline $^{99}\mathrm{Mo}/^{99\mathrm{m}}\mathrm{Tc}$ half-lives across 40 independent Poisson noise realizations. Table entries denote the mean $\pm$ sample standard deviation of recovered half-lives in hours, with relative percentage deviations from the true physical constants provided in parentheses. None of the mean deviations exhibited statistical significance at the 5% level ($|t| \leq 1.9, n = 40$); true values are explicitly designated for reference.

Method	**t½ ^{99}Mo [h]**	**t½ $^{99\mathrm{m}}$Tc [h]**
True value	**65.976**	**6.0067**
NLS	65.20 ± 4.52 (−1.18%)	6.060 ± 0.345 (+0.89%)
MCMC	66.07 ± 4.80 (+0.14%)	6.038 ± 0.351 (+0.53%)
ML	65.36 ± 5.92 (−0.93%)	6.148 ± 0.464 (+2.36%)

IV. DISCUSSION

The multi-paradigm evaluation reveals a well-defined accuracy–speed–uncertainty trade-off (Figure 5) wherein no single computational paradigm dominates across all operational metrics. Each methodology exhibits optimal performance along a distinct operational axis: the amortized ML ensemble excels in point accuracy, Bayesian MCMC provides superior uncertainty calibration, and deterministic NLS delivers the lowest raw per-query latency. While amortized ML becomes the most economical choice for large-scale data processing, this cost advantage holds strictly relative to MCMC and only after the offline training overhead has been fully amortized. Consequently, algorithm selection in decay-chain inverse problems must be dictated by the binding operational constraints of the target application rather than a presumed universal hierarchy.

The point accuracy advantage of the ML ensemble warrants a rigorous, transparent interpretation. This superior performance does not stem from the discovery of novel physical dynamics beyond the common Bateman formulation. Instead, the ensemble implicitly learns a prior distribution over the design space, regularizing predictions toward parameter regions that remain physically plausible under the training bounds. This implicit regularization effectively suppresses error tails in severe, ill-conditioned scenarios that otherwise inflate the RMSE of pointwise optimization routines. Two structural features of the experimental setup elucidate this mechanism. First, because the evaluation parameter space was strictly nested within the training domain, every test instance represented an in-distribution query. Second, because decision tree ensembles predict via training label aggregation, their outputs are inherently bounded by the training label range, preventing the extreme parameter excursions that pointwise fitters generate when likelihood surfaces exhibit near-flat geometries. Thus, while this accuracy advantage remains practically valuable, it is intrinsically conditional on the test distribution matching the training corpus; under distribution shift or extrapolation, this advantage would narrow or reverse. Furthermore, the empirical over-coverage of the ensemble intervals represents the direct counterpart of this property, indicating that

individual trees exhibit higher inter-model agreement than true parameter uncertainty warrants within the training regime.

Conversely, Bayesian MCMC occupies the complementary trade-off domain. Although computationally demanding, MCMC returns posterior credible intervals that achieve nominal coverage without relying on a pre-learned design distribution, utilizing merely a broad, weakly informative prior box. In safety-critical or regulatory settings requiring rigorously defensible uncertainty budgets (such as internal dosimetry, quality assurance, and nuclear forensics), this calibration fidelity represents the primary operational criterion, rendering the approximately 58-fold computational premium over NLS fully acceptable. Finally, NLS remains the pragmatic baseline when rapid point estimates with approximate error bounds are sufficient (particularly in real-time or resource-constrained embedded systems), offering mildly conservative coverage (scalar calibration error of 0.038) at zero marginal computational cost beyond the optimization fit.

Several methodological limitations bound the scope of these conclusions and define key vectors for subsequent research. First, the physical system examined here was restricted to a two-member decay chain featuring a single dominant daughter channel; longer, highly branched decay chains introduce additional label-swap and parameter degeneracies that complicate inverse inference. Second, the observation architecture assumed an idealized Poisson counting process evaluated on a uniform temporal grid, omitting real-world experimental complexities such as detector dead time, spectral background noise, and finite energy resolution. Furthermore, the ^{99}Mo branching ratio was absorbed into a composite normalization parameter rather than modeled explicitly. Third, because the evaluation parameter space was constructed as a strictly nested subset within the training domain, the reported machine learning metrics reflect in-distribution performance exclusively; predictive fidelity under distribution shift or domain extrapolation remains to be established. Fourth, the recorded wall-clock latencies are inherently implementation- and hardware-specific, depending directly on the chain length chosen for MCMC sampling and the forest size of the ensemble; these processing times should therefore be interpreted as relative rather than absolute benchmarks. Fifth, all numerical results were derived using a deterministic single master seed scheme, with reported aggregate metrics representing averages over 120 evaluation scenarios rather than repeated pipeline executions (thereby expecting minor third-digit variability across independent software runs). Sixth, the machine learning architecture was instantiated via a single representative random forest ensemble; alternative regression models and advanced neural simulation-based inference techniques may shift the accuracy–calibration equilibrium. Finally, this benchmark relied on synthetic observations generated from a known analytical forward model; while this setup enabled a rigorous audit against an exact physical ground truth, empirical validation against laboratory-measured spectroscopic data remains an essential next step.

V. CONCLUSION

We have presented a fully reproducible, ground-truth-anchored benchmark comparing deterministic optimization, Bayesian inference, and amortized machine learning for half-life recovery in the medically important ^{99}Mo/^{99m}Tc decay-chain system under a unified uncertainty-quantification protocol. The three paradigms delineate a clean accuracy–speed–uncertainty trade-off without a universal winner. The amortized ML ensemble achieves the highest point accuracy and, once its initial offline training investment is amortized, represents the more economical choice among uncertainty-rich methods. Bayesian MCMC yields superior posterior uncertainty calibration at an approximately 58-fold computational premium over NLS. Deterministic NLS delivers the lowest per-query operational latency, making it the fastest estimator at every scale. At the exact isotopic operating point, all three paradigms recover true physical half-lives to within 1.2% (parent) and 2.4% (daughter) on average without statistically significant bias, validating the benchmarking pipeline against the physical system. Ultimately, algorithm selection must be governed by whether point accuracy, computational latency, or rigorously calibrated uncertainty constitutes the binding operational constraint.

REFERENCES

Bateman, H., 1910. The solution of a system of differential equations occurring in the theory of radioactive transformations, in: Proc. Cambridge Philos. Soc. pp. 423–427.

Boehnlein, A., Diefenthaler, M., Sato, N., Schram, M., Ziegler, V., Fanelli, C., Hjorth-Jensen, M., Horn, T., Kuchera, M.P., Lee, D., Nazarewicz, W., Ostroumov, P., Orginos, K., Poon, A., Wang, X.N., Scheinker, A., Smith, M.S., Pang, L.G., 2022. Colloquium: Machine learning in nuclear physics. Rev. Mod. Phys. 94. https://doi.org/10.1103/RevModPhys.94.031003

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32.

Cranmer, K., Brehmer, J., Louppe, G., 2020. The frontier of simulation-based inference. Proc. Natl. Acad. Sci. 117, 30055–30062. https://doi.org/10.1073/pnas.1912789117

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1995. Bayesian data analysis. chapman and hall/cRc.

Gupta, R., Garg, S., Heston, T.F., 2025. Mo99-Tc99m generator, in: StatPearls [Internet]. StatPearls Publishing, p. https://www.ncbi.nlm.nih.gov/sites/books/NBK565854.

Haario, H., Saksman, E., Tamminen, J., 2001. An adaptive Metropolis algorithm. Bernoulli 7, 223–242. https://doi.org/10.2307/3318737

Harris, C.R., Millman, K.J., van der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M.H., Brett, M., Haldane, A., del Río, J.F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., Oliphant, T.E., 2020. Array programming with NumPy. Nature 585, 357–362. https://doi.org/10.1038/s41586-020-2649-2

Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57, 97–109.

Kondev, F.G., Wang, M., Huang, W.J., Naimi, S., Audi, G., 2021. The NUBASE2020 evaluation of nuclear physics properties *. Chinese Phys. C 45, 30001. https://doi.org/10.1088/1674-1137/abddae

Lakshminarayanan, B., Pritzel, A., Blundell, C., 2017. Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles, in: Guyon, I., Luxburg, U. Von, Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R. (Eds.), Advances in Neural Information Processing Systems. Curran Associates, Inc.

Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.H., Teller, E., 1953. Equation of state calculations by fast computing machines. J. Chem. Phys. 21, 1087–1092. https://doi.org/10.1063/1.1699114

Morillon, C., Bé, M.-M., Chechev, V., Egorov, A., 2012. Table de Radionucléides: 99Mo. http://www.lnhb.fr/nuclides/Mo-99_tables.pdf.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Vicente-Valdez, P., Bernstein, L., Fratoni, M., 2021. Nuclear data evaluation augmented by machine learning. Ann. Nucl. Energy 163, 108596. https://doi.org/10.1016/J.ANUCENE.2021.108596

Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., Vijaykumar, A., Bardelli, A. Pietro, Rothberg, A., Hilboll, A., Kloeckner, A., Scopatz, A., Lee, A., Rokem, A., Woods, C.N., Fulton, C., Masson, C., Häggström, C., Fitzgerald, C., Nicholson, D.A., Hagen, D.R., Pasechnik, D. V., Olivetti, E., Martin, E., Wieser, E., Silva, F., Lenders, F., Wilhelm, F., Young, G., Price, G.A., Ingold, G.L., Allen, G.E., Lee, G.R., Audren, H., Probst, I., Dietrich, J.P., Silterra, J., Webber, J.T., Slavič, J., Nothman, J., Buchner, J., Kulick, J., Schönberger, J.L., de Miranda Cardoso, J.V., Reimer, J., Harrington, J., Rodríguez, J.L.C., Nunez-Iglesias, J., Kuczynski, J., Tritz, K., Thoma, M., Newville, M., Kümmerer, M., Bolingbroke, M., Tartre, M., Pak, M., Smith, N.J., Nowaczyk, N., Shebanov, N., Pavlyk, O., Brodtkorb, P.A., Lee, P., McGibbon, R.T., Feldbauer, R., Lewis, S., Tygier, S., Sievert, S., Vigna, S., Peterson, S., More, S., Pudlik, T., Oshima, T., Pingel, T.J., Robitaille, T.P., Spura, T., Jones, T.R., Cera, T., Leslie, T., Zito, T., Krauss, T., Upadhyay, U., Halchenko, Y.O., Vázquez-Baeza, Y., 2020. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods 17, 261–272. https://doi.org/10.1038/s41592-019-0686-2

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Yakup Kadri Karaosmanoğlu'nun *Yaban* Romanında Kadın ve Çocuk

Havvaana KARADENİZ[1*]

[1]*Türkçe ve Sosyal Bilimler Eğitimi Bölümü, Eğitim Fakültesi, Kırşehir Ahi Evran Üniversitesi, Türkiye*

[*]*(havvaanakaradeniz@gmail.com)*

Özet – Bu çalışma, Yakup Kadri Karaosmanoğlu'nun 1932 yılında yayımlanan *Yaban* romanındaki kadın ve çocukların anlatıdaki konumlarını incelemek amacıyla gerçekleştirilmiştir. Alan yazınında genel olarak halk ile aydın arasındaki kopukluk üzerinden ele alınan romanda incelemeler, romanın erkek kişileri çevresinde toplanmıştır. Böyle bir yönelim, kadın ve çocukların köy tasvirini tamamlayan yardımcı kişiler olarak algılanmasına neden olmaktadır. Araştırmanın temel problemi, söz konusu kadın ve çocukların yalnızca köy hayatını tamamlayan yardımcı figürler olarak mı yer aldığı yoksa bu kişiler aracılığıyla köyün toplumsal yapısı, yoksulluk, emek, aile ilişkileri, şiddet ve savaş şartlarındaki korunmasızlığın ortaya konulup konulmadığıdır. Çalışma, nitel araştırma yaklaşımına dayalı bir doküman incelemesi olarak tasarlanmış; roman yakın okuma yöntemiyle çözümlenmiştir. Çözümlemede kadın ve çocukların yer aldıkları sahneler, konuşma imkânları, tasvir edilme biçimleri, olayların gelişimindeki işlevleri ve köy yapısı içindeki konumları incelenmiştir. İnceleme sonunda romandaki kadınların tek bir toplumsal grup oluşturmadığı, aralarında adlandırılma ve konuşma imkânı bakımından ayırt edici bir dereceleme bulunduğu ve söz sahibi olan kadınların dul kadınlar olduğu görülmüştür. Çocuklar bakımından ise romanın, çocukluğu, hayatın ayrı bir süreci olarak tanımadığı, okulun olmayışı ile çocuk emeğine duyulan ihtiyacın aynı yerde birleştiği tespit edilmiştir. Bunun dışında çocukların masumiyet sembolü olarak kullanılmadığı sonucuna varılmıştır. Savaş, mevcut yoksulluğu iyiden iyiye ağırlaştırırken kadın ve çocukları yaş ile cinsiyete göre farklılaşan tehditlerle karşı karşıya bırakmaktadır. Çalışmada söz konusu kişilerin yardımcı figür olmadığı, köyün toplumsal zeminini kuran ve anlatıcının bakışını sınayan bir ölçü işlevi gördükleri sonucuna ulaşılmıştır.

Anahtar Kelimeler – çocuk, kadın, Türk romanı, Yaban, Yakup Kadri Karaosmanoğlu

I. GİRİŞ

Yakup Kadri Karaosmanoğlu'nun 1932 yılında yayımlanan *Yaban*'ı, köyü ve köylüyü konu edinen gerçekçi Türk romanının erken halkalarından biridir. Özkırımlı, romanı Nabizade Nazım'ın *Karabibik*'i (1890) ile Ebubekir Hazım Tepeyran'ın *Küçük Paşa*'sından (1910) sonra gelen üçüncü örnek saymakta ve romanın ilk ikisinden ayrılan yanını da köylülüğü tarihî ve toplumsal bir mesele hâline getirmesinde bulmaktadır [2, s. 13]. Cumhuriyet dönemi Türk romanında köy meselesinin gelişimi bakımından da eser, sonraki köy romanlarının önünü açan bir dönemeç olarak değerlendirilmektedir [3], [4]. Roman, yayımlandığı yıllarda köylüyü yanlış tanıttığı ileri sürülmüş; 1960'tan sonra ise edebî ölçütlerle değerlendirilmeye başlanmıştır [2, ss. 13-18].

Alan yazınında *Yaban*, ağırlıklı olarak halk ile aydın arasındaki kopukluk üzerinden okunmuştur. Moran, romanı Cumhuriyet aydınının Anadolu'yla karşılaşmasındaki başarısızlık üzerinden değerlendirirken [5]; Akı, eseri yazarın kişiliği, üslubu ve dönem algısıyla birlikte ele almaktadır [6]. Hayber, romanı nesil çatışması çerçevesinde incelemekte [7]; Dündar ise metni ulus inşası ve millî bilinç kazandırma amacı bakımından çözümlemektedir [8]. Aydının yozlaşması [9] ile romanın kimlik ve tarih ilişkisi içindeki yeri [10] de aynı eksende tartışılmıştır. Böyle bir okuma ekseninin doğal sonucu, incelemelerin romanın erkek kişileri çevresinde toplanmasıdır. Mehmet Ali, Bekir Çavuş, Salih Ağa, muhtar, imam ve Şeyh Yusuf, Ahmet Celâl'in karşısına konulan köy dünyasının temsilcileri sayılmış; kadınlarla çocuklar bu dünyanın tasvirini tamamlayan yardımcı kişiler olarak anılmıştır.

Yakup Kadri'nin romanlarındaki kadın kişileri ele alan Baş ise yazarın romanlarında kadın karakterlerin anlatılma biçimlerini incelemiş [11]; Serdar, cinselliğin bu romanlardaki kuruluşunu ele almış [12]; Alver ise *Yaban*'daki kadınları doğayla kurdukları ilişki bakımından değerlendirmiştir [13]. Bununla birlikte söz konusu çalışmalarda kadınlar ile çocuklar birlikte ve emek, aile, hukuk ile savaş ölçütleri üzerinden ele alınmamıştır. Bu araştırma söz konusu boşluğu doldurmayı amaçlamaktadır.

Romanın kendisi, kadınlarla çocukları ikincil sayan bir dağılımı desteklememektedir. Zeynep Kadın, Cennet, Emine, Emeti Kadın ve Zehra metinde adı geçen beş kadın karakterdir. İsmail, Hasan ve hem kadın hem de çocuk olarak yer alan Zehra ise romandaki adı geçen üç çocuk karakteri oluşturmaktadır. Bu sayı, adı verilen erkek köylülerin sayısından az değildir. Romanın en uzun tutulmuş sahneleri de bu kişilere aittir. Ebesiz bir doğum [1, s. 91], bir tarla davası [1, ss. 88-90], bir zina baskını ve ardından gelen hülle [1, ss. 83-84, 123-125], görme engelli bir kız çocuğunun tarlada uğradığı istismar [1, ss. 98-100], on bir yaşında bir çobanın sırtından vurulup ölmesi [1, ss. 234-235] ve işgal askerleri önünde bir kadının ayağa kalkıp konuşması [1, ss. 246-247]; sahne düzeyinde kurulmuş, özetlenmemiş bölümlerdir. Romanda öldürülen tek kişi bir çocuktur. Romanın son cümlesi ise defterin bir kadına teslim edilmesiyle kurulur [1, s. 262].

Romanın anlatım biçimi bu araştırma için özel bir önem taşımaktadır. Okur, köyü ve köylüleri savaşta bir kolunu kaybetmiş olan Ahmet Celâl'in tuttuğu notlardan tanır. Onun İstanbul'daki yaşamı, eğitim düzeyi, edebî birikimi ve subay kimliği, karşılaştığı kişileri nasıl tanımlayıp adlandırdığını etkilemektedir. Köylü bedenlerini şekilsiz, hasta ve hayvan benzetmeleriyle tasvir etmesi; kadınlara bakarken zaman zaman arzu, iğrenme ve kültürel üstünlük duygusu arasında gelip gitmesi; bu konum ve kimliğinden bağımsız değildir. Anlatı kuramında bir metnin değer dizgesi ile anlatıcının hükümlerinin çakışmayabileceği, dolayısıyla anlatıcının güvenilirliğinin ayrıca sınanması gerektiği belirtilmektedir [14], [15]. Odaklanmanın kimin bakışıyla kurulduğu sorusu da aynı ayrımı gerektirmektedir [16], [17]. Bu sebeple Ahmet Celâl'in kadın, çocuk ve köylülere dair hükümleri doğrudan doğruya Yakup Kadri'nin görüşü olarak değerlendirilmeyecektir.

Bu çerçeve, çalışmanın, *Yaban* romanındaki kadın ve çocukların, köy yaşamının tasvirini tamamlayan yardımcı kişiler olarak mı yer aldığı yoksa bu kişiler aracılığıyla köyün sosyal yapısı, yoksulluk, savaş, emek, eğitim, aile ilişkileri, şiddet ve korunmasızlık gibi konuların mı ortaya konduğu araştırma problemini teşkil etmektedir. Çalışmanın tezi, bu kişilerin her iki işlevi de yerine getirdiğidir. Onlar köyün emek düzenini, yoksulluğunu, hukuk dışı kalışını ve savaş şartlarındaki savunmasızlığını taşıyan somut zemini inşa etmekte; aynı zamanda anlatıcının bakışının sınırlarını ölçmeye yarayan bir ölçüt hâline gelmektedir. Ahmet Celâl bu kişileri kâh Anadolu'nun sefaletini kanıtlayan bir belgesi kâh kendi arzusu ve kurtarıcılık tasarısının nesnesi saymaktadır.

II. MATERYAL VE YÖNTEM

A. Araştırmanın Modeli ve İnceleme Nesnesi

Bu çalışma, nitel araştırma yaklaşımına dayanan bir doküman incelemesi olarak tasarlanmıştır. Doküman incelemesi, araştırılması hedeflenen olgu hakkında bilgi içeren yazılı materyallerin sistematik biçimde çözümlenmesini kapsamaktadır [18], [19]. İncelenen temel doküman, Yakup Kadri Karaosmanoğlu'nun *Yaban* romanının Birikim Yayınları tarafından yayımlanan 1981 tarihli on beşinci basımıdır [1]. Metindeki bütün sayfa atıfları bu basıma aittir.

Çözümlemede yakın okuma yöntemi kullanılmıştır. İnceleme birimi olarak kadın ve çocuk karakterlerin yer aldığı sahne esas alınmış; her sahnede bu karakterlerin ad sahibi olup olmadığı, konuşma imkânı bulup bulmadığı, hangi işleve sahip olduğu, hangi fiziksel özelliklerle tasvir edildiği, olayların seyrini değiştirip değiştirmediği ve köy yapısı içinde hangi konumu işgal ettiği kaydedilmiştir. Ulaşılan veriler kadınlar, çocuklar ve savaş şartları olmak üzere üç ana sınıflandırma altında toplanmıştır.

B. Kavramsal Çerçeve

Çalışmada "kadın", "annelik" ve "çocukluk"; birbirinden bağımsız olarak değerlendirilmiştir. Toplumsal cinsiyetin, iktidar ilişkilerinin kurulduğu tarihsel bir çözümleme kategorisi olduğu kabulünden hareketle [20] kadın karakterler; eş, gelin, dul, anne, büyükanne, kız çocuğu veya bekâr kadın olarak aile içindeki farklı konum ve işlevleriyle ele alınmıştır. Türkiye bağlamında kadınlık deneyiminin hane içi güç

dengeleri, akrabalık ve mülkiyet ilişkileriyle birlikte biçimlendiğine ilişkin çözümlemeler [21], [22] bu ayrımı gerekli kılmaktadır. Annelik, bütün kadınların sahip olduğu doğal ve değişmez bir özellik gibi ele alınmamış; doğum, bakım, üretkenlik, emek, işi gücü, otorite, kayıp ve ekonomik zorunluluklarla kurulan bir ilişki olarak incelenmiştir. Nitekim Cennet ile Emine'nin çocuğu yoktur; Emeti Kadın ise kızının ölümünden sonra torununa bakma görevini üstlenen bir ninedir [1, s. 161].

Çocukluk da yaşla belirlenmemiştir. Çocukluğun evrensel ve değişmez bir hayat evresi olmayıp toplumsal şartlar içinde kurulduğu yönündeki yaklaşım [23], [24]; bu tercihin dayanağını oluşturmaktadır. Romanda on dört yaşındaki İsmail çocuk sayılırken [1, s. 40] aynı yaşlarda evlenmesi son derece olağan karşılanır. Salih Ağa'nın kambur oğlunun çocuk mu delikanlı mı olduğuna ise anlatıcı dahi karar veremez [1, s. 99]. Dolayısıyla çalışmada çocukluk, yaş sayısıyla değil; kişinin üstlendiği iş ve maruz kaldığı tehlikeyle birlikte ele alınmıştır.

III. BULGULAR

A. *Hane, Emek ve Kadın Otoritesi*

Yaban'da kadınların büyük bölümü ev, tarla, çeşme ve pazar çevresinde görülmektedir. Bu mekânlar aslında köyde yaşamsal üretimin sağlandığı, çocukların bakımının sağlandığı, erkeklerin askere gitmesinin ardından işlerin devam ettirildiği ve aile kurumunun korunduğu ve hepsinden önemlisi kadınların faaliyet gösterdiği yerlerdir. Romandaki kadın karakterler, emeğini ortaya koyuşu ile farklı işlev ve konuma sahiptir. Bunlardan Zeynep Kadın gibi kimileri adı, konuşmaları, gücü, hâkimiyeti ve kararlarıyla öne çıkarken Mehmet Ali'nin karısı, Mehmet Ali'nin kız kardeşleri, muhtarın karısı, Emine'nin halası ve Bekir Çavuş'un karısı gibi kimi kadınlar ise yalnızca aile içindeki konumları ve akrabalık sıfatlarıyla adlandırılmaktadır. Başka bir ifadeyle kadınların anlatıdaki ağırlığı ve görünme sıklıkları kadar sözlerinin etki ve sonuçları üzerinden de değerlendirilmelidir.

On iki yıldır dul olan ve ailenin "tek başı" sayılan Zeynep Kadın, köy aile geleneğinde otoritenin yetişkin erkeğe ait olmadığını somutlaştıran en belirgin karakterdir. Anlatıcı, onu "katı, sert ve mütevekkil" bulmakta; onda tabiata özgü bir güç sezmektedir [1, s. 56]. Bu tanıtım, kadının kişiliğini; yumuşaklık ve şefkat gibi anneliğin verdiği alışılmış özellikler üzerinden inşa etmemektedir. Ürettiği ürünleri pazara götürmesi, toprak işlerini düzenlemesi ve Salih Ağa'yla arazi meselesinde karşı karşıya gelmesi, onun hanedeki yerini bakım veren ve şefkat timsali bir anne konumunun ötesine taşımaktadır [1, s. 49, ss. 88-90]. Bu bakımdan mülkiyet ve üretim hakkında karar verebilen Zeynep kadın, köyün sosyo-ekonomik ilişkileri içinde de tanınan bir kadın karakterdir.

Zeynep Kadın'ın sahip olduğu bu otorite, kendiliğinden gelişen bir durum değildir. İsmail'i ağır biçimde dövmesi, sahip olunan bu otoritenin şiddetle düzenlendiğini göstermektedir [1, ss. 57-58]. Çocuğun çalışması ise yalnızca annenin sert mizacıyla izah edilemez. Yetişkin emeğine duyulan ihtiyaç, aile içinde zor kullanmayı ve şiddete başvurmayı meşrulaştırdığı bir düzene yerleşir. Hane içindeki gücü, İsmail için baskıya dönüşebilirken Salih Ağa ve işgal askerleri karşısında ise savunma imkânı hâline gelmektedir.

Kadın emeğinin sürekliliği, *"Zeynep Kadın, kızları ve gelinleriyle beraber, sabahtan akşama kadar, durmaksızın çalışıyor"* cümlesiyle açıkça belirtilmektedir [1, s. 119]. Cümlede Zeynep Kadın, adıyla ayrılmakta; kızlar ve gelinler ise akrabalık konumları içinde topluca anılmaktadır. Ev ve tarla işi, ürün hazırlama ve çocuk bakımı ayrı çalışma alanları değildir. Küçük çocukların tarlaya götürülmesi veya evde bırakılması da bu devamlılığın bir sonucudur [1, ss. 97-98, 119-120]. Anlatıcının Zeynep Kadın'ı tanıtırken gücünü, eli ayağı topraktan yeni sökülmüş köklere; vücudunu ise meşe kütüğüne benzetmesi [1, s. 56] ise romanın kadın tasvirlerinde sıra dışı bir durumu ortaya koymaktadır. Bu hâliyle kadın bedenine cinsel açıdan değil, iş gücü açısından yaklaşılmıştır.

Zeynep Kadın'ın aldığı kararlar, köyün sosyo-ekonomik durumuyla doğrudan ilgilidir. Oğlu Mehmet Ali askere çağrıldığında jandarmaların karşısına dikilir ve anlatıcının vatan savunmasına dair sözlerine tek cümleyle *"Bebelerimizi tam iş zamanında alırlar"* karşılığını verir [1, s. 79]. Askerliği, vatani bir görev veya bir kader olarak değerlendirmek yerine, hasat mevsiminde el konulan bir emek olarak tanımlar. Erkeklerin seferberliğe çağrılması sırasında kadınların topluca itiraz etmesi de bu emeğin kamusallaştırıldığının en büyük delillerindendir [1, ss. 77-81]. Kadınlar, erkeklerin köyden uzaklaşmasını; ayrılık ve aile özlemi olarak algılamak yerine iş zamanında çalışacak kişilerin/iş gücünün

eksilmesi ve çocukların sahipsiz kalması şeklinde okumaktadır. Böylece askerlik, devlet ile erkek köylü arasındaki bir yükümlülük olmaktan çıkarak bütün hanenin üretim düzenini sekteye uğratan karara dönüşmektedir.

Mehmet Ali'nin karısı ve kız kardeşleri bu düzenin adsız çalışanlarıdır. Yeni gelinin eve getirilişi sırasında kullanılan *"Nihayet, gelin bir hamam bohçası gibi cansız ve şahsiyetsiz, evden içeri sokuldu"* ifadesi, evliliğin kadın açısından nasıl bir yer değiştirme ve teslim alınma biçimine dönüştüğünü göstermektedir [1, s. 53]. Benzetme gelinin duygularını açıklamaz; onun canlı bir kişi olarak algılanmadığını kaydeder. Kısa süre sonra ebesiz doğum yapması ve yeniden çalışma düzeni içinde görünmesi, hane için taşıdığı işlevin eşlik, doğurganlık ve emek çevresinde kurulduğunu düşündürmektedir [1, ss. 91-92, 119-120].

Doğum sahnesindeki sağlık hizmetleri ve imkânının yokluğu, köydeki kadın hayatının zorlu şartlarını ve imkânsızlıklarını gözler önüne sermektedir. *"Ebe de ne olacak? İşte, tavana urganı bağladım. Ona asıla asıla kurtulur"* [1, s. 91] ifadesi, doğum için, tavana bağlanan urganın yeterli görüldüğünün en büyük kanıtıdır. Burada doğum, olağan bir kadın işi sayıldığı için tıbbi yardım gerektiren bir durum gibi algılanmamaktadır. Annelik, duygusal yakınlıktan önce bedensel tehlike ve çalışma zorunluluğuyla belirlenmektedir. Bebeğin daha sonra eşikte kendi başına bırakılması ve maruz kaldığı bakım eksikliği ise annenin kayıtsızlığından ziyade yoksulluk içindeki iş bölümünün sonucudur.

Emeti Kadın'da emek, yaşlılık ve aile kayıpları başka bir bileşim oluşturmaktadır. Ahmet Celâl'in ev işlerini görürken torunu Hasan'a da bakan Emeti, yoksulluğunu kendi geçmişiyle açıklar. Hasan'ın babası savaşta ölmüş, kızı doğum sırasında hayatını kaybetmiş, damadı ise onları evden çıkarmıştır [1, ss. 160-161]. Bu durum, çocuğun bakım sorumluluğunun, ölüm, savaş ve terk edilme gibi sebeplerden dolayı yaşlı kadının üzerine kaldığını göstermektedir. Anlatıcının bu kadını tanıttığı satırlar, romandaki en acımasız tasvirdir. Anlatıcı, Emeti Kadın'ın yüzünü çamura bulanmış bir karnabahara benzetir ve ardından yüzünü niçin yıkamadığını sorar. Aldığı cevap, romanda bir kadının kendi emeğini kendi ağzından anlattığı tek yerdir. Emeti Kadın, bir gününün nasıl geçtiğini sırasıyla "sabah karanlığında torununu yollamak, dağarcığına yiyeceğini koymak, ineğin altını temizlemek, sütünü sağmak, ocağa vurmak, sonra koşarak buraya gelmek, su kaynatmak, akşamdan kalan kabı kacağı yıkamak, her gün türlü aş pişirmek, akşam eve varınca aç ve yorgun çocuğa süt ısıtıp ekmek doğramak [1, ss. 151-152] şeklinde sıralamaktadır. Anlatıcının sorusu bir temizlik sorusu, cevabı ise bir iş günü dökümüdür. İkisi arasındaki mesafe, romanın aydın ile köylü arasına koyduğu mesafenin somut ölçüsüdür.

Romanın bir diğer kadın karakteri Cennet ise romana bir dedikoduyla girmektedir. Çocukken cinsel saldırıya uğramış ve köylüler tarafından dışlanmıştır; Süleyman adında bir gençle evlenmiştir [1, ss. 73-76]. Geçmişte maruz kaldığı bu saldırı, yetişkinliğinde onu küçük düşürmek için kullanılan bir anlatıya dönüşmüştür. Cennet'in *"Ben, yalnız kocama teslim olurum!"* [1, s. 74] sözü, sessiz kalmayıp kalabalık önünde kendi bedeni hakkında konuşma hakkını üstlendiği andır. Konuşmalarıyla evlilik sadakatinin dilini kullanırken kocasının ve köyün, onu geçmişteki saldırıyla tanımlamasını reddeder. Dinî hukukun usulünü kocasından ve köyün imamından da iyi bilen kadın, bu bilgiyi kendi lehine çevirmekte ve evlilik hukukunu kullanarak kocasını saf dışı bırakmaktadır [1, ss. 83-84, 123-125]. Söz konusu sahne, kadınların mülk, nikâh ve alacak üzerindeki haklarının kâğıda değil; bir güç dengesine bağlı olduğunu kanıtlamaktadır.

Emine ise romanda en çok sözü geçen ancak en az konuşan kadındır. İlk göründüğü sahnede bir gölgeden ibarettir. Kavaklar arasında geyik hafifliğiyle kaçar, ağaç arkasına saklanır ve sorulara ağaç arkasından cevap verir [1, s. 65]. Anlatıcı, adını, Emine'nin kendisinden değil; İsmail'den öğrenir [1, s. 110]. Babasının, Balkan Harbi'nde şehit düştüğü kabul edilmiş, annesi başkasıyla evlenip onu ortada bırakmış, bu nedenle Emine, dul bir halanın yanında büyümüştür [1, s. 114]. Kendi evliliği hakkındaki tek beyanı iki kısa cümledir; İsmail'e varmak isteyip istemediği sorulduğunda ise ne bileceğini söylemekte ve *"Halam bilir"* diye eklemektedir [1, s. 117]. Kararın, kendisine değil; halasına ait olduğu bu iki kelimeyle bildirilmiştir. Nitekim evlilik de sahne dışında gerçekleşmiş; anlatıcı, evliliğin olup bittiğini tesadüfen öğrenmiştir [1, s. 153].

Romanda kadınlar arasında bir dereceleme bulunur. Bu derecelemenin en aşağı ucunda ise muhtarın kötürüm karısının olduğu görülür. Sekiz yıldan beri yatağa bağlı olan bu kadını anlatıcı hiç görmez, varlığını Mehmet Ali'den öğrenir. Kocasının her gün "gözleri de kapansın, oldu olacak" [1, s. 34]

şeklinde bağırdığı söylenir. Bu kadın, romanın sonunda köy yakılırken evde kalır [1, s. 237]. Adı, sesi ve sahnesi olmayan bir kişinin ölümü, olay örgüsünde hiçbir sonuç doğurmaz.

B. İşgalde Söz, Beden ve Korunma

Savaş köye doğrudan ulaştığında kadınların önceki konumları ortadan kalkmaz, taşıdıkları yükler daha da ağırlaşır. Ürünlere ve paraya el konulması, evlerin yakılması ve erkeklerin dövülmesi bütün haneyi etkilemektedir. Kadınlar için bunlara, kaçırılma ve cinsel saldırı korkusu da eklenmektedir. Mehmet Ali'nin karısı ile kız kardeşlerinin gündelik hayattaki adsızlığı işgal sırasında da sürmekte; kız kardeşlerden birinin işgalci askerler tarafından götürülmesi tek bir cümleyle haber verilmektedir [1, s. 256].

Zeynep Kadın bu silinmeye konuşarak karşı koymaktadır. İşgal subayı, köylüleri sorguya çektiği sırada kadınlar genellikle ağlarken kadınlardan yalnızca Zeynep Kadın ağlamayıp *"Evimi yaktınız. Harman yerindeki buğdayımı yaktınız. Bütün paramı, altınlarımı aldınız. Gelinlik kızlarımın boyunlarındaki Mahmudiyelere kadar neyimiz varsa çaldınız"* [1, s. 246] diyerek kayıplarını tek tek saymaktadır. Sözü ev, ürün, para ve kızların ziynetleri arasında ayrım yapmadan hanenin bütün varlığını kapsamaktadır. Bu, soyut bir vatan konuşması değildir; savaşın köylü kadın açısından mülkün, emeğin ve aile geleceğinin kaybı anlamına geldiğini bildirmektedir. Tercüman, sözlerini yumuşatarak aktarınca susmamakta, göğsünü açıp canını istemekte; ardından kızlarına ve gelinlerine dönerek onlara susmamalarını söylemektedir. Emine'yi göstererek kaç defa gözlerinin önünde ona dokunulmak istendiğini bildirmektedir [1, s. 247]. Zeynep Kadın, burada yalnız kendisi için konuşmamakta, konuşamayanlar adına da konuşmayı denemektedir. Zeynep Kadın, itirazı nedeniyle dövülür fakat aldığı darbe sözünü hükümsüz bırakmaz. Romanın başında sertliğiyle tanıtılan kadın, aynı sertliği işgal kuvvetleri karşısında kamusal bir dirence dönüştürür.

Emine'nin işgal günündeki konumu ise başka bir korunmasızlık biçimini gösterir. Kalabalığın içinde büzülerek şekilsiz bir bohça hâline gelir [1, s. 241], bir asker yüzünü açmasını isteyince kocası İsmail birkaç adım ötede durup olan biteni yan gözle seyreder. Romanın sonunda ise kaçarken vurulur ve yürüyemez hâle gelir. Ardından anlatıcı, defteri ona teslim edip yola çıkar [1, s. 262]. Böylece roman boyunca söz hakkı verilmeyen kadın, romanın metnini elinde tutan kişi olarak konumlandırılır. Bu tercih, anlatıcının kadınlar hakkındaki hükümlerini kurgunun kendi içinde düzelten en güçlü işaretlerden biridir.

C. Çocukluğa Ayrılmayan Zaman

Yaban'da oyun ve eğitim için kendilerine özgü düzenlenmiş bir alana sahip olmayan çocuklar, iş gücü bakımından yetişkinlerle eş düzeyde bir hayat yaşamaktadır. Çeşme başında, çöplükte, tarlada, hayvanların yanında ve işgal sırasında toplanan kalabalığın içinde görülmektedirler. Romanda çalışma/emek, hastalık ve şiddet; çocukluğun gündelik şartlarını belirleyen temel unsurlardır. Kadın karakterlerde olduğu gibi kimi çocuklar, adlarıyla anlatıda yer edinirken kimileri adsızdır. Bu çocuklardan İsmail ve Hasan; adları, konuşmaları ve işlevleriyle bireyselleşirken yeni doğmuş bebek, göç yolundaki çocuklar ve işgal sırasında ağlayanlar ise toplu görüntüler olarak yer edinmektedir.

İsmail'in ilk tasviri, çocukluk ile iş gücü arasındaki ilişkinin açık biçimde kurulduğu sahnedir. Ahmet Celâl, okul meselesini sorduğunda İsmail'in bu soruya verdiği cevap, *"Ne okulu be. Ben okula gideyim de burada işe kim baksın? Hem bu köyde okul yok"* [1, s. 40] şeklinde kesin ve kendinden emindir. Okula gitmeyişi kişisel isteksizlikten kaynaklanmamaktadır. Köyde düzenli bir okul bulunmadığı gibi ailenin onun emeğine ihtiyacı vardır. Başka bir ifadeyle çocuk, eğitim ile çalışma arasında kendisine sunulmuş bir seçim ve bir imkân bulunmadığını bildirmektedir. İsmail yük taşımakta, tarla ve hayvan işlerinde çalışmakta; sıtma nöbetleri geçirmesine rağmen sorumlulukları hafiflememektedir. Anlatıcı onun hayatını *"Bu yaratık, çocukluk nedir bilmedi"* cümlesiyle değerlendirmektedir [1, s. 58]. Söz konusu bu karşılaştırma, Ahmet Celâl'in çalışma yükü, hastalık ve anneden görülen şiddet bakımından İstanbuldaki ve köydeki çocukluğa dair maddi ölçüye dayanmaktadır.

İsmail'in uğradığı şiddet onu sürekli edilgen bırakmamaktadır. Cennet'in sakladığı erkeği haber vermesi, Emine'yle ilişki kurması, Mehmet Ali'nin dayağından sonra hane içindeki tavrının sertleşmesi ve evlenmesi, olayların akışında etkin hâle geldiği aşamalardır [1, ss. 83-84, 108-117, 127, 141-155]. Roman, İsmail'i önce on dört yaşında tanıtmakta; ilerleyen kısımda on yedi yaşına girmeden evlilik ve

yetişkinlik sorumluluklarına yaklaştığını belirtmektedir [1, s. 40, s. 112]. Dolayısıyla çocukluk ile yetişkinlik takvim yaşından çok çalışma, cinsellik ve evlilik üzerinden belirlenmektedir.

Romanın bir diğer çocuk karakteri Hasan da emeğiyle tanıtılmakta fakat İsmail'den farklı bir kişilik çizgisi kazanmaktadır. On bir veya on iki yaşında olduğu belirtilen çocuk için anlatıcı *"ne kadar ağırbaşlı, ne kadar vakarlı, ne kadar görev ve sorumluluğunu anlamış bir insandır"* demektedir [1, s. 152]. Yaş ile sorumluluk arasındaki uyumsuzluk burada övgü biçiminde verilmektedir. Hasan'ın ağırbaşlılığı doğuştan gelen bir erdem gibi görünse de sığırtmaçlık yapması, annesiz ve babasız büyümesi, Emeti'yle birlikte yoksulluk içinde yaşaması bu tavrın şartlarını açıklamaktadır. İsmail'in erken büyümesi daha çok çirkinleşme ve sertleşme üzerinden anlatılırken Hasan'ınki vakar ve sadakat üzerinden değerlendirilmektedir. İki çocuk arasındaki fark, maddi şartlarından çok, anlatıcının onlara duyduğu yakınlığın derecesiyle ilgilidir. Savaş ilk aşamada Hasan için uzaktan izlenen bir hareketlilik taşımakta [1, ss. 193-195]; işgal kuvvetlerinin köye girmesiyle çocuk gözlemci konumunu kaybederek doğrudan şiddetin hedefi olmaktadır.

İsmail ile Hasan'a verilen bireysellik, köydeki çocukların çoğuna tanınmamaktadır. Ahmet Celâl köyün çöplüğünü anlatırken *"Köyün mezbelesinde, köpek enikleriyle insan yavruları birbirine karışmış, oynaşıyorlar"* demektedir [1, s. 48]. Çocuklar yiyecek artıkları için hayvanlarla yarışmakta; oyun ile açlık aynı görüntüde birleşmektedir. "İnsan yavruları" ifadesi, onların yaşını ve kimliğini belirsizleştirirken insanla hayvan arasındaki sınırı da zayıflatmaktadır. Sahne, köy yoksulluğuna ilişkin güçlü bir kanıt taşımakla birlikte kullanılan dil aynı zamanda anlatıcının köylülere yönelen insanlıktan uzaklaştırıcı bakışını açığa çıkarmaktadır. Küçük çocukların kırk yaşındaki erkekleri andıran yüzlerle tasvir edilmesi [1, s. 58] de çocukluğu, beden bakımından erken yaşlanma olarak tanımlamaktadır. Yaşlı görünmek, bilgelik belirtisi değil, hastalık ve yetersiz beslenmenin yüzdeki izidir.

Bir diğer çocuk Zehra ise çocuklar arasında farklı bir korunmasızlık taşımaktadır. On iki on üç yaşlarında, görme engelli ve cılız bir kız çocuğu olan Zehra, tarlada yalnız bırakıldıktan sonra Salih Ağa'nın oğlu tarafından cinsel saldırıya uğrar [1, ss. 98-100]. Olayın toplumsal karşılığı çok sonra Bekir Çavuş'un ağzından duyulur. Adam, anlatıcıya kızının berbat edildiğini, alması gerektiğini kambur oğlana söylediğini, onun almam dediğini ve hastalık bahanesi ileri sürdüğünü anlatır. Anlatıcının gebelik ihtimalini sorması üzerine verdiği cevap, *"Yok olamaz. Daha on iki yaşında"* [1, s. 204] şeklindedir. Yaşın burada bir korunma işlevi değil; bir gebelik hesabı olarak algılanması, olayın köy içinde nasıl adlandırıldığını gösterir. Aylarca maruz kalınan bir istismar, bir evlilik pazarlığına indirgenmiştir. Böyle bir şiddete uğramasına rağmen romanda Zehra'ya hiçbir söz hakkı verilmemiştir. Failin yaşı ise kesin biçimde verilmemekte; anlatıcı ondan kimi yerde çocuk, kimi yerde evlenmesi beklenen bir genç erkek gibi söz etmektedir [1, ss. 98-100, 190-191, 204]. Söz konusu bu belirsizlik, çocukluk veya gençliğin kendiliğinden masumiyet üretmediğini hatırlatmaktadır.

Bir başka savunmasız çocuk ise Mehmet Ali'nin yeni doğan oğludur. Çocuk ebesiz doğar, aylarca adı konulmaz ve kadınlar çalışırken sokak kapısının eşiğine bırakılır. Anlatıcı onu sürüngen bir bitkiye benzeterek ana karnından çıkmış olmaktan çok, topraktan bitme hissi verdiğini ifade eder [1, ss. 119-120]. Verdiği hüküm ve bulduğu çözüm, *"O kadar kişiliği yok ki daha adını bile koymadılar"* [1, s. 120] şeklindedir. Bunun üzerine çocuğu "adsız" diye çağırmaya başlar. Başka bir ifadeyle ada muhtaç bir çocuğa verilen ad, adsızlıktır. Göç kafilelerindeki çocuklar da aileleriyle birlikte hareket etmelerine rağmen güvenli bir yere yönelmemektedir. Kadınların sırtında taşınan bebekler ve kalabalıktan ayrılıp kaybolan çocuk, savaşın çocukluk mekânını ortadan kaldırdığını göstermektedir [1, ss. 171-173]. Kaybolan çocuk, Ahmet Celâl'in yardım girişiminden korkarak kaçmaktadır; yetişkinin yaklaşması onun için korunma vaadi değil, yeni bir tehlike ihtimalidir. İşgal sırasında yarı çıplak ve ağlayan çocukların süngü tehdidi altındaki kalabalıkta bulunması da aynı güvensizliği sürdürmektedir [1, ss. 220, 224, 240, 249].

IV. TARTIŞMA

Bulgular, üretimin sürdürülmesi, aile üyelerinin bakımı ve gündelik tehlikelerden korunmanın çoğu zaman kadınlarla çocukların omuzlarında bulunduğunu göstermektedir. Devletin, düzenli sağlık hizmetinin, okulun ve güvenliği sağlayacak bir gücün köy hayatındaki sınırlı varlığı; bu sorumluluğu aileye bırakmaktadır. Romanın çizdiği köyde bakım ve eğitim için hane emeğinin dışında bir kaynak

bulunmadığından yetişkin kadınların ve çocukların ihtiyaçları birbirine rakip hâle gelmektedir. Söz konusu tablo, yoksulluk içindeki bakım ilişkilerinin bireysel ahlak üzerinden açıklanamayacağını ortaya koymaktadır.

Savaşın aile üzerindeki etkisi, işgal sahnelerinden önce başlamaktadır. Erkeklerin askere alınması; üretimdeki iş bölümünü değiştirmekte, iş yükünün kadın ve çocuklar üzerindeki etkisini ağırlaştırmaktadır. Bunun yanında kadınlar, iş zamanında götürülen erkeklerin ardından tarlanın ve çocukların ne olacağını sorgulamaktadır [1, ss. 77-81]. Savaşın, kadın emeğini seferber ederken kadınları karar süreçlerinin dışında tuttuğuna [25] ve ulusal anlatıların, kadınları sembolik düzeyde yüceltirken siyasal özne saymadığına [26] ilişkin çözümlemeler, romanın bu sahneleriyle örtüşmektedir. Nitekim *Yaban*'daki kadınlar, vatan söyleminin muhatabı sayılmamakta; buna karşılık savaşın maliyeti doğrudan onların emeği ve bedeni üzerinden ödenmektedir.

İşgal başladığında kadın ve çocuklar aynı meydanda toplanarak aynı yangın ve açlık korkusunu yaşamakta ve aynı kaderi paylaşmaktadır. Evlerin aranması, yiyeceklere el konulması ve insanların zorla dışarı çıkarılması yaş ayrımı gözetmemektedir [1, ss. 214-220, 235-249]. Erkeklere yönelen şiddet, ağırlıklı olarak mala; kadınlar ile çocuklara yönelen şiddet ise bedene gelmektedir. Kadın ve kızlar kaçırılma ve cinsel saldırı korkusu taşırken çocuklar dayak, süngü tehdidi, açlık ve gözetimsiz kalma tehlikesiyle karşılaşmaktadır. Romanda öldürülen tek kişinin on bir yaşında bir çoban olması, korunma imkânının en az bulunduğu yerin çocukluk olduğunu gözler önüne sermektedir.

Baş, yazarın kadın karakterleri çoğunlukla erkek kahramanın bakışı ve arzusu üzerinden kurduğunu belirtmektedir [11]. *Yaban* bu eğilimi, Emine örneğinde doğrulamakla birlikte Zeynep Kadın, Cennet ve Emeti örneklerinde aşmaktadır. Alver'in kadın-doğa ilişkisi üzerinden yürüttüğü çözümleme [13] de kadınların doğayla özdeşleştirilerek nesneleştirildiğine işaret etmektedir ancak bu araştırmanın bulguları, aynı özdeşleştirmenin kimi kişilerde iş gücü ve dayanıklılık vurgusuna dönüştüğünü ortaya koymaktadır. Başka bir ifadeyle romandaki kadın tasviri tek yönlü bir nesneleştirme değildir.

Söz hakkının dağılımı bakımından elde edilen bulgular, tabii konumdaki öznenin kendi adına konuşma imkânını yapının kendisinin sınırladığı yönündeki tartışmayla ilişkilendirilebilir [27]. Romanda söz sahibi olan kadınların dul kadınlar olması tesadüf değildir; koca otoritesinin ortadan kalkması, kadını hane reisi konumuna taşıyarak ona kamusal bir konuşma zemini açmaktadır. Buna karşılık Zehra, hiçbir yerde konuşturulmamakta; uğradığı şiddet yetişkin erkeklerin sözleri içinden aktarılmaktadır. Adlandırma ile konuşma imkânı arasındaki koşutluk, romanın toplumsal eleştirisinin taşıyıcı ögelerinden biridir.

Anlatıcının konumu da bulguların yorumlanmasında belirleyicidir. Ahmet Celâl, kadın ve çocuk bedenlerini kimi zaman hayvanlara veya nesnelere benzetmekte; genç kadınlara yönelik tasvirlerinde arzu ile kültürel üstünlük duygusunu bir araya getirmektedir. İsmail'i kıskançlıkla, Hasan'ı ise yakınlık ve idealleştirmeyle değerlendirmesi çocuklar hakkındaki hükümlerini de değiştirmektedir. Bununla beraber romanın olayları, anlatıcının ilk yargılarını her zaman doğrulamamaktadır. Zeynep'in ekonomik gücü ve direnci, Cennet'in meydan okuması, Hasan'ın sorumluluk bilinci ve Emine'nin son bölümdeki konumu, Ahmet Celâl'in köylüleri edilgen ve biçimsiz bir topluluk olarak gören yaklaşımını aşmaktadır. Bu sebeple anlatıcının sözlerini doğrudan yazarın görüşü olarak aktarmak [5], [14], romanın kendi içindeki düzeltme imkânlarını ortadan kaldırmaktadır.

V. SONUÇLAR

Yaban'daki kadın ve çocuklar üzerine yapılan bu inceleme, bu kişilerin köy hayatını tamamlayan yardımcı figürler olarak görülmesinin romanın sunduğu malzemeyi açıklamakta yetersiz kaldığını göstermiştir. Köyün üretim düzeni, aile ilişkileri, yoksulluk, hastalık, şiddet ve savaş büyük ölçüde onların gündelik işleri ve bedenleri üzerindeki sonuçlarıyla anlatılmaktadır. Erkeklerin askere alınmasıyla artan iş yükü, doğum ve bakımın güç şartları, okulun yokluğu ve işgal sırasında korumanın çökmesi; romanın toplumsal yapısını oluşturan temel olaylardır. Dolayısıyla kadınlarla çocukları anlatıdan çıkarmak, köy hayatının nasıl sürdüğünü ve savaşla nasıl dağıldığını da büyük ölçüde belirsiz bırakacaktır. Bu bulgu, bütün kadınların aynı konumda bulunduğu anlamına gelmemektedir. Zeynep Kadın dul bir aile reisi olarak üretim, mülkiyet ve evlilik kararlarına müdahale etmekte; işgal sırasında kayıplarını kamusal biçimde dile getirmektedir. Cennet, kendisi hakkında kurulan ahlaki anlatıya söz ve eylemle karşılık vermektedir. Emeti, aile kayıplarını kendi tanıklığıyla aktarırken torununun bakımını

üstlenmektedir. Emine'nin hareket alanı Ahmet Celâl'in arzusu ve halasının denetimiyle çevrilidir; yine de evlilik teklifini reddetmesi ve kaçış sırasındaki konumu, onu anlatıcının seyir nesnesi olmaktan çıkarmaktadır. Buna karşılık Mehmet Ali'nin karısı ve kız kardeşleri ile işgal sahnelerindeki birçok kadın adsız veya sessiz kalmaktadır.

Çocuklarda da benzer bir eşitsizlik bulunmaktadır. İsmail çalışma, hastalık, aile içi şiddet ve eğitimden yoksunluk içinde erken yetişkinleşmekte; evlendikten sonra genç erkeklere ayrılan otorite alanına yaklaşmaktadır. Hasan sığırtmaçlığı, konuşmaları ve Ahmet Celâl'le kurduğu ilişki sayesinde belirgin bir kişilik kazanmakta; ölümünün etkisi önceden kurulmuş bu kişiliğe dayanmaktadır. Zehra ise yaşı, cinsiyeti ve engelliliği nedeniyle daha sınırlı bir hareket alanına sahiptir. Adsız bebek, çöplükteki ve göç kafilelerindeki çocuklar ise bireysel bilinçleriyle değil; bedenleri ve içinde bulundukları şartlarla tanınmaktadır. Dolayısıyla çocukluk; eğitim, oyun, bakım ve güvenliğin farklı ölçülerde eksildiği bir hayat dönemidir.

Kadınlarla çocukları birleştiren başlıca şart korunma araçlarının azlığıdır. Savaş bu eksikliği yaratmaktan çok ağırlaştırmakta; önceden var olan yoksulluk, emek yükü ve aile kayıpları işgalle birlikte göç, talan, yangın, dayak, cinsel tehdit ve ölüme dönüşmektedir. Bazı sahnelerde kadın ve çocuklar, köy kalabalığının adsız üyeleri olarak kullanılırken aslında onlar aracılığıyla köyde üretimin kimler tarafından sürdürüldüğü, aile otoritesinin nasıl dağıldığı, evlilik ve bakım ilişkilerinin hangi şartlarda kurulduğu, yoksulluğun çocukluğu nasıl daralttığı ve savaşın korumasız bedenlere nasıl ulaştığı anlatılmaktadır. Kişilerin bir kısmının sessiz kalması da anlatısal bir boşluk olmaktan çok, söz ve yetkinin köy içinde eşit dağılmadığını gösteren bir özelliktir.

KAYNAKLAR

[1] Y. K. Karaosmanoğlu, *Yaban* (A. Özkırımlı, Haz., 15. bs.). İstanbul: Birikim Yayınları, 1981.
[2] A. Özkırımlı, "Sunuş," Y. K. Karaosmanoğlu, *Yaban* içinde, İstanbul: Birikim Yayınları, 1981, ss. 7-20.
[3] R. Kaplan, *Cumhuriyet Dönemi Türk Romanında Köy*. Ankara: Akçağ Yayınları, 1997.
[4] K. Akyüz, *Modern Türk Edebiyatının Ana Çizgileri 1*. Ankara: Ankara Üniversitesi Basımevi, 1979.
[5] B. Moran, *Türk Romanına Eleştirel Bir Bakış 1*. İstanbul: İletişim Yayınları, 2001.
[6] N. Akı, *Yakup Kadri Karaosmanoğlu: İnsan-Eser-Üslup*. İstanbul: İstanbul Matbaası, 1960.
[7] A. Hayber, *Halide Edip, Yakup Kadri ve Reşat Nuri'nin Romanlarında Nesil Çatışması*. Ankara: Millî Eğitim Bakanlığı Yayınları, 1993.
[8] F. Dündar, "Yakup Kadri Karaosmanoğlu'nun Yaban adlı romanındaki ulus inşa etme amacı," *Akademik Bakış Dergisi*, sy. 43, ss. 1-14, 2014.
[9] E. Biçer, "Yakup Kadri Karaosmanoğlu'nun Yaban adlı romanında aydın yozlaşması," Yüksek lisans tezi, Ardahan Üniversitesi, Ardahan, 2018.
[10] T. Timur, *Osmanlı-Türk Romanında Tarih, Toplum ve Kimlik*. Ankara: İmge Kitabevi, 2002.
[11] Ş. Baş, "Yakup Kadri Karaosmanoğlu'nun romanlarında kadın karakterlerin anlatımı," *Türkbilig*, sy. 6, ss. 13-36, 2003.
[12] A. Serdar, "Yakup Kadri Karaosmanoğlu'nun romanlarında cinsellik," Yüksek lisans tezi, Bilkent Üniversitesi, Ankara, 2002.
[13] T. Alver, "Yakup Kadri Karaosmanoğlu'nun Yaban adlı romanında ekofeminizmin izleri," *Karamanoğlu Mehmetbey Üniversitesi Edebiyat Fakültesi Dergisi*, c. 4, sy. 1, ss. 35-47, 2021.
[14] W. C. Booth, *The Rhetoric of Fiction*, 2. bs. Chicago: University of Chicago Press, 1983.
[15] G. Genette, *Narrative Discourse: An Essay in Method*. Ithaca: Cornell University Press, 1980.
[16] J. Parla, *Don Kişot'tan Bugüne Roman*. İstanbul: İletişim Yayınları, 2011.
[17] F. Naci, *Yüz Yılın 100 Türk Romanı*. İstanbul: Türkiye İş Bankası Kültür Yayınları, 2007.
[18] A. Yıldırım ve H. Şimşek, *Sosyal Bilimlerde Nitel Araştırma Yöntemleri*. Ankara: Seçkin Yayıncılık, 2021.
[19] G. A. Bowen, "Document analysis as a qualitative research method," *Qualitative Research Journal*, c. 9, sy. 2, ss. 27-40, 2009.
[20] J. W. Scott, "Gender: A useful category of historical analysis," *The American Historical Review*, c. 91, sy. 5, ss. 1053-1075, 1986.
[21] D. Kandiyoti, *Cariyeler, Bacılar, Yurttaşlar: Kimlikler ve Toplumsal Dönüşümler*. İstanbul: Metis Yayınları, 1997.
[22] F. Berktay, *Tarihin Cinsiyeti*. İstanbul: Metis Yayınları, 2010.
[23] P. Ariès, *Centuries of Childhood: A Social History of Family Life*. New York: Vintage Books, 1962.
[24] B. Onur, *Türkiye'de Çocukluğun Tarihi*. Ankara: İmge Kitabevi, 2005.
[25] C. Enloe, *Maneuvers: The International Politics of Militarizing Women's Lives*. Berkeley: University of California Press, 2000.
[26] N. Yuval-Davis, *Gender and Nation*. London: Sage Publications, 1997.
[27] G. C. Spivak, "Can the subaltern speak?," C. Nelson ve L. Grossberg (Ed.), *Marxism and the Interpretation of Culture* içinde, Urbana: University of Illinois Press, 1988, ss. 271-313.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Ortaokul Türkçe Ders Kitaplarındaki Şiirlerin Çocuk Edebiyatı ve Türkçe Öğretimi Açısından İncelenmesi

Merve MORKOYUNLU[1*] ve Havvaana KARADENİZ[2]

[1]*Türkçe ve Sosyal Bilimler Eğitimi Ana Bilim Dalı, Sosyal Bilimler Enstitüsü, Kırşehir Ahi Evran Üniversitesi, Türkiye*
[2]*Türkçe ve Sosyal Bilimler Eğitimi Bölümü, Eğitim Fakültesi, Kırşehir Ahi Evran Üniversitesi, Türkiye*

[*]*(morkoyunlu3857@gmail.com)*

Özet – Bu araştırmanın amacı, Türkiye Yüzyılı Maarif Modeli doğrultusunda hazırlanan ve Millî Eğitim Bakanlığı tarafından yayımlanan 5. sınıf Türkçe ders kitabında yer alan şiirleri çocuk edebiyatı ve Türkçe öğretimi bakımından incelemektir. Nitel araştırma yaklaşımının benimsendiği çalışmada doküman analizi yöntemi kullanılmış, veriler alan yazınına dayalı olarak geliştirilen "Şiir İnceleme Ölçütleri Formu" aracılığıyla toplanmış ve betimsel analiz ile içerik analizi tekniklerinden yararlanılarak çözümlenmiştir. İnceleme sonucunda ders kitabının iki cildinde altı tema kapsamında sunulan otuz metnin beşinin şiir türünde olduğu, buna karşılık öğretim programının bilgilendirici, hikâye edici ve şiir biçimleri arasında öngördüğü dengeli dağılım ilkesinin sağlanamadığı belirlenmiştir. Şiirlerin temalara eşit biçimde dağılmadığı, "Sağlıklı Yaşıyorum" temasında hiçbir şiire yer verilmediği; basılı olarak sunulan altı şiirden beşinin kısaltılarak aktarıldığı ve bu durumun metinlerin estetik bütünlüğünü zedelediği görülmüştür. Biçim bakımından şiirlerin üçte ikisinin serbest ölçüyle yazıldığı, hece ölçüsüne ve düzenli uyak dizgesine dayalı örneklerin sınırlı kaldığı saptanmıştır. Değer aktarımı boyutunda sevgi, saygı, çalışkanlık, aile bütünlüğü ve vatanseverlik değerlerinin ön plana çıktığı; söz varlığı çalışmalarının ise yalnızca ana okuma metni işlevindeki şiirlerle sınırlandırıldığı tespit edilmiştir. Çocuğa görelik ölçütü bakımından şiirlerin büyük bölümünün kısmen uygun olduğu, arkaik söz varlığı ile yetişkin bakış açısının yer yer çocuk gerçekliğinin önüne geçtiği sonucuna ulaşılmıştır. Araştırma, metin seçiminde estetik bütünlüğün korunması ve tür dengesinin gözetilmesi gerektiğini ortaya koymaktadır.

Anahtar Kelimeler – çocuğa görelik, çocuk edebiyatı, şiir, Türkçe ders kitabı, Türkiye Yüzyılı Maarif Modeli

I. GİRİŞ

Ana dili eğitimi, dil bilgisi kurallarının öğretildiği bir alan olmasının yanında, bireyin duygu ve düşünce dünyasının biçimlendiği bir duyarlık alanıdır. Bu bakımdan Türkçe derslerinde kullanılan metinlerin niteliği, öğretim sürecinin başarısını doğrudan etkilemektedir. Çocuğun estetik duyarlığını geliştiren, ona dilin sınırlarını ve olanaklarını sezdiren metinler arasında şiirin ayrı bir yeri bulunmaktadır [1]. Şiir; gerek içerik ve öz gerekse söze dönüştürme ve sunuluş bakımından özgün, etkilemeye ve duygulandırmaya yönelik bir söz sanatı ürünü olarak tanımlanmakta [2] ve bu yönüyle diğer metin türlerinden ayrılmaktadır.

Şiirin çocuk edebiyatı içindeki konumu, çocuğun gelişim özellikleriyle doğrudan ilişkilidir. Çocuğun şiirle kurduğu ilişki önce dinleme becerisiyle başlamakta, ilerleyen süreçte okuma, ezberleme ve tekrarlama biçiminde devam etmektedir [3]. Kısa dizelerden oluşması ve ahenk yönü sayesinde şiir, çocuğa nesirden farklı bir şekilde hitap etmekte; onun algı dünyasını renklendirmekte, hayal gücünü geliştirmekte ve sevinç, hüzün, coşku, cesaret gibi duyguları deneyimlemesine imkân tanımaktadır [4]. Bunun yanı sıra şiir; öğrencilerin telaffuz, ezber ve konuşma becerilerini besleyerek kendini ifade etme becerisini geliştirmektedir. Dolayısıyla şiir, dört temel dil becerisinin tamamıyla ilişkilendirilebilecek önemli bir tür olarak karşımıza çıkmaktadır [5].

Ancak şiirin bu işlevi kendiliğinden gerçekleşmemekte, seçilen metnin niteliğine bağlı olarak ortaya çıkmaktadır. Çocuk edebiyatı alan yazınında bu niteliği belirleyen temel ölçüt "çocuğa görelik" kavramıyla açıklanmaktadır. Çocuğa görelik, çocuk için yazılan edebî ürünlerin çocuğun fiziksel, bilişsel, ruhsal, toplumsal ve dilsel gelişim evrelerine uygun olmasını ifade etmekte; bu ilkenin en aktif boyutunu ise çocuğun zihinsel şemalarında karşılığı olan sözcük ve kavramların seçilmesi oluşturmaktadır [6]. Söz konusu ilke gereğince soyut ve didaktik kavramların doğrudan aktarılması yerine bu kavramların bir olay, kurgu veya serüven aracılığıyla somutlaştırılması gerekmektedir. Başka bir ifadeyle çocuk edebiyatı, kavramları kurgusal yaşantılar üzerinden somutlaştırarak dille buluşturan özgün bir anlatı diline sahip olmalıdır [7]. Bu bakımdan nitelikli edebiyat yapıtları çocuğu pasif bir okur olmaktan çıkarmakta ve öğrenme sürecine aktif bir şekilde katılmasını sağlamaktadır [8], [9].

Çocuk şiirinde biçim ögeleri de içerik kadar önem arz etmektedir. Kısa dize yapısı, kurallı ve dengeli cümle düzeni, dize tekrarları ve ritim unsurları metnin algılanmasını kolaylaştırmakta; içerik boyutunda ise yaşama sevinci, aile ve yurt sevgisi, doğa ve hayvan sevgisi ile estetik ve toplumsal duyarlılık gibi olumlu temalar öne çıkmaktadır [3], [10]. Çocukların bütüncül bir duygu ve düşünce dünyası geliştirebilmeleri için tekdüze anlatımlarla sınırlandırılmayıp epik, lirik, didaktik, pastoral ve dramatik gibi farklı şiir türleriyle buluşturulmaları pedagojik açıdan önemlidir [10].

Türkiye Yüzyılı Maarif Modeli doğrultusunda hazırlanan Türkçe Dersi Öğretim Programı, şiiri bilgilendirici ve hikâye edici metinlerle birlikte üç ana metin biçiminden biri olarak tanımlamaktadır [11] [12]. Söz konusu tanımlama şiire araçsal bir metin türü olmanın ötesinde resmî bir statü kazandırmakta; ayrıca metin biçimlerinin kitap genelinde ve temalar arasında dengeli bir dağılım sergilemesini esas kural olarak belirlemektedir. Türkiye Yüzyılı Maarif Modeli doğrultusunda hazırlanan 5. sınıf Türkçe Dersi Öğretim Programı altı tema üzerine yapılandırılmıştır. Programda metin seçimi ve metin biçimlerinin dağılımı, dil becerilerinin geliştirilmesiyle birlikte ele alınmaktadır. Bu bakımdan ders kitabındaki şiirlerin hem tür dağılımı hem de çocuğa görelik bakımından incelenmesi, programın uygulamaya yansımasını göstermesi açısından önem taşımaktadır.

Alan yazınında Türkçe ders kitaplarındaki şiirleri konu edinen çalışmalar bulunmaktadır. Tezgör, Cumhuriyet döneminden 2005 yılına uzanan süreçte ortaokul Türkçe ders kitaplarındaki şiiri tarihsel bir çerçevede ele almıştır [13]. Kemiksiz, ortaokul 6, 7 ve 8. sınıf ders kitaplarındaki 438 metnin 79'unun şiir türünde olduğunu belirlemiştir [14]. Ercan 8. sınıf ders kitaplarındaki şiirleri şekil, tema ve tür açısından [15]; Bayat ise şiirbilim ölçütleri bakımından değerlendirmiştir [16]. Çapoğlu ve Okur 8. sınıf ders kitaplarındaki şiirlerde yer alan değerleri belirlemiş [17]; Çolak ve Günesu ise şiirleri çocuğa görelik ilkesi doğrultusunda incelemişlerdir [18], [19]. Bununla birlikte Türkiye Yüzyılı Maarif Modeli kapsamında hazırlanan yeni ders kitaplarındaki şiirleri çocuk edebiyatı ve Türkçe öğretimi ölçütlerini birlikte ele alarak değerlendiren bir çalışmaya rastlanmamıştır. Bu araştırma söz konusu boşluğu doldurmayı amaçlamaktadır.

A. Araştırmanın Amacı ve Alt Amaçları

Bu araştırmanın temel amacı, Türkiye Yüzyılı Maarif Modeli doğrultusunda hazırlanan 5. sınıf Türkçe ders kitabındaki şiirleri çocuk edebiyatı ve Türkçe öğretimi bakımından incelemektir. Bu temel amaç doğrultusunda aşağıdaki sorulara cevap aranmıştır:

1. Ders kitabındaki şiirlerin sayısal görünümü ve temalara göre dağılımı nasıldır? 2. Şiirler, biçim özelliklerinin öğretimi bakımından uygun mudur? 3. Şiirler, değer aktarımı bakımından uygun mudur? 4. Şiirler, çocukların söz varlığını geliştirici nitelikte midir? 5. Şiirler, dil ve anlatım yönünden ilgili sınıf düzeyine uygun mudur? 6. Şiirler, çocuğa görelik ilkesine uygun mudur?

II. MATERYAL VE YÖNTEM

A. Araştırmanın Modeli

Araştırmada nitel araştırma yaklaşımı benimsenmiş ve doküman analizi yöntemi kullanılmıştır. Doküman analizi, araştırılması hedeflenen olgu veya olgular hakkında bilgi içeren yazılı materyallerin çözümlenmesini kapsayan sistematik bir yöntem olarak tanımlanmaktadır [20]. Ders kitapları, öğretim programının uygulamaya aktarıldığı temel materyaller olduğundan öğretim süreçlerine ilişkin çıkarımlar yapmaya elverişli birer dokümandır [21].

B. İnceleme Nesnesi

Araştırmanın inceleme nesnesini, 5. sınıf Türkçe ders kitabının birinci ve ikinci kitapları oluşturmaktadır [22], [23]. Söz konusu kitaplar sırasıyla 182 ve 160 sayfadan meydana gelmekte, altı tema ve bu temalar altında toplam otuz metin içermektedir. İnceleme nesnesinin belirlenmesinde amaçlı örnekleme yöntemlerinden ölçüt örnekleme kullanılmış [24]; "Türkiye Yüzyılı Maarif Modeli doğrultusunda hazırlanmış olma", "Millî Eğitim Bakanlığınca yayımlanmış olma" ve "ortaokulun ilk sınıf düzeyine ait olma" ölçütleri esas alınmıştır. Ortaokulun ilk basamağı olması sebebiyle 5. sınıf, ilkokuldan ortaokula geçişte metin türü algısının biçimlendiği bir eşik durumundadır. Bu durum söz konusu sınıf düzeyinin seçilmesini gerekli kılmıştır.

İnceleme kapsamına, ders kitabında ana okuma metni, okuma keyfi metni, dinleme metni ve etkinlik metni işleviyle yer alan bütün şiirler dâhil edilmiştir. Karekod aracılığıyla sunulan ve tam metni kitapta basılı olmayan dinleme şiirleri, yalnızca sayısal dağılım ve işlev bakımından değerlendirmeye alınmış; biçim, söz varlığı ve çocuğa görelik çözümlemelerinin dışında tutulmuştur.

C. Veri Toplama Aracı

Veriler, alan yazını taraması sonucunda araştırmacılar tarafından geliştirilen "Şiir İnceleme Ölçütleri Formu" aracılığıyla toplanmıştır. Form; künye bilgileri, biçim özellikleri, içerik ve değer aktarımı, söz varlığı, dil ve anlatım ile çocuğa görelik olmak üzere altı boyuttan oluşmaktadır. Boyutların oluşturulmasında çocuk edebiyatı ve şiir bilgisi alan yazınından yararlanılmış [3], [6], [16], [27]; değer boyutunda Türkiye Yüzyılı Maarif Modeli'nin erdem-değer-eylem çerçevesi esas alınmıştır [11]. Form, ikisi Türkçe eğitimi ve biri Türk dili ve edebiyatı alanında olmak üzere üç uzmanın görüşüne sunulmuş, alınan dönütler doğrultusunda ölçütlerin ifade biçimi yeniden düzenlenmiş ve forma son hâli verilmiştir.

D. Verilerin Analizi

Elde edilen veriler betimsel analiz ve içerik analizi teknikleriyle çözümlenmiştir. Şiirlerin sayısal görünümü, temalara ve işlevlere göre dağılımı ile biçim özelliklerine ilişkin veriler frekans ve yüzde değerleriyle betimlenmiş; değer aktarımı, söz varlığı ve çocuğa görelik boyutlarına ilişkin veriler ise kodlama yoluyla kategorilere ayrılarak yorumlanmıştır [26]. Çözümleme sürecinde şiirlerden doğrudan alıntılar yapılarak bulguların inandırıcılığı güçlendirilmiştir. Kodlama güvenirliğini sağlamak amacıyla veri setinin tamamı iki araştırmacı tarafından bağımsız biçimde kodlanmış, kodlayıcılar arası uyum Miles ve Huberman'ın [Görüş Birliği / (Görüş Birliği + Görüş Ayrılığı)] × 100 formülü ile hesaplanmış ve %85 düzeyinde uyum elde edilmiştir [25]. Ayrıca inceleme sürecinin ve ölçütlerin ayrıntılı biçimde betimlenmesi yoluyla aktarılabilirlik ve teyit edilebilirlik ölçütleri gözetilmiştir.

III. BULGULAR

A. Şiirlerin Sayısal Görünümüne İlişkin Bulgular

Ders kitabının birinci ve ikinci ciltlerinde altı tema kapsamında toplam otuz metin yer almaktadır. Her tema; üç okuma metni, bir dinleme/izleme metni ve bir okuma keyfi metni olmak üzere beş metinden meydana gelmektedir. Bu metinler içinde şiir türünde olanların künye bilgileri Tablo 1'de sunulmuştur.

Tablo 1. Ders kitabındaki şiirlerin künyesi

Şiir	**Şair**	**Tema**	**İşlev**	**Sunum**
Oyuncakların Oyuncağı Olmak	Mustafa Ruhi Şirin	1. Oyun Dünyası	Okuma keyfi	Basılı / kısaltılmış
Düş Oyunu	Belirtilmemiş	1. Oyun Dünyası	Dinleme (etkinlik)	Karekod
Atatürk Kurtuluş Savaşı'nda	Cahit Külebi	2. Atatürk'ü Tanımak	Dinleme (etkinlik)	Karekod
Türkiye	Adil Turan	3. Duygularımı Tanıyorum	Okuma (ana metin)	Basılı / kısaltılmış
Uçan Balon	Necdet Neydim	3. Duygularımı Tanıyorum	Etkinlik metni	Basılı / kısaltılmış
Misafir Odası	Gökhan Akçiçek	4. Geleneklerimiz	Okuma (ana metin)	Basılı / kısaltılmış
Kırık At	Tevfik Fikret	4. Geleneklerimiz	Okuma keyfi	Basılı / tam
Gülü İncitme Gönül	Bestami Yazgan	5. İletişim ve Sosyal İlişkiler	Okuma (ana metin)	Basılı / kısaltılmış

Tablo 1'e göre ders kitabında sekiz şiir yer almakta, bunlardan altısı basılı olarak sunulmaktadır. Basılı şiirlerden beşi "(Kısaltılmıştır.)" ibaresiyle verilmiş, yalnızca "Kırık At" bütünüyle aktarılmıştır. Başka

bir ifadeyle basılı şiirlerin %83,33'ü eksiltilerek kitaba alınmıştır. Bu durum, ders kitabına seçilecek metinlerin estetik bütünlüğünün korunması gerektiği yönündeki ilkeyle çelişmektedir [33].

Şiirlerin temalar ve metin işlevleri bakımından dağılımı Tablo 2'de gösterilmiştir.

Tablo 2. Şiirlerin temalara ve metin işlevlerine göre dağılımı

Tema	Okuma (3)	Dinleme (1)	Okuma keyfi (1)	Etkinlik	Toplam şiir
1. Oyun Dünyası	-	-	1	1 (karekod)	2
2. Atatürk'ü Tanımak	-	-	-	1 (karekod)	1
3. Duygularımı Tanıyorum	1	-	-	1	2
4. Geleneklerimiz	1	-	1	-	2
5. İletişim ve Sosyal İlişkiler	1	-	-	-	1
6. Sağlıklı Yaşıyorum	-	-	-	-	0
Toplam	**3 / 18**	**0 / 6**	**2 / 6**	**3**	**8**

Tablo 2, şiirin ders kitabındaki konumu bakımından önemli sonuçlar ortaya koymaktadır. Temalarda yer alan otuz metnin yalnızca beşi şiir türündedir; bu oran %16,67'ye karşılık gelmektedir. Öğretim programının üç ana metin biçimi arasında dengeli dağılım öngördüğü dikkate alındığında beklenen oranın yaklaşık %33 olması gerekmektedir. Dolayısıyla şiir, kitap genelinde beklenenin altında temsil edilmektedir. On sekiz ana okuma metninden yalnızca üçü şiirdir ve altı dinleme/izleme metninin hiçbiri şiir türünde değildir. Şiirin dinleme becerisiyle kurduğu doğal bağ [3] göz önünde bulundurulduğunda bu dağılım işlevsel bir eksiklik olarak karşımıza çıkmaktadır. Ayrıca "Sağlıklı Yaşıyorum" temasında hiçbir şiire yer verilmemiş, "Atatürk'ü Tanımak" temasında ise şiir yalnızca karekodla sunulan bir dinleme materyali olarak bulunmuştur.

B. Biçim Özelliklerine İlişkin Bulgular

Basılı olarak sunulan altı şiirin biçim özellikleri Tablo 3'te verilmiştir.

Tablo 3. Şiirlerin biçim özellikleri

Şiir	Nazım birimi	Ölçü	Uyak dizgesi	Dize
Oyuncakların Oyuncağı Olmak	Bent (düzensiz)	Serbest	Uyaksız	26
Türkiye	Bent (düzensiz)	Serbest	Kısmen uyaklı	24
Uçan Balon	Bent (tek)	Serbest	Uyaksız	8
Kırık At	Dörtlük + beyit	8'li hece	Çapraz + düz	47
Misafir Odası	Dörtlük	Serbest	Çapraz (xaxa)	16
Gülü İncitme Gönül	Dörtlük	7'li hece	abab / aaab + redif	16

Tablo 3'e göre altı şiirin dördü serbest ölçüyle, ikisi hece ölçüsüyle yazılmıştır. Şiirlerin üçü dörtlük nazım birimine sahipken üçü düzensiz bentlerden oluşmaktadır. Uyak bakımından ise dört şiirde uyak bulunmaktayken iki şiir uyaksızdır. Ritim ve ses tekrarlarının çocukta şiire yönelik ilk ilgiyi oluşturduğu ve ezberlemeyi kolaylaştırdığı dikkate alındığında [3], [27] hece ölçüsüne dayalı örneklerin sınırlı kalması bir eksiklik olarak değerlendirilebilir. Bu özelliklerin bir arada görüldüğü şiirlerden biri "Gülü İncitme Gönül"dür. Şiirde *"Çiçeklerle hoş geçin, / Balı incitme gönül. / Bir küçük meyve için / Dalı incitme gönül."* dizelerinde görüldüğü üzere yedili hece ölçüsü, tam uyak ve "incitme gönül" redifi bir arada kullanılmakta; böylece metin âdeta ezberlenmeye elverişli bir ses dizgesi kurmaktadır.

Ders kitabında biçim öğretiminin şiir metinleri üzerinden yürütüldüğü de görülmektedir. "Türkiye" şiirinin ardından gelen etkinlikte öğrencilerden okudukları şiiri düzyazıdan ayıran özellikleri belirlemeleri ve bu özelliklerin şiire katkısını açıklamaları istenmektedir. Benzer biçimde söz sanatlarının öğretimi de büyük ölçüde şiirlere bağlanmıştır. "Düş Oyunu" şiirinde kişileştirme, "Atatürk Kurtuluş Savaşı'nda" şiirinde ise *"Bir gemi yanaştı Samsun'a sabaha karşı / Selam durdu kayığı, çaparı, takası"* dizeleri üzerinden kişileştirme ve benzetme çalıştırılmaktadır. Söz konusu tercih, şiirin biçim ve ahenk ögelerinin öğretiminde işlevsel bir araç hâline getirildiğini göstermektedir; ancak bu durum aynı zamanda şiirin estetik bir yaşantı olmaktan çıkarılıp dil bilgisi ve söz sanatı öğretimine indirgenmesi riskini de taşımaktadır.

C. İçerik ve Değer Aktarımına İlişkin Bulgular

Şiirlerin taşıdığı temalar ile Türkiye Yüzyılı Maarif Modeli'nin değerler çerçevesi bakımından öne çıkan değerler Tablo 4'te sunulmuştur.

Tablo 4. Şiirlerin içerik ve değer görünümü

Şiir	**Tema/konu**	**Öne çıkan değerler**
Oyuncakların Oyuncağı Olmak	Geleneksel oyuncak, dijitalleşme	Kültürel miras, aile bağı, sorumluluk
Türkiye	Vatan, doğa, halk kültürü	Vatanseverlik, estetik, çalışkanlık
Uçan Balon	Özgürlük ve yalnızlık	Özgürlük, dostluk, duyarlılık
Kırık At	Emek, zanaat, usta-çırak	Çalışkanlık, sabır, emeğe saygı
Misafir Odası	Ev, aile belleği, misafirlik	Aile bütünlüğü, misafirperverlik, kültürel miras
Gülü İncitme Gönül	Nezaket ve öğüt	Saygı, mütevazılık, merhamet, dürüstlük

Tablo 4 incelendiğinde şiirlerde sevgi, saygı, çalışkanlık, aile bütünlüğü ve vatanseverlik değerlerinin öne çıktığı görülmektedir. Bununla birlikte adalet, tasarruf, temizlik ve sağlıklı yaşam gibi değerlere şiirler aracılığıyla hiç değinilmemiştir. Bu durum, şiirlerin temalara dengesiz dağılmasının doğrudan bir sonucudur; nitekim "Sağlıklı Yaşıyorum" temasında şiire yer verilmemesi, ilgili değer alanının şiir yoluyla işlenmesinin önüne geçmektedir.

Şiirlerde değerlerin iki farklı biçimde aktarıldığı görülmektedir. "Kırık At" şiirinde emek ve sabır değerleri bir olay örgüsü içinde somutlaştırılmakta ve şiir *"Demek, biraz iyi işler / Birçok alın teri ister!"* dizeleriyle çocuğun kendi çıkarımını yapmasına imkân tanıyacak biçimde sonlanmaktadır. Buna karşılık "Gülü İncitme Gönül" şiirinde değerler doğrudan öğüt cümleleriyle aktarılmakta, hatta ders kitabındaki etkinlikte öğrencilerden "şairin şiirde vermek istediği öğütleri" bölümlerle eşleştirmeleri istenmektedir. Soyut değerlerin didaktik tanımlarla değil bir olay ya da kurgu aracılığıyla somutlaştırılarak sunulması gerektiği yönündeki ilke [6] dikkate alındığında ilk yaklaşımın çocuk edebiyatı ölçütlerine daha uygun olduğu söylenebilir.

D. Söz Varlığına İlişkin Bulgular

Şiirlerin söz varlığını geliştirme işlevi, hem metinlerin barındırdığı söz varlığı ögeleri hem de bu ögelere yönelik etkinlik desteği bakımından incelenmiş ve elde edilen bulgular Tablo 5'te gösterilmiştir.

Tablo 5. Şiirlerin söz varlığı görünümü ve etkinlik desteği

Şiir	**Çalışılan söz varlığı ögeleri**	**Sayı**
Türkiye	ekin, gütmek, bayır, katık, urba, banmak, saz, saadet	8
Misafir Odası	göz bebeği, kadife, büfe, konsol	4
Gülü İncitme Gönül	geçinmek, mahsus, hayır, engin	4
Oyuncakların Oyuncağı Olmak	Etkinlik desteği yok	0
Uçan Balon	Etkinlik desteği yok	0
Kırık At	Etkinlik desteği yok	0
Toplam		**16**

Tablo 5'e göre söz varlığı çalışmaları yalnızca ana okuma metni işlevindeki üç şiirle sınırlıdır. Bu üç şiir üzerinden on altı söz varlığı ögesi çalıştırılmakta; ayrıca "göz bebeği" kelime grubunun çok anlamlılığı ile "gönül" kelimesinin geçtiği atasözleri üzerinden kalıplaşmış söz varlığına yönelik çalışmalar yapılmaktadır. Bu etkinlikler, şiirlerin öğrencilerin söz varlığını geliştirmesine katkı sağlamaktadır [31].

Buna karşılık okuma keyfi ve etkinlik işleviyle sunulan üç şiirde herhangi bir söz varlığı çalışması bulunmamaktadır. Bu bulgu, "Kırık At" şiiri bakımından ayrıca dikkat çekicidir çünkü söz konusu şiir *nalbant, örs, körük, eğe, kıskaç, mıhlamak* gibi bugünün çocuğu için büyük ölçüde yabancı olan zengin bir meslek söz varlığı barındırmaktadır. Dolayısıyla ders kitabında en yoğun bilinmeyen kelimeyi taşıyan şiirin etkinlikle desteklenmemesi, metinden beklenen kazanımın gerçekleşmesini güçleştirmektedir.

E. Dil ve Anlatım ile Çocuğa Görelik Ölçütüne İlişkin Bulgular

Şiirlerin dil ve anlatım özellikleri ile çocuğa görelik ölçütü bakımından değerlendirilmesine ilişkin bulgular Tablo 6'da sunulmuştur.

Tablo 6. Şiirlerin dil ve anlatım ile çocuğa görelik bakımından değerlendirilmesi

Şiir	**Anlatıcı/bakış açısı**	**Dil ve anlatım**	**Çocuğa görelik**
Uçan Balon	Kişileştirilmiş nesne	Yalın, somut	Uygun
Kırık At	Çocuk anlatıcı	Arkaik meslek söz varlığı	Kısmen uygun
Misafir Odası	Kişileştirilmiş mekân	Nostaljik eşya adları	Kısmen uygun
Gülü İncitme Gönül	Öğüt veren yetişkin	Akıcı fakat soyut	Kısmen uygun
Oyuncakların Oyuncağı Olmak	Çocuk-baba diyaloğu	İmge yoğun	Kısmen uygun
Türkiye	Yetişkin lirik özne	Arkaik ve soyut	Sınırlı

Tablo 6'ya göre yalnızca bir şiir çocuğa görelik ölçütüne bütünüyle uygun bulunmuş, dört şiir kısmen uygun, bir şiir ise sınırlı düzeyde uygun olarak değerlendirilmiştir. Çocuğa görelik bakımından en güçlü örnek "Uçan Balon" şiiridir. Şiirde *"Ben bir uçan balondum / Minik bir çocuğun sıcacık ellerinde"* dizeleriyle kurulan kişileştirme, çocuğun kendisini metnin öznesiyle özdeşleştirmesine imkân tanımakta; yalın söz varlığı ve somut anlatım ise metnin anlaşılmasını kolaylaştırmaktadır.

Buna karşılık "Türkiye" şiiri, çocuğa görelik bakımından en sorunlu metin olarak karşımıza çıkmaktadır. Şiirdeki lirik özne bir yetişkindir ve *"Ormanlar urban olmuş yeşil yeşil"* ya da *"Sıcak kucağında saadetim"* gibi dizelerde yer alan sözcükler çocuğun yaşantı alanında doğrudan bir karşılık bulmamaktadır. Ayrıca *"Sınırlarında omuz omuza yatan insanlarına..."* dizesinin taşıdığı örtük ölüm imgesi, beşinci sınıf düzeyindeki bir öğrencinin bağımsız olarak çözümleyebileceği bir anlatım değildir. Benzer biçimde "Misafir Odası" şiirinde geçen *konsol, dantelâ, guguk kuşu, misafir kolonyası* gibi ögelerin, bugünün çocuğunun gündelik yaşantısından uzak olduğu söylenebilir. Dolayısıyla söz konusu şiirlerde yetişkin belleği çocuk gerçekliğinin önüne geçmektedir.

Şairler bakımından yapılan incelemede ise ders kitabındaki sekiz şiirin yedisinin şairinin belirtildiği, birinin ise şair adı verilmeksizin karekodla sunulduğu görülmüştür. Adı belirtilen yedi şairin beşi yaşayan veya çağdaş dönem şairi, ikisi ise klasikleşmiş isimlerdir. Ayrıca şairlerin tamamının erkek olması, ders kitabındaki şiir seçkisinin temsil çeşitliliği bakımından gözden geçirilmesi gerektiğini göstermektedir.

IV. TARTIŞMA

Araştırmada elde edilen en önemli bulgu, şiirin ders kitabındaki temsil oranının öğretim programının öngördüğü denge ilkesinin altında kalmasıdır. İncelenen kitapta temalarda yer alan otuz metnin beşi şiir türündedir ve bu oran %16,67'ye karşılık gelmektedir. Söz konusu bulgu, ortaokul ders kitaplarındaki 438 metnin 79'unun şiir olduğunu belirleyen Kemiksiz'in [14] bulgusuyla (%18,04) büyük ölçüde örtüşmektedir. Benzer biçimde Türkçe ders kitaplarında hikâye edici ve bilgilendirici metinlerin baskın olduğunu ortaya koyan çalışmalar [29], [30] da bu tabloyu desteklemektedir. Dolayısıyla öğretim programı değişmiş olsa dahi ders kitaplarındaki tür dağılımının uzun süredir devam eden bir eğilimi koruduğu söylenebilir.

Şiirin dinleme/izleme metinleri arasında hiç yer almaması, bu araştırmaya özgü ve üzerinde durulması gereken bir bulgudur. Çocuğun şiirle kurduğu ilk ilişkinin dinleme yoluyla gerçekleştiği [3] göz önünde bulundurulduğunda söz konusu tercih, şiirin en doğal öğretim kanalının kullanılmaması anlamına gelmektedir. Nitekim kitapta şiire dayalı iki dinleme çalışması bulunmakta fakat bunlar tema metni statüsünde değil etkinlik materyali olarak ve karekod aracılığıyla sunulmaktadır. Bu durum, dijital erişimin bulunmadığı sınıf ortamlarında şiirin bütünüyle devre dışı kalmasına yol açabilecek bir uygulamadır.

Basılı şiirlerin %83,33'ünün kısaltılarak aktarılması da tartışılması gereken bir başka bulgudur. Şiir, ögeleri arasında sıkı bir bağ bulunan ve anlamı yoğunlaştırılmış bir tür olduğundan dizelerin eksiltilmesi metnin ahenk dizgesini ve anlam bütünlüğünü zedelemektedir. Bayat'ın [16] ders kitaplarındaki şiirleri şiirbilim ölçütleri bakımından yetersiz bulması, bu araştırmanın bulgusuyla aynı doğrultudadır. Ders kitabına alınacak metinlerin estetik bütünlüğünün korunması gerektiği yönündeki ilke [33] dikkate

alındığında kısaltmanın zorunlu olduğu durumlarda bütünlüğü koruyan bölümlerin tercih edilmesi gerekmektedir.

Değer aktarımı boyutunda elde edilen bulgular, Çapoğlu ve Okur'un [17] 8. sınıf ders kitaplarındaki şiirlerde en çok insan sevgisi, doğa sevgisi ve Türk büyüklerine saygı değerlerinin vurgulandığı yönündeki sonucuyla benzerlik göstermektedir. Bu araştırmada da en sık karşılaşılan değerler sevgi, saygı ve vatanseverliktir. Ancak bu değerlerin şiirlerde sunuluş biçimleri birbirinden farklıdır. Değerin bir olay örgüsü içinde somutlaştırıldığı örnekler çocuk edebiyatı ölçütlerine uygunluk gösterirken doğrudan öğüt aktarımına dayanan örnekler didaktik bir söyleme yaklaşmaktadır. Cumhuriyet dönemi çocuk edebiyatının uzun süre güdümlü ve didaktik bir yapı taşıdığı [8], [13] göz önünde bulundurulduğunda bu eğilimin tarihsel bir sürekliliğe işaret ettiği söylenebilir.

Çocuğa görelik boyutunda ulaşılan sonuçlar, ders kitaplarındaki şiirleri aynı ölçüt doğrultusunda inceleyen çalışmaların bulgularıyla paralellik göstermektedir. Çolak [18] ve Günesu [19] da ortaokul ve 5. sınıf ders kitaplarındaki şiirlerin çocuğa görelik bakımından kısmen yeterli olduğunu belirlemişlerdir. Bu araştırmada söz konusu yetersizliğin başlıca kaynağının çocuğun gündelik dilinden uzak söz varlığı ile yetişkin bakış açısı olduğu tespit edilmiştir. Çocuğun dilsel dünyasının ancak kendi tecrübe ettiği yaşantı alanı içinde biçimlenebileceği [6], [7] dikkate alındığında ders kitabına alınacak şiirlerde çocuğun gündelik yaşantısıyla bağ kurabileceği somut ögelerin öncelenmesi gerekmektedir.

V. SONUÇ VE ÖNERİLER

Bu araştırmada Türkiye Yüzyılı Maarif Modeli doğrultusunda hazırlanan 5. sınıf Türkçe ders kitabındaki şiirler çocuk edebiyatı ve Türkçe öğretimi bakımından incelenmiştir. Elde edilen bulgular doğrultusunda şu sonuçlara ulaşılmıştır:

Ders kitabında sekiz şiir yer almakta, bunlardan altısı basılı olarak sunulmaktadır. Temalarda yer alan otuz metnin yalnızca beşi şiir türünde olduğundan öğretim programının öngördüğü tür dengesi sağlanamamıştır. Şiirler temalara eşit dağılmamış, bir temada hiçbir şiire yer verilmemiştir. Dinleme/izleme metinleri arasında şiir bulunmamaktadır.

Biçim bakımından şiirlerin üçte ikisi serbest ölçüyle yazılmış, hece ölçüsüne ve düzenli uyak dizgesine dayalı örnekler sınırlı kalmıştır. Değer aktarımı bakımından şiirler zengin bir görünüm sunmakta ancak değerler kimi metinlerde somutlaştırılarak kimi metinlerde ise doğrudan öğüt yoluyla aktarılmaktadır. Söz varlığı çalışmaları yalnızca ana okuma metni işlevindeki şiirlerle sınırlıdır. Çocuğa görelik bakımından şiirlerin büyük bölümü kısmen uygun bulunmuş, günlük dilden uzak söz varlığı ile yetişkin bakış açısı temel sınırlılık olarak belirlenmiştir.

Bu sonuçlar doğrultusunda şu öneriler geliştirilmiştir: Ders kitabı hazırlama süreçlerinde metin biçimleri arasındaki denge tema düzeyinde gözetilmeli, her temada en az bir şiire yer verilmelidir. Dinleme/izleme metinleri arasına şiir eklenmeli, karekodla sunulan şiirlerin tam metinleri kitapta da basılı olarak bulunmalıdır. Şiirlerin kısaltılmasından mümkün olduğunca kaçınılmalı, zorunlu durumlarda anlam ve ahenk bütünlüğünü koruyan bölümler tercih edilmelidir. Okuma keyfi ve etkinlik işlevindeki şiirler de söz varlığı çalışmalarıyla desteklenmelidir. Metin seçiminde çocuğun gündelik yaşantısıyla bağ kurabileceği somut ögeler öncelenmeli, şair ve tür çeşitliliği gözetilmelidir. Bundan sonraki araştırmalarda diğer sınıf düzeylerine ait ders kitapları da aynı ölçütlerle incelenerek ortaokul kademesinin bütününe ilişkin karşılaştırmalı bir görünüm ortaya konulabilir.

KAYNAKLAR

[1] S. Sever, *Çocuk Edebiyatı ve Okuma Kültürü*. İzmir: Tudem Yayınları, 2013.
[2] D. Aksan, *Şiir Dili ve Türk Şiir Dili*. Ankara: Engin Yayınevi, 2006.
[3] A. Yalçın ve G. Aytaş, *Çocuk Edebiyatı*. Ankara: Akçağ Yayınları, 2017.
[4] H. Akyol, *Türkçe Öğretim Yöntemleri*. Ankara: Kök Yayıncılık, 2006.
[5] A. F. Oğuzkan, *Çocuk Edebiyatı*. Ankara: Anı Yayıncılık, 2013.
[6] E. N. Gökşen, *Örnekleriyle Çocuk Edebiyatımız*. İstanbul: Remzi Kitabevi, 1985.
[7] E. Çer, "Türkçe öğretiminde çocuğa görelik ilkesine uygun edebiyat yapıtlarının önemi," *İlköğretim Online*, c. 15, sy. 4, ss. 1399-1410, 2016.
[8] M. R. Şirin, *99 Soruda Çocuk Edebiyatı*. İstanbul: Çocuk Vakfı Yayınları, 2000.

[9] Ş. Dilidüzgün, *Çağdaş Çocuk Yazını*. İstanbul: Yapı Kredi Yayınları, 1996.
[10] H. Güleryüz, *Yaratıcı Çocuk Edebiyatı*. Ankara: Pegem Yayıncılık, 2002.
[11] Millî Eğitim Bakanlığı, *Türkiye Yüzyılı Maarif Modeli Öğretim Programları Ortak Metni*. Ankara: Talim ve Terbiye Kurulu Başkanlığı, 2024.
[12] Millî Eğitim Bakanlığı, *Türkçe Dersi Öğretim Programı (5, 6, 7 ve 8. Sınıflar)*. Ankara: Talim ve Terbiye Kurulu Başkanlığı, 2024.
[13] H. Tezgör, *"Bin Atlı Akınlarda Çocuklar": Ortaokul Türkçe Ders Kitaplarında Şiir (1929-2005)*. İstanbul: İletişim Yayınları, 2013.
[14] Ö. Kemiksiz, "Ortaokul Türkçe ders kitaplarında şiirin yeri," *Kesit Akademi Dergisi*, sy. 4, ss. 284-303, 2016.
[15] A. N. Ercan, "8. sınıf Türkçe ders kitaplarındaki şiirlerin şekil, tema ve tür açılarından incelenmesi," *Ana Dili Eğitimi Dergisi*, c. 2, sy. 2, ss. 15-27, 2014.
[16] N. Bayat, "Türkçe ders kitaplarında yer alan şiirlerin şiirbilim açısından görünümü üstüne bir inceleme," *Turkish Studies*, c. 7, sy. 1, ss. 325-344, 2012.
[17] E. Çapoğlu ve A. Okur, "Ortaokul 8. sınıf Türkçe ders kitaplarındaki şiirlerde yer alan değerler," *Sakarya University Journal of Education*, c. 5, sy. 3, ss. 90-104, 2015.
[18] M. T. Çolak, "Ortaokul Türkçe ders kitaplarındaki şiirlerin çocuğa görelik ilkeleri açısından incelenmesi ve tavsiye şiirler seçkisi," Yüksek lisans tezi, Ondokuz Mayıs Üniversitesi, Samsun, 2022.
[19] Y. Günesu, "Beşinci sınıf Türkçe ders kitaplarındaki şiirlerin 'çocuğa görelik' ilkesi doğrultusunda incelenmesi," Yüksek lisans tezi, İstanbul Üniversitesi, İstanbul, 2015.
[20] A. Yıldırım ve H. Şimşek, *Sosyal Bilimlerde Nitel Araştırma Yöntemleri*. Ankara: Seçkin Yayıncılık, 2021.
[21] S. B. Merriam, *Nitel Araştırma: Desen ve Uygulama İçin Bir Rehber* (Çev. Ed. S. Turan). Ankara: Nobel Yayıncılık, 2018.
[22] A. Cin, S. Dalioğlu, M. Kavas vd., *Ortaokul ve İmam Hatip Ortaokulu Türkçe 5 Ders Kitabı (1. Kitap)*. Ankara: Millî Eğitim Bakanlığı Yayınları, 2025.
[23] A. Cin, S. Dalioğlu, M. Kavas vd., *Ortaokul ve İmam Hatip Ortaokulu Türkçe 5 Ders Kitabı (2. Kitap)*. Ankara: Millî Eğitim Bakanlığı Yayınları, 2025.
[24] M. Q. Patton, *Nitel Araştırma ve Değerlendirme Yöntemleri* (Çev. Ed. M. Bütün ve S. B. Demir). Ankara: Pegem Akademi, 2014.
[25] M. B. Miles ve A. M. Huberman, *Qualitative Data Analysis: An Expanded Sourcebook*. Thousand Oaks: Sage Publications, 1994.
[26] K. Krippendorff, *Content Analysis: An Introduction to Its Methodology*. Thousand Oaks: Sage Publications, 2013.
[27] C. Dilçin, *Örneklerle Türk Şiir Bilgisi*. Ankara: Türk Dil Kurumu Yayınları, 2004.
[28] C. Kavcar, *Edebiyat ve Eğitim*. Ankara: Engin Yayınevi, 1999.
[29] M. Solak ve D. Yaylı, "İlköğretim ikinci kademe Türkçe ders kitaplarının türler açısından incelenmesi," *Uluslararası Sosyal Araştırmalar Dergisi*, c. 2, sy. 9, ss. 444-453, 2009.
[30] M. A. Çeçen ve Ö. Çiftçi, "İlköğretim 6. sınıf Türkçe ders kitaplarında yer alan metinlerin tür ve tema açısından incelenmesi," *Millî Eğitim Dergisi*, sy. 173, ss. 39-49, 2007.
[31] M. Mangır, "İlköğretim 5. sınıf Türkçe ders kitabındaki deyim varlığının öğrencilerin alıcı söz varlığına katkısı," *Turkish Studies*, c. 7, sy. 4, ss. 2371-2385, 2012.
[32] B. Göğüş, *Orta Dereceli Okullarımızda Türkçe ve Yazın Eğitimi*. Ankara: Gül Yayınevi, 1978.
[33] S. Sever, *Çocuk ve Edebiyat*. İzmir: Tudem Yayınları, 2010.
[34] Türk Dil Kurumu, *Türkçe Sözlük*. Ankara: Türk Dil Kurumu Yayınları, 2011.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Eğitimde Üretken Yapay Zekâ Araştırmaları Yeniden Üretilebilir mi? Raporlama Şeffaflığının İçerik Analizi (2023–2026)

Şener Balat*

Bilişim Teknolojileri Meslek Yüksekokulu, Bingöl Üniversitesi, Türkiye

(senerbalat@hotmail.com) Başlıca yazarın mail adresi

Özet – Üretken yapay zekâ (YZ) modelleri sürüm sürüm değiştiği için, kullanılan modelin sürümü, erişim tarihi, prompt metni ve üretim parametreleri bildirilmeyen bir çalışma ilkesel olarak yeniden üretilemez. Sağlık alanında bu bilgilerin raporlanmasını düzenleyen kılavuzlar geliştirilmişken, eğitim araştırmalarında eşdeğer bir çerçeve bulunmamaktadır. Bu çalışma, eğitim bağlamında bir üretken YZ sistemi işleten görgül araştırmaların raporlama şeffaflığını nicel içerik analiziyle belirlemeyi amaçlamaktadır. Web of Science Core Collection'ın Education & Educational Research kategorisinde 2023–2026 arasında yayımlanan 5.641 kayıt taranmış, başlık ve öz taramasından sonra 1.725 çalışmalık uygun havuz oluşturulmuş, bu havuzdan yıla göre tabakalı rastgele 300 çalışma seçilmiş ve tam metnine erişilebilen 89 çalışma kodlanmıştır. Kodlama, mevcut raporlama kılavuzlarından eğitim bağlamına uyarlanan altı alanlı ve 27 maddelik GAI-RTE çerçevesiyle yapılmış, her olumlu karar makaledeki kaynak cümleyle belgelenmiştir. Bulgular, genel yöntem raporlamasının yeterli düzeyde (ortalama %83,1), YZ'ye özgü raporlamanın ise çok düşük olduğunu göstermektedir (%19,7; Wilcoxon $p < 0,001$; $d_z = 2,41$). Çalışmaların %59,6'sı model sürümünü, %20,2'si en az bir promptu birebir, %4,5'i prompt setinin tamamını, %5,6'sı üretim parametrelerini bildirmiştir; model sürümünü, promptu ve parametreleri birlikte bildiren çalışma sayısı yalnızca ikidir. Bileşik şeffaflık skorunun ortalaması 31,1 olup hiçbir çalışma %56'yı aşmamıştır. Yıllara göre küçük ama anlamlı bir iyileşme gözlenmiştir ($\rho = 0,243$; $p = 0,022$). Sonuçlar, eğitim araştırmalarında üretken YZ kullanımı için asgari bir raporlama setine ivedilikle ihtiyaç duyulduğunu ortaya koymaktadır.

Anahtar Kelimeler – üretken yapay zekâ, raporlama şeffaflığı, yeniden üretilebilirlik, içerik analizi, eğitim araştırmaları, açık bilim

I. GİRİŞ

Üretken yapay zekâ (YZ) sistemleri, 2023 yılından bu yana eğitim araştırmalarının en hızlı büyüyen konu alanlarından biri hâline gelmiştir. Bu çalışmaların büyük bölümü, bir dil modelini yalnızca tartışma konusu olarak ele almakla kalmaz, onu araştırma sürecinin içinde fiilen işletir: modele geri bildirim ürettirilir, sınav maddesi yazdırılır, öğrenci metni puanlatılır ya da öğrenciler bir sohbet botuyla etkileşime sokulur. Bu durumda model, çalışmanın bir aracı değil, işlemsel bir bileşenidir ve tıpkı bir ölçme aracı ya da deneysel işlem gibi ayrıntılı biçimde tanımlanmayı gerektirir.

Sorun, bu bileşenin kararlı olmamasıdır. Ticari üretken YZ sistemleri sürüm sürüm güncellenir; aynı ürün adı altında sunulan model, aylar içinde farklı davranışlar sergileyebilir. Üstelik bu sistemler olasılıksal üretim yapar; aynı yönerge aynı modele iki kez verildiğinde farklı çıktılar alınabilir. Bu iki özellik birlikte düşünüldüğünde, hangi modelin hangi sürümünün, hangi tarihte, hangi yönergeyle ve hangi üretim ayarlarıyla çalıştırıldığı bildirilmediği sürece bir çalışmanın yinelenmesi ilkesel olarak olanaksız hâle gelir. Raporlanmayan bilgi yalnızca eksik bir ayrıntı değildir; çalışmanın yeniden üretilebilirliğini ortadan kaldıran yapısal bir boşluktur.

Sağlık alanı bu sorunu erken fark etmiş ve yanıt üretmiştir. Üretken YZ tabanlı sohbet botlarının başarımını değerlendiren çalışmalar için geliştirilen CHART bildirisi, 531 paydaşın katıldığı bir Delfi uzlaşı süreci sonunda 12 madde ve 39 alt maddeden oluşan bir kontrol listesi ortaya koymuştur [1]. Büyük dil modeli kullanan çalışmalar için geliştirilen TRIPOD-LLM ise 19 ana madde ve 50 alt madde içermektedir [2]. Bu iki kılavuz, makine öğrenmesi temelli kestirim modelleri için hazırlanan TRIPOD+AI çerçevesinin [3] uzantısı niteliğindedir. Bunlara ek olarak, üretken YZ'nin akademik üretimde kullanımına ilişkin CANGARU girişimi de sürdürülmektedir [4]. Yayıncı düzeyindeki düzenlemelerin ise dağınık olduğu gösterilmiştir: en büyük 100 yayıncının yalnızca %17'si üretken YZ kullanımına ilişkin yönerge sunmaktadır [5].

Eğitim araştırmalarında ise eşdeğer bir kılavuz bulunmamaktadır. Alandaki derleme çalışmaları, üretken YZ uygulamalarının etkililiğine ve etkileşim biçimlerine odaklanmakta [6], etik ve uygulama güçlüklerini haritalamakta [7], ancak çalışmaların kendi yöntemlerini ne ölçüde şeffaf raporladığını sistematik olarak ölçmemektedir. Yakın tarihli bir sistematik derleme, incelenen çalışmaların istatistiksel raporlamasının yeterli ayrıntıda olmadığına dikkat çekmiş [6], fakat bunu bir yan gözlem olarak aktarmış, raporlama kalitesini bir ölçüm nesnesi hâline getirmemiştir. Dolayısıyla eğitim alanında şu soru yanıtsızdır: Üretken YZ kullanan görgül eğitim araştırmaları, bir başka araştırmacının aynı işlemi yineleyebilmesi için gereken bilgileri ne ölçüde bildirmektedir?

Bu çalışma söz konusu boşluğu doldurmayı amaçlamaktadır. Bunun için mevcut raporlama kılavuzlarından eğitim bağlamına uyarlanan GAI-RTE (Generative AI Reporting Transparency in Education) adlı altı alanlı ve 27 maddelik bir çerçeve geliştirilmiş ve 2023-2026 arasında yayımlanan görgül çalışmalardan oluşan olasılıklı bir örneklem üzerinde uygulanmıştır. Araştırma soruları şunlardır: (1) Çalışmalar model kimliğini, prompt bilgisini, üretim ayarlarını ve çıktı yönetimini ne ölçüde raporlamaktadır? (2) YZ'ye özgü raporlama ile genel yöntem raporlaması arasında fark var mıdır? (3) Raporlama şeffaflığı yayın yılına, dergi dizin düzeyine ve erişim durumuna göre farklılaşmakta mıdır?

II. MATERYAL VE YÖNTEM

A. Araştırma Deseni

Çalışma, yönlendirilmiş (directed) nicel içerik analizi desenindedir. Kodlama kategorileri veriden türetilmemiş, kuramsal olarak önceden belirlenmiş kılavuz maddelerinden uyarlanmıştır. Örneklem seçimi PRISMA 2020 akışı [8] uyarlanarak belgelenmiştir.

B. Örnekleme Çerçevesi ve Tarama

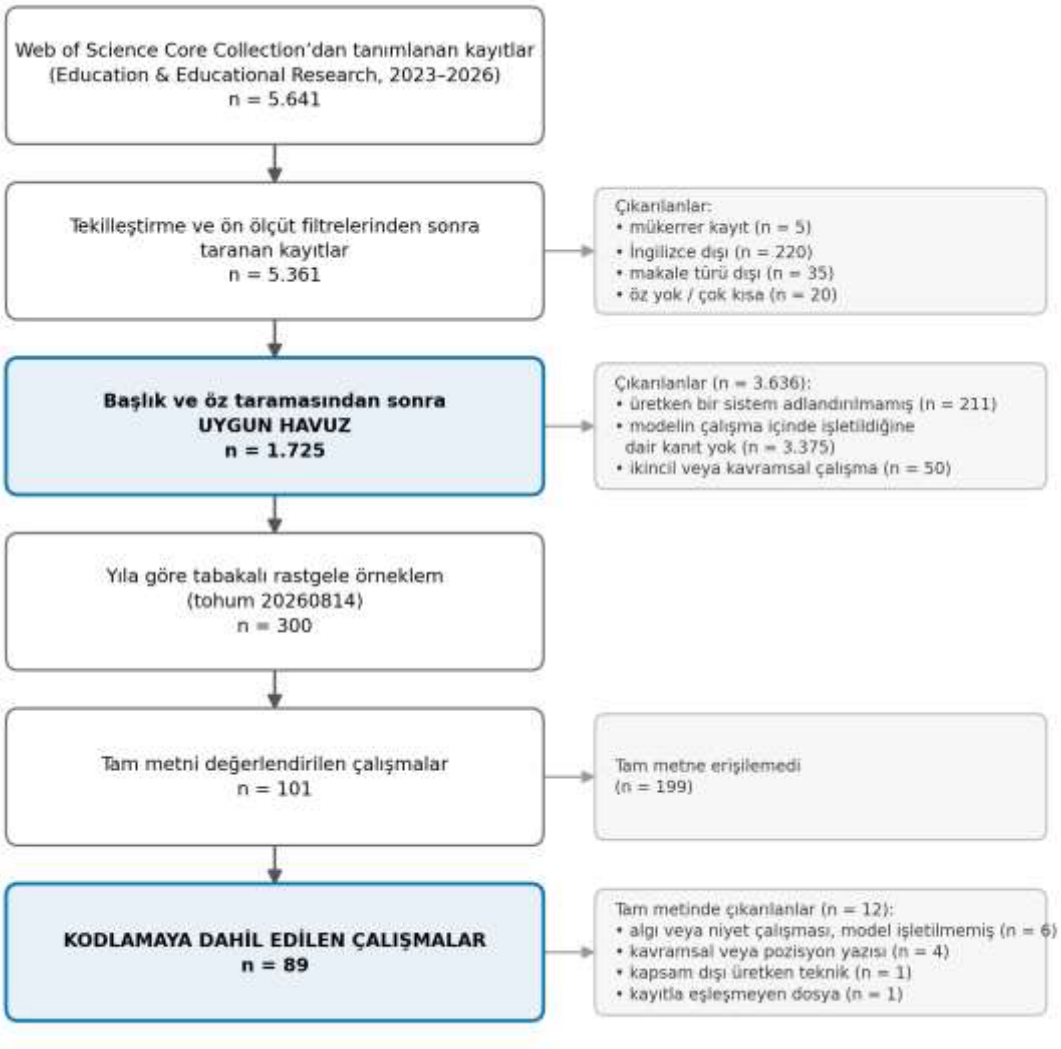
Popülasyon, Web of Science Core Collection'ın Education & Educational Research kategorisinde 2023-2026 arasında yayımlanan makaleler olarak tanımlanmıştır. Bu tanım bilinçli bir tercihtir: kategori, eğitim araştırmalarının dizinlenmiş dergi kümesini sınırları belli biçimde tanımladığı için eksik bir arama değil, tanımlı bir popülasyon oluşturmaktadır.

Tarama, üretken YZ terimleri (üretken yapay zekâ, ChatGPT, GPT, büyük dil modeli, Gemini, Copilot, Claude, DeepSeek, Llama, DALL-E, Midjourney vb.) ile eğitim terimlerinin birleşimiyle yapılmış ve 5.641 kayıt elde edilmiştir. Tekilleştirme sonrası 5 mükerrer kayıt çıkarılmış; İngilizce dışı 220, makale türü dışı 35 ve özü bulunmayan ya da çok kısa olan 20 kayıt elenerek 5.361 kayıt taranmıştır.

Başlık ve öz taramasında üç ölçüt birlikte aranmıştır: (a) belirli bir üretken sistemin adlandırılması, (b) modelin çalışma içinde fiilen işletildiğine ilişkin güçlü kanıt, (c) çalışmanın ikincil ya da kavramsal nitelikte olmaması. Tarama kuralları, araştırmacı tarafından elle okunan 49 öz üzerinde kalibre edilmiş ve kesinliği yaklaşık %90'a çıkarılmıştır. Bu aşamada 211 kayıt üretken bir sistem adlandırmadığı, 3.375 kayıt modelin işletildiğine dair kanıt taşımadığı ve 50 kayıt ikincil ya da kavramsal nitelikte olduğu için elenmiş, geriye 1.725 çalışmalık uygun havuz kalmıştır.

C. Örneklem

Uygun havuzdan yayın yılına göre tabakalı ve orantılı dağılımlı rastgele örneklem çekilmiştir (n = 300; tohum değeri 20260814; 2023: 7, 2024: 39, 2025: 113, 2026: 141). Tam metinlerin 101'ine erişilebilmiş, bunlardan 12'si tam metin incelemesinde çıkarılmış ve nihai örneklem 89 çalışma olmuştur (Şekil 1). Çıkarma gerekçeleri: model işletmeyen algı ya da niyet çalışması (6), kavramsal veya pozisyon yazısı (4), kapsam dışı üretken teknik (1) ve kayıtla eşleşmeyen dosya (1).



Şekil 1. Örneklem seçim akışı (PRISMA 2020 uyarlaması).

D. GAI-RTE Kodlama Çerçevesi

Kodlama çerçevesi, CHART [1], TRIPOD-LLM [2] ve TRIPOD+AI [3] maddelerinin eğitim bağlamına uyarlanmasıyla oluşturulmuş; altı alan ve 27 maddeden meydana gelmektedir (Tablo I). Çerçeve yalnızca bilginin raporlanıp raporlanmadığını ölçer, çalışmanın yöntemsel niteliği hakkında yargı üretmez.

Tablo I. GAI-RTE Çerçevesinin Alanları ve Maddeleri

Alan	Maddeler
A. Model kimliği (6)	A1 sistem adlandırılmış; A2 model sürümü; A3 erişim biçimi (web arayüzü/API); A4 erişim veya veri toplama tarihi; A5 sağlayıcı; A6 erişim düzeyi veya abonelik
B. Prompt ve etkileşim (6)	B1 prompt geliştirme süreci; B2 en az bir prompt birebir; B3 prompt setinin tamamı erişilebilir; B4 prompt dili; B5 etkileşim protokolü (tur sayısı, yeni oturum); B6 tekrar üretim ve çıktı değişkenliği
C. Yapılandırma ve çıktı (4)	C1 üretim parametreleri veya varsayılan ayar beyanı; C2 çıktıların kaydı ve arşivlenmesi; C3 çıktının insan tarafından doğrulanması; C4 hatalı veya halüsinasyonlu çıktının ele alınışı
D. Yöntem şeffaflığı (4)	D1 araştırma deseni adlandırılmış; D2 örneklem veya analiz birimi tanımlanmış; D3 geçerlik ve güvenirlik kanıtı; D4 analitik ayrıntı
E. Açıklık (3)	E1 veri erişilebilirliği beyanı; E2 depo bağlantısı; E3 ön kayıt
F. Etik ve hesap verebilirlik (4)	F1 etik kurul onayı; F2 bilgilendirilmiş onam; F3 yazarların kendi YZ kullanım beyanı; F4 veri gizliliği önlemi

E. Kodlama Süreci

Tam metinler PDF'ten çıkarılmış, kaynakça bölümleri tarama dışında bırakılmıştır. Kaynakça sınırı, 89 metnin 87'sinde başlık yoluyla, kaynakçasını başlıksız basan iki dergide ise yıl ve cilt-sayfa yoğunluğuna dayalı bir yedek yöntemle belirlenmiştir.

Kodlama üç katmanlıdır. İlk katmanda, her madde için bağlam duyarlı düzenli ifadeler ve olumsuz filtrelerle aday cümleler saptanmıştır. İkinci katmanda, saptanan her olumlu adayın kaynak cümlesi araştırmacı tarafından okunmuş ve uymayanlar gerekçesi kaydedilerek sıfırlanmıştır; bu süreçte 82 elle yargılama yapılmıştır. Üçüncü katmanda, düşük yaygınlıklı maddelerde sıfır kodlanan metinler daha geniş anahtar sözcüklerle yeniden taranarak kaçırılmış kanıt aranmış ve bu yolla dört ek olumlu kod saptanmıştır.

Elle denetimde ayıklanan yanlış eşleşme türleri, katmanlı yaklaşımın gerekliliğini göstermektedir: öğrencilere verilen ödev yönergeleri model promptu olarak; iki puanlayıcının öğrenci verisini kodlaması YZ çıktısının doğrulanması olarak; giriş bölümündeki genel halüsinasyon uyarıları hata yönetimi olarak; bir fizik metnindeki sıcaklık değeri model sıcaklık parametresi olarak eşleşmiştir. Her olumlu koda ait kaynak cümle ve varsa yargılama gerekçesi ek materyalde yayımlanmaktadır.

F. Analiz

Madde düzeyinde oranlar ve Wilson %95 güven aralıkları [9] hesaplanmıştır. Alan skorları, ilgili alandaki maddelerin ortalaması; bileşik GAI-RTE skoru ise 27 maddenin ortalaması olarak 0-100 ölçeğine dönüştürülmüştür. YZ'ye özgü alanlar (A, B, C) ile genel yöntem alanı (D) Wilcoxon işaretli sıralar testiyle karşılaştırılmış, etki büyüklüğü d_z ile verilmiştir. Yıl eğilimi Spearman sırası korelasyonuyla; erişim durumu ve dizin düzeyine göre farklar Mann-Whitney U testi ve sıra çift serili korelasyonla incelenmiştir. Tam metne erişim yanlılığına karşı, havuzun açık erişim bileşimini geri getiren ters olasılık ağırlıklı kestirimler de raporlanmıştır. Çözümlemeler Python 3.12 (NumPy, SciPy, pandas) ile yapılmıştır.

III. BULGULAR

A. Örneklemin Profili

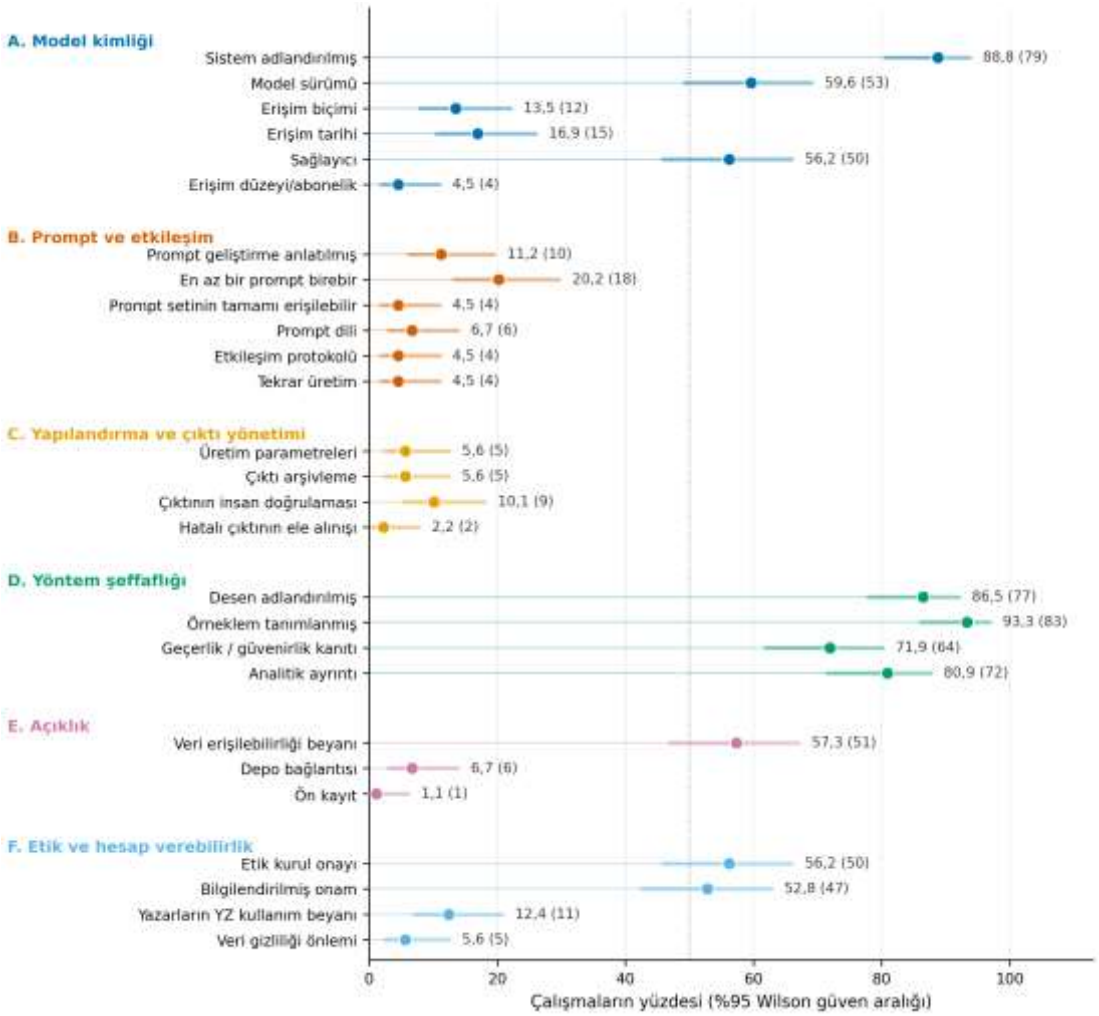
Nihai örneklem 89 çalışmadan oluşmaktadır (2023: 3, 2024: 13, 2025: 36, 2026: 37). Çalışmaların %81,5'i açık erişimli, %59,8'i SSCI dizinlidir.

B. Madde Düzeyinde Raporlama

Madde bazlı oranlar Şekil 2'de sunulmuştur. Genel yöntem raporlamasına ilişkin maddeler yüksek düzeydedir: örneklem tanımlama %93,3, desen adlandırma %86,5, analitik ayrıntı %80,9 ve geçerlik-güvenirlik kanıtı %71,9. Buna karşılık YZ'ye özgü maddelerde tablo tersine dönmektedir.

Model kimliği alanında çalışmaların %88,8'i kullanılan sistemi adlandırmakta, ancak yalnızca %59,6'sı model sürümünü belirtmektedir. Modelin hangi tarihte çalıştırıldığını bildiren çalışma oranı %16,9, erişim biçimini (web arayüzü veya API) bildirenlerin oranı ise %13,5'tir. Prompt alanında çalışmaların %20,2'si en az bir promptu birebir vermekte, prompt setinin tamamını erişilebilir kılanların oranı %4,5'te kalmaktadır. Etkileşim protokolünü ve tekrar üretim yordamını bildirenlerin oranı da %4,5'tir. Üretim parametrelerini bildiren ya da varsayılan ayarların kullanıldığını açıkça belirten çalışmaların oranı %5,6, çıktıların nasıl arşivlendiğini bildirenlerin oranı %5,6, hatalı çıktının ele alınışını bildirenlerin oranı ise %2,2'dir.

Açıklık alanında çalışmaların %57,3'ünde bir veri erişilebilirliği beyanı bulunmakta, ancak yalnızca %6,7'si işleyen bir depo bağlantısı vermektedir; ön kayıt bildiren tek bir çalışma vardır (%1,1). Etik alanında etik kurul onayı %56,2, bilgilendirilmiş onam %52,8 oranında bildirilmişken, yazarların kendi YZ kullanımına ilişkin beyanı yalnızca %12,4'ünde bulunmaktadır.



Şekil 2. GAI-RTE maddelerinin raporlanma oranları ve Wilson %95 güven aralıkları (N = 89). Parantez içindeki değerler çalışma sayısını göstermektedir.

C. Alan Skorları ve Bileşik Skor

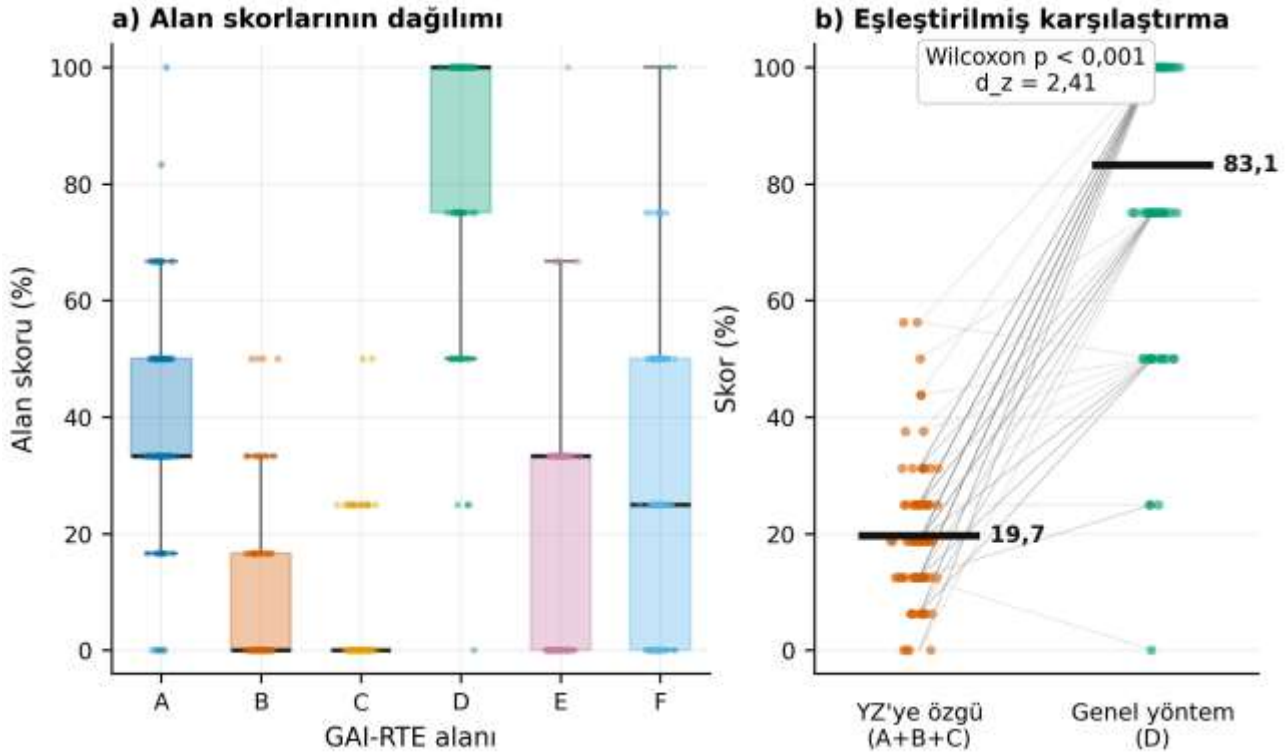
Alan ortalamaları Tablo II'de verilmiştir. En yüksek ortalama yöntem şeffaflığı alanına (%83,1), en düşük ortalamalar ise yapılandırma ve çıktı (%5,9) ile prompt ve etkileşim (%8,6) alanlarına aittir. Prompt alanında 59 çalışma, yapılandırma alanında 70 çalışma altı ya da dört maddenin tamamından sıfır almıştır.

Tablo II. Alan Skorlari Ve Bileşik Skor (N = 89)

Alan	Ortalama	SS	Ortanca
A. Model kimliği	39,9	18,4	33,3
B. Prompt ve etkileşim	8,6	14,0	0,0
C. Yapılandırma ve çıktı	5,9	11,9	0,0
D. Yöntem şeffaflığı	83,1	23,5	100,0
E. Açıklık	21,7	20,8	33,3
F. Etik ve hesap verebilirlik	31,7	25,8	25,0
Bileşik GAI-RTE	31,1	9,9	29,6

Bileşik GAI-RTE skorunun ortalaması 31,1 (SS = 9,9), ortancası 29,6, çeyrekler arası aralığı 22,2-37,0'dir. Dağılımın en düşük değeri 7,4, en yüksek değeri 55,6'dır; yani örneklemdeki hiçbir çalışma maddelerin %56'sından fazlasını raporlamamıştır ve yalnızca dört çalışma 50'nin üzerinde skor almıştır (Şekil 4a).

YZ'ye özgü alanların (A, B, C) ortalaması 19,7 iken genel yöntem alanının ortalaması 83,1'dir. Eşleştirilmiş karşılaştırma bu farkın büyük olduğunu göstermektedir (Wilcoxon işaretli sıralar testi, $p < 0,001$; $d_z = 2,41$). Fark, Şekil 3'te her çalışmanın kendi içindeki düşüş çizgisiyle görülebilmektedir.

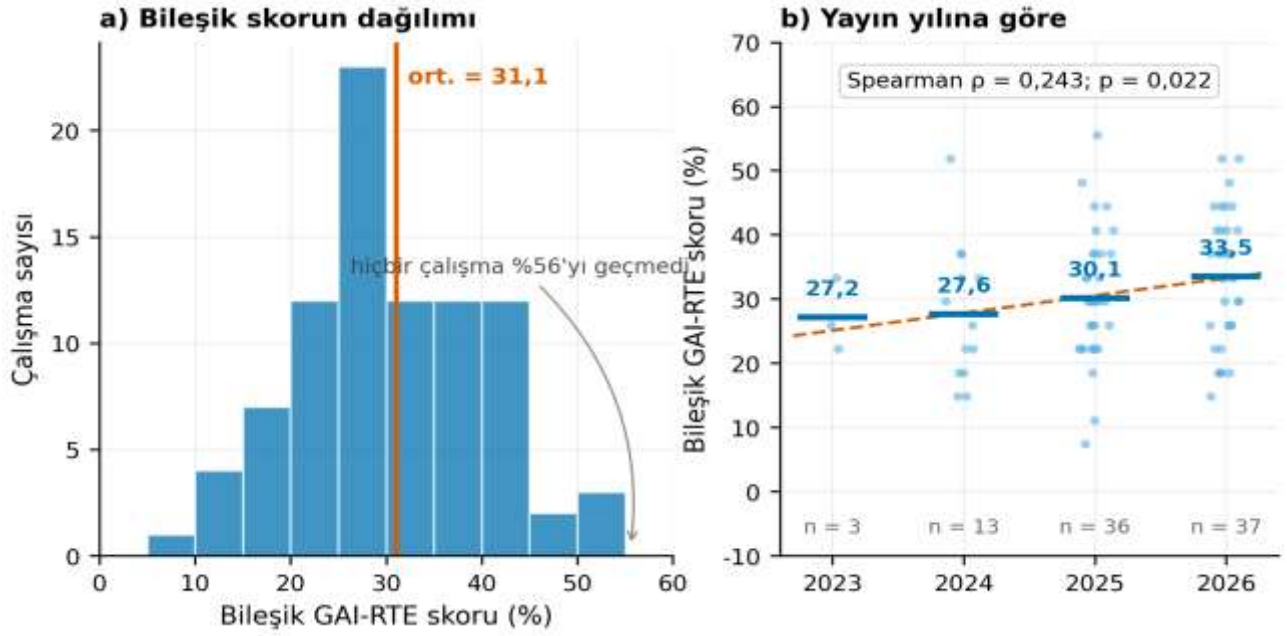


Şekil 3. Alan skorlarının dağılımı (a) ve yapay zekâya özgü alanlar ile genel yöntem alanının eşleştirilmiş karşılaştırması (b). Kalın yatay çizgiler ortalamaları göstermektedir.

Yeniden üretilebilirlik açısından en somut gösterge, maddelerin birlikte raporlanmasıdır. Model sürümünü ve modelin çalıştırıldığı tarihi birlikte bildiren çalışma sayısı 11'dir. Model sürümünü, en az bir promptu ve üretim parametrelerini birlikte bildiren çalışma sayısı ise yalnızca ikidir. Bu üç bilginin birlikte bulunmaması durumunda bir çalışmanın aynı koşullarda yinelenmesi olanaklı değildir.

D. Zaman İçindeki Değişim

Bileşik skorun yıllara göre ortalamaları 2023'te 27,2, 2024'te 27,6, 2025'te 30,1 ve 2026'da 33,5'tir. Yayın yılı ile bileşik skor arasındaki ilişki küçük ama anlamlıdır (Spearman ρ = 0,243; p = 0,022). İyileşme yönü olumlu olmakla birlikte, 2026 ortalaması bile maddelerin üçte birinin biraz üzerindedir (Şekil 4b).

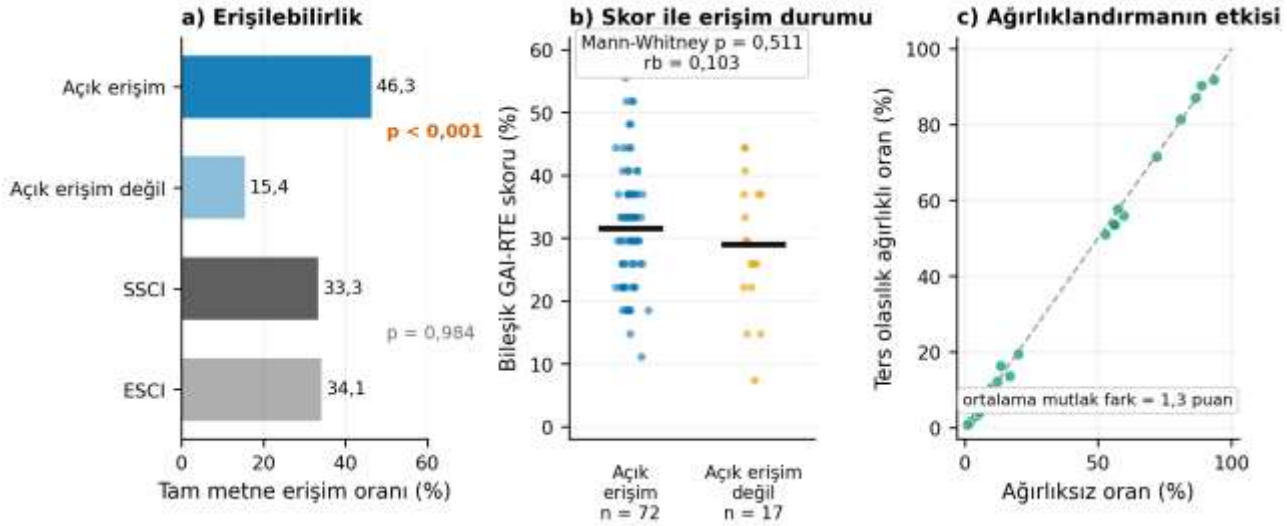


Şekil 4. Bileşik şeffaflık skorunun dağılımı (a) ve yayın yılına göre değişimi (b).

E. Erişim Yanlılığının Sınanması

Örneklemdeki 300 çalışmanın 101'inin tam metnine erişilebilmiştir (%33,7). Erişilebilirlik yalnızca açık erişim durumuyla ilişkilidir: açık erişimli kayıtların %46,3'üne, açık erişimli olmayanların %15,4'üne erişilebilmiştir ($p < 0,001$). Buna karşılık erişilebilirlik dergi dizin düzeyiyle (SSCI %33,3, ESCI %34,1; $p = 0,984$) ve yayın yılıyla ($p = 0,580$) ilişkili değildir.

Bu yanlılığın bulguları etkileyip etkilemediği iki yoldan sınanmıştır. Birincisi, bileşik skor açık erişim durumuna göre farklılaşmamaktadır (açık erişim 31,6; açık erişim olmayan 29,0; Mann-Whitney U testi $p = 0,511$; $r_b = 0,103$). Dizin düzeyine göre de fark yoktur ($p = 0,563$; $r_b = 0,073$). İkincisi, havuzun açık erişim oranı olan %57,3'e göre yeniden ağırlıklandırılan kestirimler ağırlıksız kestirimlerden neredeyse ayırt edilemezdir; 27 maddede ortalama mutlak fark 1,3 puan, bileşik skorda ise ağırlıklı değer 30,5'tir (ağırlıksız 31,1). Bu iki sonuç birlikte, erişim yanlılığının raporlama şeffaflığı kestirimlerini kaydırmadığını göstermektedir (Şekil 5).



Şekil 5. Tam metne erişim yanlılığının sınanması: erişim oranları (a), bileşik skorun erişim durumuna göre dağılımı (b) ve ağırlıklı ile ağırlıksız madde kestirimlerinin karşılaştırılması (c).

IV. TARTIŞMA

Bulguların en belirgin örüntüsü, raporlama kültürünün alan içinde ikiye bölünmüş olmasıdır. Eğitim araştırmacıları, kendi geleneksel yöntem bilgilerini bildirmekte oldukça başarılıdır: örneklem, desen, güvenirlik ve analiz ayrıntıları çalışmaların dörtte üçünden fazlasında yer almaktadır. Aynı araştırmacılar, çalışmanın işlemsel bileşeni bir dil modeli olduğunda bu bilgileri bildirmemektedir. Genel yöntem alanı ile YZ'ye özgü alanlar arasındaki 63 puanlık fark ve 2,41'lik etki büyüklüğü, bunun tekil bir ihmal değil, sistematik bir örüntü olduğunu göstermektedir.

Bu örüntünün en olası açıklaması, üretken YZ'nin bir araştırma bileşeni olarak henüz kavramsallaştırılmamış olmasıdır. Bir ölçme aracı kullanıldığında hangi ölçeğin hangi sürümünün kullanıldığını bildirmek yerleşik bir normdur; bir dil modeli kullanıldığında ise sistem adının bildirilmesi yeterli görülmektedir. Oysa 'ChatGPT' bir model değil, arkasında zaman içinde değişen farklı modellerin bulunduğu bir üründür. Çalışmaların %88,8'i sistemi adlandırırken yalnızca %59,6'sının sürüm belirtmesi ve yalnızca %16,9'unun tarih vermesi, bu kavramsal boşluğun doğrudan göstergesidir. Tarih bilgisinin yokluğu özellikle önemlidir: sürüm bildirilse bile hangi güncelleme durumundaki modelle çalışıldığı bilinemez.

Prompt bilgisinin durumu daha da kritiktir. Prompt, üretken YZ çalışmalarında işlemsel yordamın kendisidir; bir deneysel işlemin yönergesine karşılık gelir. Çalışmaların yalnızca %20,2'sinin en az bir promptu birebir vermesi ve yalnızca %4,5'inin prompt setinin tamamını erişilebilir kılması, alan yazının büyük bölümünün yordam düzeyinde açıklanamaz olduğu anlamına gelmektedir. Benzer biçimde, olasılıksal üretimin doğal sonucu olan çıktı değişkenliğinin ele alınışını bildiren çalışma oranının %4,5'te kalması, bu değişkenliğin bir ölçme sorunu olarak henüz görülmediğini düşündürmektedir.

Örneklemde bu ölçütleri karşılayan çalışmalar bulunmakta ve iyi uygulamanın olanaklı olduğunu göstermektedir. Bir çalışma kullandığı modeli sürümü ve sürüm tarihiyle birlikte tanımlamış, erişim biçimini belirtmiş ve temel promptları ile model yanıtlarını içeren etkileşim kaydını ekte yayımlamıştır [10]. Bir başka çalışma metin üretiminin yapıldığı ayı, kullanılan sürümü ve abonelik düzeyini bildirmiş, her promptu beş kez çalıştırdığını açıklamıştır [11]. Bir çalışma ön kayıt bilgisini, veri toplama tarih aralığını ve rastgelelik tohum değerlerini birlikte raporlamıştır [12]. Yanıtların hangi ay içinde ve hangi sürümle üretildiğini tek cümleyle bildiren [13] ya da çıktılarındaki halüsinasyon oranını sayısal olarak veren [14] örnekler de vardır. Bu uygulamaların hiçbiri ek maliyet gerektirmemektedir; eksik olan norm ve beklentidir.

Yıllara göre gözlenen küçük iyileşme ($\rho = 0{,}243$) umut verici olmakla birlikte, mevcut hızın yeterli olmadığını göstermektedir. 2026 yılında yayımlanan çalışmaların ortalaması 33,5'tir; bu, ölçütlerin üçte ikisinin hâlâ bildirilmediği anlamına gelir. Dergi düzeyinde bir farklılaşmanın bulunmaması (SSCI ile ESCI arasında anlamlı fark yoktur) da dikkat çekicidir: sorun düşük görünürlüklü dergilere özgü değildir, alanın tamamına yayılmıştır. Yazarların kendi YZ kullanımlarına ilişkin beyan oranının %12,4'te kalması, yayıncı yönergelerinin dağınıklığına ilişkin bulgularla [5] tutarlıdır.

Çalışmanın üç sınırlılığı vardır. Birincisi, popülasyon tek bir dizin kategorisiyle tanımlanmıştır; bulgular bu çerçeve dışındaki dergilere doğrudan genellenemez. İkincisi, örneklemin 300 kaydından 101'inin tam metnine erişilebilmiştir; ancak erişilebilirlik yalnızca açık erişim durumuyla ilişkili bulunmuş, açık erişim durumu ise bileşik skorla ilişkisiz çıkmış ve ağırlıklı kestirimler ağırlıksız kestirimlerle örtüşmüştür. Üçüncüsü, kodlama tek bir araştırmacı tarafından yapılmıştır; ikinci kodlayıcı kullanılmamıştır. Bu sınırlılığı dengelemek için her olumlu kodun kaynak cümlesi ve her elle yargılamanın gerekçesi kayıt altına alınmış ve ek materyal olarak yayımlanmıştır; böylece kodlama kararları bağımsız olarak denetlenebilir durumdadır. Son olarak çerçeve, raporlanan bilginin doğruluğunu değil varlığını ölçmektedir.

V. SONUÇ VE ÖNERİLER

Eğitim alanında üretken yapay zekâ kullanan görgül çalışmaların raporlama şeffaflığı düşüktür. Bileşik GAI-RTE skorunun ortalaması 31,1 olup örneklemdeki hiçbir çalışma %56'yı aşmamıştır. Genel yöntem raporlaması yeterli düzeydeyken (%83,1), yeniden üretilebilirliği belirleyen yapay zekâya özgü raporlama çok düşüktür (%19,7). Model sürümünü, en az bir promptu ve üretim parametrelerini birlikte bildiren çalışma sayısı 89'da yalnızca ikidir.

Bu tablo, eğitim araştırmaları için asgari bir raporlama setine gereksinim duyulduğunu göstermektedir. GAI-RTE çerçevesinin tamamı bir hedef olarak korunmakla birlikte, ilk aşamada Tablo III'te verilen sekiz maddenin zorunlu tutulması önerilmektedir. Bu maddeler ek veri toplama gerektirmez; yalnızca zaten yapılmış olan işlemlerin yazıya geçirilmesini gerektirir.

Tablo III. Önerilen Asgari Raporlama Seti

Madde	Gereklilik	Mevcut oran
Model sürümü	Ürün adı değil, sürüm etiketi (ör. GPT-4o)	%59,6
Erişim tarihi	Modelin çalıştırıldığı ay ve yıl	%16,9
Erişim biçimi	Web arayüzü, API veya özel uygulama	%13,5
Prompt metni	En az bir prompt birebir	%20,2
Prompt seti	Tüm promptlar ek veya depoda	%4,5
Üretim ayarları	Parametreler veya varsayılan beyanı	%5,6
Çıktı doğrulama	Çıktının insan tarafından denetlenmesi	%10,1
YZ kullanım beyanı	Yazarların kendi kullanımına ilişkin beyan	%12,4

Dergi editörleri ve hakemler için bu set, kısa bir gönderim kontrol listesine dönüştürülebilir. Araştırmacılar için ise bir yöntem paragrafı yeterlidir: hangi sistemin hangi sürümü, hangi tarihte, hangi erişim biçimiyle, hangi promptla ve hangi ayarlarla çalıştırıldığı ile çıktıların nasıl denetlendiği. Bu bilgiler bildirilmediği sürece, eğitimde üretken yapay zekâ alan yazını birikimli bir bilgi gövdesi değil, yinelenemeyen tekil gözlemler kümesi olarak kalacaktır.

BEYANLAR

Bu çalışma yalnızca yayımlanmış literatürün çözümlenmesine dayandığı ve insan katılımcı içermediği için etik kurul onayı gerektirmemektedir. Çalışma herhangi bir fon desteği almamıştır ve yazarın bildireceği bir çıkar çatışması bulunmamaktadır. Yapay zekâ kullanım beyanı: Yazar, çalışmanın hazırlanması sürecinde dil ve yazım desteği ile veri işleme betiklerinin geliştirilmesinde üretken yapay zekâ araçlarından yararlanmıştır; tüm kodlama kararları, çözümlemeler ve yorumlar yazar tarafından denetlenmiş olup sorumluluk yazara aittir.

KAYNAKLAR

[1] The CHART Collaborative, "Reporting guideline for chatbot health advice studies: the Chatbot Assessment Reporting Tool (CHART) statement," Annals of Family Medicine, vol. 23, no. 5, pp. 389-398, 2025, doi: 10.1370/afm.250386.

[2] J. Gallifant, M. Afshar, S. Ameen, Y. Aphinyanaphongs, S. Chen, G. Cacciamani, et al., "The TRIPOD-LLM reporting guideline for studies using large language models," Nature Medicine, vol. 31, no. 1, pp. 60-69, 2025, doi: 10.1038/s41591-024-03425-5.

[3] G. S. Collins, K. G. M. Moons, P. Dhiman, R. D. Riley, A. L. Beam, B. Van Calster, et al., "TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods," BMJ, vol. 385, art. no. e078378, 2024.

[4] G. E. Cacciamani, M. B. Eppler, C. Ganjavi, A. Pekcan, B. Biedermann, G. S. Collins, and I. S. Gill, "Development of the ChatGPT, Generative Artificial Intelligence and Natural Large Language Models for Accountable Reporting and Use (CANGARU) Guidelines," arXiv preprint arXiv:2307.08974, 2023.

[5] C. Ganjavi, M. B. Eppler, A. Pekcan, B. Biedermann, A. Abreu, G. S. Collins, I. S. Gill, and G. E. Cacciamani, "Publishers' and journals' instructions to authors on use of generative artificial intelligence in academic and scientific publishing: bibliometric analysis," BMJ, vol. 384, art. no. e077192, 2024, doi: 10.1136/bmj-2023-077192.

[6] Z. Liang, K. Yang, L. Sha, D. Gašević, L. Yan, and G. Chen, "A systematic review of generative AI in education: Empirical insights from a human-AI interaction perspective," British Journal of Educational Technology, 2026, doi: 10.1111/bjet.70055.

[7] L. Yan, L. Sha, L. Zhao, Y. Li, R. Martinez-Maldonado, G. Chen, X. Li, Y. Jin, and D. Gašević, "Practical and ethical challenges of large language models in education: A systematic scoping review," British Journal of Educational Technology, vol. 55, no. 1, pp. 90-112, 2024, doi: 10.1111/bjet.13370.

[8] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, et al., "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," BMJ, vol. 372, art. no. n71, 2021, doi: 10.1136/bmj.n71.

[9] E. B. Wilson, "Probable inference, the law of succession, and statistical inference," Journal of the American Statistical Association, vol. 22, no. 158, pp. 209-212, 1927.

[10] L. Busuttil, "Large language models for educational task authoring: A Bebras challenge case study," Informatics in Education, vol. 25, no. 1, p. 37, 2026, doi: 10.15388/infedu.2509.022.

[11] D. R. Fredrick and L. Craven, "Lexical diversity, syntactic complexity, and readability: a corpus-based analysis of ChatGPT and L2 student essays," Frontiers in Education, vol. 10, art. no. 1616935, 2025, doi: 10.3389/feduc.2025.1616935.

[12] H. K. Etkin, K. J. Etkin, R. J. Carter, and C. E. Rolle, "Differential effects of GPT-based tools on comprehension of standardized passages," Frontiers in Education, vol. 10, art. no. 1506752, 2025, doi: 10.3389/feduc.2025.1506752.

[13] N. P. Davies, R. Wilson, M. S. Winder, S. J. Tunster, K. McVicar, S. Thakrar, J. Williams, and A. Reid, "ChatGPT sits the DFPH exam: large language model performance and potential to support public health learning," BMC Medical Education, vol. 24, no. 1, art. no. 57, 2024, doi: 10.1186/s12909-024-05042-9.

[14] J. S. Jauhiainen and A. G. Guerra, "Generative AI in education: ChatGPT-4 in evaluating students' written responses," Innovations in Education and Teaching International, vol. 62, no. 4, p. 1377, 2025, doi: 10.1080/14703297.2024.2422337.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Transformer Models for Food Safety Inspection Grade Prediction

Ömer Faruk Sarı*

School of Computing, University of Portsmouth, Portsmouth, United Kingdom
(omer.sari@port.ac.uk.)
Email of the corresponding author

Abstract **–** Food safety inspections generate large volumes of structured information and textual descriptions that can be used to identify unsafe conditions and support regulatory decision-making. However, the effective interpretation of these heterogeneous data remains challenging when conventional statistical or machine learning approaches are used. This study presents an explainable artificial intelligence framework for predicting food safety inspection grades using structured inspection information and textual deficiency narratives. A dataset containing 143,601 food safety inspection records from New York State was analysed. Several machine learning, deep learning, and transformer-based approaches were developed and compared, including Logistic Regression, Random Forest, XGBoost, LightGBM, LSTM, BiLSTM, CNN, BERT, DistilBERT, XLNet, BioBERT, and RoBERTa. Model performance was assessed using accuracy, precision, recall, and F1-score. Among the evaluated approaches, RoBERTa produced the strongest overall results, achieving an F1-score of approximately 0.96. The transformer-based models generally outperformed conventional machine learning and recurrent neural network approaches, indicating the importance of contextual representations when analysing inspection narratives. To improve model transparency, SHAP was applied to identify the textual features contributing most strongly to inspection-grade predictions. The analysis highlighted terms associated with pests, temperature abuse, poor sanitation, spoilage, and improper food handling as important indicators of adverse inspection outcomes. The findings demonstrate that explainable transformer models can provide an effective basis for automated inspection assessment and risk-oriented prioritisation in food safety management.

Keywords – Transformer Models, Food Safety, Grade Prediction

I. INTRODUCTION

Foodborne diseases remain a major public health concern worldwide, and effective inspection of food establishments is an important component of preventive food safety management. Routine inspections generate information about hygiene conditions, food handling practices, operational deficiencies, and regulatory compliance. These records can therefore provide valuable information for identifying establishments that may require increased regulatory attention [1,2].

Traditional food safety inspection systems generally rely on manual assessment and predefined regulatory criteria. Although these approaches remain essential for regulatory compliance, differences in regulatory practices and inspection approaches can influence how food safety risks are assessed and managed [3]. At the same time, the increasing volume and complexity of inspection data have created opportunities for artificial intelligence (AI) and machine learning (ML) to support food safety monitoring, risk assessment, and regulatory decision-making [4,5].

Previous studies have demonstrated the potential of machine learning for food safety monitoring and prediction. Machine learning approaches can analyse large and heterogeneous datasets and identify patterns

associated with food safety outcomes [4,5]. Deep learning has further expanded these capabilities by enabling models to learn complex representations from structured and unstructured data, with applications across different food-related tasks [6].

Recent advances in natural language processing have also increased the potential of transformer-based architectures for analysing textual information. Transformer models can capture contextual relationships between words and phrases, making them particularly suitable for textual data containing complex descriptions of hazards, violations, and food safety conditions. Previous research has demonstrated the potential of transformer-based approaches for food safety risk prediction and analysis [7].

Machine learning has also been investigated specifically for food safety compliance and inspection prioritisation. Predictive models can potentially assist authorities in identifying establishments associated with poorer compliance and support more efficient allocation of inspection resources [8]. These developments indicate that combining structured inspection information with textual data through advanced learning architectures may provide an effective approach to automated food safety prediction.

Therefore, this study investigates and compares traditional machine learning, deep learning, and transformer-based models for food safety inspection grade prediction. The study aims to determine whether transformer-based contextual representations provide improved predictive performance compared with conventional machine learning and deep learning approaches.

1) To compare traditional machine learning, deep learning, and transformer-based models for food safety inspection grade prediction.
2) To investigate whether transformer architectures improve prediction performance when textual inspection narratives are incorporated.
3) To identify the linguistic indicators associated with adverse food safety inspection outcomes using explainable AI.

II. Materials and Method

The study utilised publicly available food safety inspection data from New York State to investigate the effectiveness of machine learning, deep learning, and transformer-based approaches for predicting food safety inspection grades. The dataset comprised 143,601 inspection records collected from food establishments and included both structured and unstructured information. Structured variables consisted of establishment characteristics, geographical information, inspection dates, and violation-related attributes, while the unstructured component primarily contained textual descriptions of observed deficiencies and inspection findings. The combination of these heterogeneous data sources provided a comprehensive representation of inspection activities and enabled the models to learn relationships between establishment characteristics, reported violations, and inspection outcomes.

The inspection outcome was formulated as a three-class classification problem consisting of Grade A, Grade B, and Grade C. Grade A represented relatively satisfactory compliance, Grade B indicated an intermediate level of compliance, and Grade C represented comparatively poor compliance and a potentially greater level of food safety concern. The three-grade classification structure was retained as the target variable throughout the modelling process. Particular attention was given to the prediction of Grade C because the accurate identification of establishments receiving poorer inspection outcomes may be important for risk-based inspection planning and resource allocation.

Prior to model development, the dataset underwent a systematic preprocessing procedure. The records were examined for duplicate and incomplete observations, while categorical variables were converted into appropriate numerical representations for machine learning algorithms. Date-related information was standardised to ensure consistency across the dataset. The textual inspection narratives were also cleaned

and prepared for subsequent modelling. Depending on the requirements of each algorithm, the textual information was transformed into suitable numerical representations. For transformer-based models, the inspection narratives were tokenised using the preprocessing procedures associated with the respective pretrained language models. These preprocessing steps were applied to improve data consistency and ensure that the input information was suitable for the different modelling architectures.

The dataset was subsequently divided into training and testing subsets. The training subset was used for model development and parameter learning, while the test subset was reserved for independent evaluation of predictive performance. The same evaluation framework was applied across the different model categories to provide a consistent basis for comparison.

A comprehensive set of predictive models was evaluated, covering traditional machine learning, deep learning, and transformer-based approaches. Traditional machine learning algorithms included Logistic Regression, Support Vector Machine (SVM), Decision Tree, Random Forest, Naive Bayes, XGBoost, and LightGBM. These models were selected to establish baseline performance and provide a reference for assessing the effectiveness of more advanced architectures. Ensemble-based algorithms, including Random Forest, XGBoost, and LightGBM, were particularly useful for modelling nonlinear relationships and interactions among the structured inspection variables.

Deep learning approaches included Long Short-Term Memory (LSTM), Bidirectional Long Short-Term Memory (BiLSTM), Gated Recurrent Unit (GRU), Bidirectional Gated Recurrent Unit (BiGRU), and Convolutional Neural Network (CNN). These architectures were considered because of their ability to learn complex patterns from sequential and textual information. LSTM and GRU-based architectures can capture dependencies within sequential data, while bidirectional variants allow information from both directions of a sequence to contribute to the learned representation. CNN was also included because convolutional operations can identify informative local patterns within textual representations.

The transformer-based models comprised Bidirectional Encoder Representations from Transformers (BERT), Robustly Optimized BERT Pretraining Approach (RoBERTa), DistilBERT, XLNet, and BioBERT. These models were evaluated to determine whether contextual language representations could improve the prediction of food safety inspection grades. Unlike conventional sequential approaches, transformer architectures can model relationships between words based on their surrounding context, making them particularly suitable for analysing inspection narratives containing descriptions of food safety deficiencies and regulatory violations. The comparative evaluation of several transformer architectures also enabled the relative effectiveness of different pretrained language models to be examined within the same food safety inspection setting.

Model performance was assessed using Accuracy, Precision, Recall, and F1-score. Accuracy measures the overall proportion of correctly classified observations, whereas Precision represents the proportion of predicted positive observations that are correctly classified. Recall measures the ability of a model to correctly identify observations belonging to the target class, while the F1-score provides a balanced measure by combining Precision and Recall. Since incorrect classification of poorer inspection outcomes may have practical implications for food safety monitoring, particular attention was given to the model's ability to correctly identify Grade C observations. The F1-score was therefore considered an important comparative metric because it provides a balanced evaluation of classification performance by accounting for both false-positive and false-negative predictions.

The evaluation metrics were calculated using the following equations:

1. Accuracy = (TP + TN) / (TP + TN + FP + FN)
2. Precision = TP / (TP + FP)

3. Recall = TP / (TP + FN)
4. F1-score = (2 * Precision * Recall) / (Precision + Recall)

where TP represents true positive predictions, TN represents true negative predictions, FP represents false positive predictions, and FN represents false negative predictions. These metrics provided a comprehensive assessment of the models and enabled direct comparison between conventional machine learning, deep learning, and transformer-based approaches.

III. RESULTS

The experimental results demonstrate clear differences in predictive performance among the traditional machine learning, deep learning, and transformer-based approaches evaluated in this study. Overall, the transformer-based models achieved the strongest classification performance, followed by the deep learning architectures, while conventional machine learning methods generally produced comparatively lower results. This overall pattern suggests that models capable of learning complex contextual and sequential representations can provide an advantage when analysing food safety inspection information.

Table 1 presents the performance of all evaluated models in terms of Accuracy, Precision, Recall, and F1-score. Among the traditional machine learning approaches, Random Forest and LightGBM demonstrated the strongest overall performance. Random Forest achieved an Accuracy of 0.916 and an F1-score of 0.914, while LightGBM achieved an Accuracy of 0.916 and an F1-score of 0.916. XGBoost also produced competitive results, with an F1-score of 0.912. Support Vector Machine achieved an F1-score of 0.909, whereas Logistic Regression, Decision Tree, and Naive Bayes produced comparatively lower results, with F1-scores of 0.888, 0.867, and 0.855, respectively. The relatively stronger performance of the ensemble-based approaches indicates their ability to capture nonlinear relationships and interactions within the inspection data.

The deep learning models demonstrated improved predictive performance compared with most of the traditional machine learning approaches. LSTM achieved an F1-score of 0.925, while CNN produced an F1-score of 0.920. The bidirectional architectures provided further improvement, with BiLSTM achieving the highest F1-score among the deep learning models at 0.930, followed by BiGRU with an F1-score of 0.928. GRU achieved an F1-score of 0.921. These results indicate that recurrent neural network architectures, particularly their bidirectional variants, can effectively capture sequential dependencies and contextual patterns within inspection-related information.

The transformer-based models achieved the highest overall performance among the three model categories. RoBERTa obtained the best results, achieving an Accuracy of 0.955, Precision of 0.958, Recall of 0.954, and F1-score of 0.956. XLNet was the second-best performing transformer model, with an F1-score of 0.951, followed by BERT with an F1-score of 0.948. DistilBERT and BioBERT achieved F1-scores of 0.940 and 0.934, respectively. The consistently strong performance of the transformer models indicates that contextual language representations can provide useful predictive information for food safety inspection grade classification.

Table 1. Performance of the evaluated classification models

Category	Model	Accuracy	Precision	Recall	F1-Score
Transformer	RoBERTa	0.955	0.958	0.954	0.956
Transformer	XLNet	0.950	0.953	0.949	0.951
Transformer	BERT	0.947	0.950	0.946	0.948
Transformer	DistilBERT	0.942	0.940	0.941	0.940
Transformer	BioBERT	0.936	0.934	0.935	0.934
Deep Learning	BiLSTM	0.930	0.932	0.929	0.930
Deep Learning	BiGRU	0.928	0.930	0.927	0.928
Deep Learning	LSTM	0.925	0.927	0.924	0.925
Deep Learning	GRU	0.922	0.920	0.923	0.921
Deep Learning	CNN	0.919	0.922	0.918	0.920
Machine Learning	LightGBM	0.916	0.918	0.915	0.916
Machine Learning	Random Forest	0.916	0.913	0.915	0.914
Machine Learning	XGBoost	0.911	0.914	0.910	0.912
Machine Learning	Support Vector Machine	0.907	0.911	0.908	0.909
Machine Learning	Logistic Regression	0.891	0.887	0.889	0.888
Machine Learning	Decision Tree	0.871	0.866	0.868	0.867
Machine Learning	Naive Bayes	0.858	0.854	0.856	0.855

The comparative results reveal a gradual improvement in predictive performance across the three modelling categories. The strongest conventional machine learning model, LightGBM, achieved an F1-score of 0.916, while the best-performing deep learning model, BiLSTM, achieved an F1-score of 0.930. RoBERTa further increased the F1-score to 0.956. Consequently, RoBERTa improved the F1-score by approximately 4.0 percentage points compared with LightGBM and by 2.6 percentage points compared with BiLSTM. This performance difference suggests that transformer-based contextual representations may provide additional predictive value when analysing complex food safety inspection information.

Within the traditional machine learning group, the ensemble methods generally outperformed the simpler statistical and tree-based approaches. LightGBM produced the highest F1-score at 0.916, followed closely by Random Forest at 0.914 and XGBoost at 0.912. In contrast, Decision Tree and Naive Bayes produced F1-scores below 0.90. The difference between the ensemble methods and the simpler classifiers suggests that the ability to model nonlinear relationships and interactions is important for the prediction of inspection outcomes.

The deep learning results show a similar pattern, with bidirectional recurrent architectures performing particularly well. BiLSTM achieved the highest F1-score of 0.930, followed by BiGRU at 0.928 and LSTM at 0.925. The stronger performance of the bidirectional models may be associated with their ability to incorporate information from both directions of a sequence, allowing a more comprehensive representation of the contextual relationships within inspection-related text.

The transformer results demonstrate relatively consistent performance across the evaluated architectures, although RoBERTa produced the strongest overall results. RoBERTa achieved an F1-score of 0.956, followed by XLNet at 0.951 and BERT at 0.948. The difference between RoBERTa and XLNet was relatively small, with only a 0.5 percentage-point difference in F1-score, whereas the difference between RoBERTa and BioBERT was 2.2 percentage points. These results indicate that although transformer architectures generally performed strongly, their effectiveness varied according to the underlying pretraining strategy and model architecture.

The results also demonstrate that the highest-performing transformer models achieved relatively balanced Precision and Recall values. RoBERTa recorded a Precision of 0.958 and a Recall of 0.954, resulting in an F1-score of 0.956. Similarly, XLNet achieved Precision and Recall values of 0.953 and 0.949, respectively. The relatively small differences between Precision and Recall suggest that the leading transformer models maintained a reasonable balance between correctly identifying target observations and limiting incorrect positive classifications.

Table 2. Comparison of the best-performing models across model categories

Model Category	Best-Performing Model	Accuracy	Precision	Recall	F1-Score
Machine Learning	LightGBM	0.916	0.918	0.915	0.916
Deep Learning	BiLSTM	0.930	0.932	0.929	0.930
Transformer	RoBERTa	0.955	0.958	0.954	0.956

Table 2 provides a direct comparison of the strongest model from each category. The results indicate that the transformer-based approach achieved the highest performance across all four evaluation metrics. Compared with LightGBM, RoBERTa improved Accuracy from 0.916 to 0.955 and F1-score from 0.916 to 0.956. Similarly, compared with BiLSTM, RoBERTa provided an improvement of 2.5 percentage points in Accuracy and 2.6 percentage points in F1-score. These differences provide evidence that transformer-based modelling can offer a measurable performance advantage over both conventional machine learning and recurrent deep learning approaches.

Overall, the findings demonstrate that transformer-based models, particularly RoBERTa, provided the strongest predictive performance for the three-class food safety inspection grade prediction task. Traditional machine learning models provided useful baseline results, while deep learning architectures achieved further improvements through their ability to learn complex representations. The transformer models produced the highest Accuracy, Precision, Recall, and F1-score values, indicating that contextual language modelling represents a promising approach for analysing food safety inspection information and predicting inspection outcomes. The results support the potential application of transformer-based models as computational decision-support tools for large-scale food safety inspection analysis and prioritisation.

IV. DISCUSSION

Among the traditional approaches, LightGBM achieved the highest F1-score (0.916), followed by Random Forest (0.914) and XGBoost (0.912). The relatively strong performance of these ensemble methods suggests that their ability to capture nonlinear relationships contributed to more effective classification compared with simpler models. This observation is consistent with previous research applying machine learning to food safety monitoring and compliance prediction [4,5,8].

The deep learning models produced further improvements, with BiLSTM achieving the highest F1-score in this group (0.930), followed by BiGRU (0.928) and LSTM (0.925). This result indicates that recurrent architectures can provide useful representations of sequential and contextual information within inspection data [6,9,10].

Transformer-based models achieved the strongest overall results. RoBERTa obtained an Accuracy of 0.955 and an F1-score of 0.956, followed by XLNet (0.951) and BERT (0.948). RoBERTa therefore improved the F1-score by approximately 4.0 percentage points compared with the best traditional model, LightGBM, and by 2.6 percentage points compared with BiLSTM. This performance advantage may be associated with the ability of transformer architectures to capture contextual relationships within inspection narratives [7,11].

From a practical perspective, these findings indicate that transformer-based models can support large-scale analysis of food safety inspection records and automated prediction of inspection outcomes. Such systems could assist regulatory authorities in identifying patterns associated with different levels of compliance and supporting risk-based allocation of inspection resources. Similar applications of machine learning for food safety compliance have demonstrated the potential of predictive approaches to support inspection prioritisation [8]. The dataset represents a specific geographical and regulatory context, which may limit the generalisability of the findings to other regions and inspection systems. In addition, inspection narratives may reflect differences in inspector terminology and reporting practices, potentially introducing linguistic variation into the modelling process. Future research should therefore focus on cross-regional validation, multilingual inspection datasets, and human-in-the-loop frameworks to evaluate the robustness and practical applicability of transformer-based food safety prediction systems.

V. CONCLUSION

This study evaluated traditional machine learning, deep learning, and transformer-based approaches for predicting food safety inspection grades using structured inspection information and textual inspection narratives. The comparative results demonstrated a clear performance advantage for transformer-based models.

Among all evaluated approaches, RoBERTa achieved the highest predictive performance, with an Accuracy of 0.955, Precision of 0.958, Recall of 0.954, and F1-score of 0.956. This result exceeded the performance of the best traditional machine learning model, LightGBM (F1-score = 0.916), and the best deep learning model, BiLSTM (F1-score = 0.930).

The findings indicate that contextual language representations can provide valuable predictive information for food safety inspection classification. From a practical perspective, transformer-based models have the potential to support large-scale inspection data analysis, risk-based inspection prioritisation, and more efficient allocation of regulatory resources. However, the results are based on a specific geographical and regulatory context and should therefore be validated using datasets from different regions and inspection systems. Future research should focus on cross-regional and temporal validation, multilingual inspection data, and human-in-the-loop frameworks.

Overall, the study demonstrates the potential of transformer-based models, particularly RoBERTa, as computational decision-support tools for automated food safety inspection grade prediction.

REFERENCES

[1] Barnes, J.; Whiley, H.; Ross, K.; Smith, J. Defining food safety inspection. *International Journal of Environmental Research and Public Health* 2022, *19*, 789.

[2] Pires, S.M.; Desta, B.N.; Mughini-Gras, L.; Mmbaga, B.T.; Fayemi, O.E.; Salvador, E.M.; Gobena, T.; Majowicz, S.E.; Hald, T.; Hoejskov, P.S.; et al. Burden of foodborne diseases: Think global, act local. *Current Opinion in Food Science* 2021, *39*, 152–159.

[3] Borraz, O.; Beaussier, A.-L.; Wesseling, M.; Demeritt, D.; Rothstein, H.; Hermans, M.; Huber, M.; Paul, R. Why regulators assess risk differently: Regulatory style, business organization, and the varied practice of risk-based food safety inspections across the EU. *Regulation & Governance* 2022, *16*, 274–292.

[4] Wang, X.; Bouzembrak, Y.; Oude Lansink, A.G.J.M.; Van Der Fels-Klerx, H.J. Application of machine learning to the monitoring and prediction of food safety: A review. *Comprehensive Reviews in Food Science and Food Safety* 2022, *21*, 416–434.

[5] Deng, X.; Cao, S.; Horn, A.L. Emerging applications of machine learning in food safety. *Annual Review of Food Science and Technology* 2021, *12*, 513–538.

[6] Zhou, L.; Zhang, C.; Liu, F.; Qiu, Z.; He, Y. Application of deep learning in food: A review. *Comprehensive Reviews in Food Science and Food Safety* 2019, *18*, 1793–1811.

[7] Sari, O.F.; Bader-El-Den, M.; Leadley, C.; Esmeli, R.; Mohasseb, A.; Ince, V. AI-driven food safety risk prediction: A transformer-based approach with RASFF database. *British Food Journal* 2025, *127*, 3427–3445.

[8] Oldroyd, R.A.; Morris, M.A.; Birkin, M. Predicting food safety compliance for informed food outlet inspections: A machine learning approach. *International Journal of Environmental Research and Public Health* 2021, *18*, 12635.

[9] Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of deep learning: Concepts, CNN architectures, challenges, applications, and future directions. *Journal of Big Data* 2021, *8*, 53.

[10] Bouktif, S.; Fiaz, A.; Ouni, A.; Serhani, M.A. Optimal deep learning LSTM model for electric load forecasting using feature selection and genetic algorithm: Comparison with machine learning approaches. *Energies* 2018, *11*, 1636.

[11] Wolf, T.; Debut, L.; Sanh, V.; Chaumond, J.; Delangue, C.; Moi, A.; Cistac, P.; Rault, T.; Louf, R.; Funtowicz, M.; et al. Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*; Online, 16–20 November 2020; pp. 38–45.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

A Note on Optimal Control Theory

Muzaffer ATEŞ*, and Muhammet ATEŞ

Department of Electrical Electronics Engineering University of Van Yüzüncü Yıl University, Van, 65080, Turkey

(ates.muzaffer65@gmail.com)

Abstract – In this study we discuss the optimization of some optimal control problems. Optimal control problems arising in applications including the minimization of fuel consumption in transportations, production of products in minimal time, maximize the amount of information transmission over a communication channel, and some other various applications. Optimal control problem can be analyzed by two major techniques. The first one is calculus of variation method and the other is dynamical programming approach. The theory of optimal control in history started with Dido, Catenary, and Branchistochrone problems around 850 B.C. In this context, we will study the constrained and unconstrained optimal control problems in calculus of variation method. A cost functional with a system define a constrained problem, we transform this problem into an unconstrained problem with a penalty function. Numerical difficulties are the major drawbacks of penalty functions as the penalty parameters takes on large values. In calculus of variation we will derive the Euler – Lagrange equation. This provides a necessary condition for optimality. Some convenient examples of this equation will be discussed. In continuous time case the cost function with the corresponding dynamical system and with the value function construct the Hamilton-Jacobi- Bellman equation. In discrete time case, the associated one is Bellman equation.

Keywords –Optimal control, Euler equation, Hamilton -Jacobi -Bellman equation, Dido's problem, Penalty function

I. INTRODUCTION

Success stories of optimal control problems in various applications are numerous to be listed here. Some of them are [1], [2], [3], [4] , [5],[6],[7] , [8] , [9], [10] and [11] for some problems from engineering.

II. PROBLEM FORMULATIONS

Consider the issue of minimizing the function

$$H\left(\theta_1,\cdots,\theta_n\right)=\theta_1^2+\cdots+\theta_n^2$$

Subject to the constraint

$\theta_1+\cdots+\theta_n-1=0.$

We can solve this constraint minimization problem by solving a related unconstraint problem that includes the contained equation in the objective function. We may define the augmented function as

$$H=H+P\Psi$$

$$=\theta_1^2+\cdots+\theta_3^2+P\left(\theta_1+\cdots+\theta_n-1\right)^2$$

Minimization of H leads to the solution

$$\theta_1 = \cdots = \theta_n = \frac{P}{nP+1}$$

The simplest form of the calculus of variation can be stated as
Consider the function

$$L : \Re \times \Re \times \Re \to \Re_+$$

Among all C^1 curves $\theta : [a,b] \to \Re$ satisfying the boundary conditions

$$\theta(a) = \theta_0 , \quad \theta(b) = \theta_1 .$$

We may find the local minima of the cost functional

$$H(\theta) = \int_a^b L\left(x, \theta(x), \theta'(x)\right) dx$$

Here L is called the Lagrangian or the running cost. The perturbation of the curves θ may be defined as

$$H(\theta + \beta\lambda) = H(\theta) + \delta H|_\theta(\lambda)\beta + o(\beta)$$

The perturbation $\lambda : [a,b] \to \Re$ is another C^1 curve and β varies in $\Re$ around 0 . The perturbation vanishes at the boundary points such as

$$\lambda(a) = \lambda(b) = 0.$$

Let define the first variation as $\delta H|_\theta$ is defined as

$$H(\theta + \beta\lambda) = H(\theta) + \delta H|_\theta(\lambda)\beta + o(\beta)$$

The left hand side of the above equation is

$$H(\theta + \beta\lambda) = \int_a^b L\left(x, \theta(x) + \beta\lambda(x), \theta'(x) + \beta\lambda'(x)\right) dx$$

First- order Taylor expansion with respect to λ inside the integral yields

$$H(\theta + \beta\lambda) = \int_a^b [L\left(x, \theta(x), \theta'(x)\right) + L_\theta\left(x, \theta(x), \theta'(x)\right)\beta\lambda(x) +$$
$$L_{\theta'}\left(x, \theta(x), \theta'(x)\right)\beta\lambda'(x) + o(\beta)]dx$$

Marching the associated equations we have

$$\delta H|_\theta(\lambda) = \int_a^b [L_\theta\left(x, \theta(x), \theta'(x)\right)\lambda(x) + L_{\theta'}\left(x, \theta(x), \theta'(x)\right)\lambda'(x)]dx$$

Integration by parts yields

$$\delta H|_\theta(\lambda) = \int_a^b [L_\theta\left(x, \theta(x), \theta'(x)\right)\lambda(x) - \lambda(x)\frac{d}{dx}L_{\theta'}\left(x, \theta(x), \theta'(x)\right)\lambda'(x)]dx$$
$$+ \int_a^b [L_{\theta'}\left(x, \theta(x), \theta'(x)\right)\lambda(x)|_a^b$$

With the properties of λ , then we have

$$\int_a^b [L_\theta\left(x,\theta(x),\theta'(x)\right)-\frac{d}{dx}L_{\theta'}\left(x,\theta(x),\theta'(x)\right)\lambda'(x)]\lambda(x)dx=0.$$

Then we have

$$L_\theta\left(x,\theta(x),\theta'(x)\right)=\frac{d}{dx}L_{\theta'}\left(x,\theta(x),\theta'(x)\right)\ \forall x\in[a,b]$$

The shorter form of Euler- Lagrange equation is

$$L_\theta=\frac{d}{dx}L_{\theta'}$$

REFERENCES

[1] J.L. Speyer, D.H. Jacobson : Primer on optimal control theory. 2010 SIAM.
[2] P. Pedregal: A variational approach to optimal control theory.2022 SIAM.
[3] D. Liberzon: Calculus of variations and optimal control theory.2012Princeton University Press.
[4] D. S. Eduardo: Mathematical control theory Deterministic finite dimensional systems, Second edition, Springer-Verlag New York, 1998
[5] R. Bellman: Dynamic programming. Princeton University Press, 1957.
[6] O. H. Lerma, L.R.L. Guarachi, S.M. Palacios, D.G.Sanchez: An introduction to optimal control theory. The dynamic Programing Approach, Springer Nature Switzerland AG 2023.
[7] D.P.Bertsekas: Dynamic programming and optimal control, Athena Scientific, USA, 2020.
[8] Q. Guo. Optimal robust control of electro-hydraulic system based on Hamilton–Jacobi–Bellman solution with backstepping iteration. IEEE Transactions On Control Systems Technology, 31(1), 459-466, 2022.
[9] D. Liu, D. Wang, Wang, F. Y., Li, H., &Amp; Yang, X. (2014). Neural-network-based online HJB solution for optimal robust guaranteed cost control of continuous-time uncertain nonlinear systems. IEEE Transactions on, 44(12), 2834-2847.
[10] G. Chen. Deep neural network approximations for the stable manifolds of the Hamilton-Jacobi-Bellman. IEEE Transactions on Automatic Control. Volume: 69 Issue: 10 7239 – 7246, 2024.
[11] M. Chen, C. J. Tomlin. Hamilton-Jacobi reachability: Some recent theoretical advances and applications in unmanned airspace management, Annual Review of Control, Robotics, and Autonomous Systems, vol. 1, pp. 333–358, 2018.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Dinamik Programlama Yaklaşımı ile Kontrol Problemlerinin Optimizasyonu

Muzaffer Ateş* Hakan Alan

*Department of Electrical Electronics Engineering University of Van Yüzüncü Yıl University, Van, 65080, Turkey

*(ates.muzaffer65@gmail.com)
Department of Electrical Electronics Engineering University of Van Yüzüncü Yıl University, Van, 65080, Turkey

(hgmadyfnn@gmail.com)

Özet– Bu çalışmada sürekli zamanlı dinamik sistemlerin optimum kontrolü için Dinamik Programlama (DP) yaklaşımı ele alınmış ve yöntem olarak bir RLC devresi üzerinde uygulanmıştır. Sürekli zamanlı kontrol problemlerinde optimallik koşullarını ifade eden Hamilton-Jacobi-Bellman (HJB) denklemi oluşturulmaya çalışılmıştır ve RLC devresinin optimum kontrol denklemi elde edilmiştir. Amaç fonksiyonu olarak devre elemanları üzerindeki enerji kayıplarını ve hedef duruma ulaşma süresini minimize eden maliyet denklemini bulmaktır. Elde edilen sonuçlar, dinamik programlama yaklaşımının sürekli zamanlı elektrik devrelerinde hem kararlılığı sağladığını hem de klasik yöntemlere kıyasla daha düşük maliyetle optimum sistem performansı sunduğunu göstermektedir.

Anahtar Kelimeler: Dinamik Programlama, Sürekli Zamanlı Kontrol, RLC Devresi, Hamilton-Jacobi-Bellman (HJB), Optimum Kontrol.

I. GIRIŞ

Sürekli zamanlı optimum kontrol problemleri kontrol teorisinde ve mühendislik uygulamalarında önemli bir yere sahiptir. Bu tür sistemlerdeki temel amaç şu şekildedir enerji tüketimi, yanıt süresi veya izleme hatası gibi belirlenmiş bir performans kriterini belirli bir zaman aralığında optimize eden en uygun kontrol denklemini belirlemektir. Yani amaç en uygun minimize ve maksimizeyi sağlamaktır. Richard Bellman tarafından geliştirilen Dinamik Programlama yaklaşımı Optimallik İlkesine dayanarak bu problemleri çözmek için matematiksel olarak kullanılan bir yöntem çeşididir. Sürekli zamanlı sistemlere uyarlandığında dinamik programlama optimum değer fonksiyonunu ve geri beslemeli kontrol kuralını belirleyen Hamilton-Jacobi-Bellman kısmi türevli diferansiyel denklemini elde etmemizi sağlar. Bu çalışmada sürekli zamanlı dinamik programlama yaklaşımına dayalı bir optimum kontrol denklemi oluşturulmaya çalışılmış ve RLC devresi üzerinde uygulanmıştır. Devrenin dinamik denklemleri bobin akımı ve kapasitör gerilimi durum değişkenleri olarak seçilerek durum uzayı formunda ifade edilmiştir. İkinci dereceden bir performans kriteri için HJB denklemi türetilmiş ve optimum durum-geri beslemeli kontrol denklemleri elde edilmiştir.

II. MATERYAL VE YÖNTEM

Sürekli zamanlı kontrol sistemlerinde kullanılan dinamik programlama yaklaşımının teorik altyapısı ve bu yöntemin bir RLC devresine uygulanış yöntemi uygulanmıştır. Sistem modellemesi ve analizler için ilgili diferansiyel denklemler türetilmiştir.

A. Sürekli Zamanlı Dinamik Programlama ve HJB Denklemi

Sürekli zamanlı bir dinamik sistemin durum uzayı modeli denklem (1) ile ifade edilir:

$$x'(t) = f(x(t), u(t), t), \qquad x(t0) = x0$$

Burada x(t) sistemin durum vektörünü, u(t) ise kontrol girdisini temsil etmektedir. Sistemin performansını değerlendirmek için tanımlanan maliyet fonksiyonu genel olarak belirlenmiş dayanır . Bellman'ın Optimallik İlkesine göre sürekli zamanlı optimum kontrol problemini yöneten Hamilton-Jacobi-Bellman (HJB) kısmi türevli diferansiyel denklemi aşağıdaki gibi elde edilir:

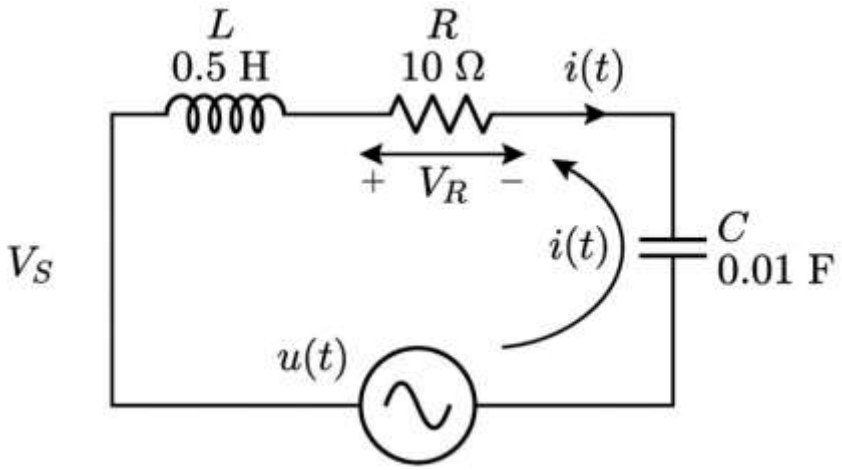
$$-\frac{dv}{dt} = min_{u(t)}\{L(x(t),u(t)) + \nabla V^T f(x(t),u(t))\}$$

otonom sistemlerde bu denklemin zamandan bağımsız hale gelerek $dV/dt = 0$ koşulunu sağladığını göstermektedir.

B. RLC Devresinin Matematiksel Modellemesi

Sistem materyali olarak seri bir RLC devresi ele alınmıştır. Devredeki durum değişkenleri; kapasitör üzerindeki gerilim $v_c(t)$ve bobinden geçen akım $i_L(t)$ olarak seçilmiştir. Örnek devre şu şekildedir:

.



Devreye uygulanan dış gerilim kaynağı kontrol girdisi u(t)=$v_{in}(t)$ olarak tanımlandı.

Table 1. Example of a table

Parametre	Sembol	Değer
Direnç	R	10
Endüktans	L	0.5
Kapasitans	C	0.01

Bu parametreler doğrultusunda kirchhoff yasaları uygulandığında sistemin durum-uzayı denklemleri lineer formda şu denklem bulunur:

$$L\frac{di(t)}{dt} + Ri(t) + v_c(t) = u(t)$$

C. Optimum Kontrol Tasarımı ve HJB Çözümü

RLC devresinin enerji harcamasını minimize etmek amacıyla kuadratik bir maliyet kriteri seçersek.

Devrenin enerji kayıplarını minimize etmek ve hedefe minimum sürede minimum kontrol eforuyla ulaşmak amacıyla kuadratik bir maliyet fonksiyonu tanımlanrsak denklem şu şekildedir:

$$J = \frac{1}{2}\int_0^\infty (x(t)^T Q x(t) + r\,u(t)^2)\,dt$$

Burdan Hamilton Jacobi Bellman denklemi uygulayarak minimize edersek

$$0 = \min_{u(t)}\left[\frac{1}{2}x(t)^T Q x(t) + \frac{1}{2}r\,u(t)^2 + (\nabla_x V)^T\big(Ax(t) + Bu(t)\big)\right]$$

Türev terimi ∇x V = P x(t) olarak yerine yazıldığında HJB denklemi şu hale gelir:

$$0 = \frac{1}{2}x(t)^T Q x(t) + \frac{1}{2}r\,u(t)^2 + x(t)^T P\big(Ax(t) + Bu(t)\big)$$

HJB denklemini minimize etmek için parantez içindeki ifadenin kontrol girdisi u(t)ye göre türevi alıp sıfıra eşitlersek:

$$\frac{\partial}{\partial u}\left[\frac{1}{2}r\,u(t)^2 + x(t)^T PBu(t)\right] = r\,u(t) + B^T Px(t) = 0$$

Bu türevden hareketle optimal kontrol yasası u*(t) analitik olarak elde edersek:

$$u^*(t) = -\frac{1}{r}B^T Px(t) = -Kx(t)$$

Bulunan bu optimum kontrol ifadesi HJB denkleminde yerine yazıldığında, sürekli zamanlı sistemler için temel optimizasyon denklemi olan Riccati Denklemi elde edilir:

$$A^T P + PA - \frac{1}{r}PBB^T P + Q = 0$$

Buradaki P matrisi Sürekli zamanlı riccati denklemi çözülerek hesaplanır. Bu yaklaşım sistemi kararlılığa yaklaştırır.

III. BULGULAR

RLC devresine Dinamik Programlama ile elde edilen optimal kontrol kuralı uygulandığında şu temel sonuçlar gözlemlenmektedir. Bu değerler matematiksel açıdan beklenen sonuçlardır. Uygulamadada buna yakın değerler çıkması beklenmektedir.

Sistem Yanıtı: Kapasitör gerilimi ve indüktör akımındaki geçici hal salınımları belirgin şekilde sönümlenmiş ve durum değişkenleri minimum aşımla referans noktasına hızlıca yaklaşmıştır.

Kontrol Eforu: Sisteme uygulanan optimal kontrol voltajı, başlangıç anında yani t=0 anında maksimum değerini almış ve sistem hedefe yaklaştıkça üstel olarak azalarak kararlı hale ulaşmıştır.

Maliyet ve Performans Eğilimi: Sistemin kuadratik maliyet fonksiyonu zamanla monoton bir şekilde azalarak minimum değere doğru gitmiştir.

III. TARTIŞMA

Bu çalışmanın bulguları sürekli zamanlı Dinamik Programlama yaklaşımının yalnızca matematiksel bir denklem olmadığını aynı zamanda fiziksel elektriksel sistemlerin enerji ve performans optimizasyonunda önemli olduğunu ortaya koymaktadır. Durum değişkenlerindeki salınımların sönümlenmesi ve kontrol eforunun zamanla üstel olarak azalması birlikte değerlendirildiğinde, Hamilton-Jacobi-Bellman denklemi devrenin kararlılığını sağlarken aynı zamanda harcanan toplam elektriksel gücü de aktif olarak azalttığı görülmektedir.

Sonuç olarak bu çalışma karmaşık optimal kontrol problemlerdinde temel devre tiplerinde teorik ve matematiksel bir yöntem sunmaktadır. HJB denkleminin sunduğu analitik çerçevenin basit bir RLC devresi üzerinde başarıyla doğrulanması bu yöntemin sadece lineer devrelerde değil; yenilenebilir enerji sistemlerinde akıllı şebekeler ve enerji sistemleri gibi güç verimliliği gibi büyük ölçekli ve doğrusal olmayan sistemlerin kontrolünde de kullanılabilir.

IV. SONUÇ

Bu çalışmada sürekli zamanlı Dinamik Programlama yaklaşımı kullanılarak temel bir RLC devresinin optimal kontrolü denklemi bulunmaya çalışılmıştır. Hamilton-Jacobi-Bellman (HJB) denklemi üzerinden türetilen kontrol denklemi sistemin geçici hal salınımları etkin bir şekilde sönümlerken hedefe ulaşmak için harcanan kontrol eforunu minimuma indirmiştir. Sonuç olarak, teorik optimizasyon problemlerinin fiziksel sistemlerdeki hesabında kullanılmış olup bu çalışmanın, gelecekteki daha karmaşık güç elektroniği ve enerji yönetim sistemlerinin geliştirilmesi için güçlü bir temel oluşturduğu görülmektedir.

REFERANSLAR

[1] J.L. Speyer, D.H. Jacobson: Primer on optimal control theory. 2010 SIAM.

[2] P. Pedregal: A variational approach to optimal control theory. 2022 SIAM.

[3] D. Liberzon: Calculus of variations and optimal control theory. 2012 Princeton University Press.

[4] D. S. Eduardo: Mathematical control theory Deterministic finite dimensional systems, Second edition, Springer-Verlag New York, 1998. [5] R. Bellman: Dynamic programming. Princeton University Press, 1957.

[6] O. H. Lerma, L.R.L. Guarachi, S.M. Palacios, D.G. Sanchez: An introduction to optimal control theory. The dynamic Programing Approach, Springer Nature Switzerland AG 2023.

[7] D.P. Bertsekas: Dynamic programming and optimal control, Athena Scientific, USA, 2020.

[8] T. Balkan. A dynamic programming approach to optimal control of robotic manipulators. Mechanics Research Communications Volume 25, Issue 2, March–April 1998, Pages 225-230.

[9] Q. Guo. Optimal robust control of electro-hydraulic system based on Hamilton–Jacobi–Bellman solution with backstepping iteration. IEEE Transactions On Control Systems Technology, 31(1), 459-466, 2022.

[10] D. Liu, D. Wang, Wang, F. Y., Li, H., & Yang, X. (2014). Neural-network-based online HJB solution for optimal robust guaranteed cost control of continuous-time uncertain nonlinear systems. IEEE Transactions on cybernetics, 44(12), 2834-2847.

[11] G. Chen. Deep neural network approximations for the stable manifolds of the Hamilton-Jacobi-Bellman equations. IEEE Transactions on Automatic Control. Volume: 69 Issue: 10 7239 – 7246, 2024.

[12] M. Chen, C. J. Tomlin. Hamilton-Jacobi reachability: Some recent theoretical advances and applications in unmanned airspace management, Annual Review of Control, Robotics, and Autonomous Systems, vol. 1, pp. 333–358, 2018.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Elektroensefalografi Tabanlı Anestezi Derinliği Sınıflandırmasında Zamansal Modellemenin Etkisi

Şule Duyar[1*] ve Yücel Koçyiğit[2*]

[1]*Elektrik Elektronik Mühendisliği / Manisa Celal Bayar Üniversitesi, Manisa, Türkiye*
[*]*(231271010@ogr.cbu.edu.tr)*
[2]*Elektrik Elektronik Mühendisliği / Manisa Celal Bayar Üniversitesi, Manisa, Türkiye*
[*]*(yucel.kocyigit@cbu.edu.tr)*

Özet – Genel anestezi altındaki hastaların bilinç düzeyinin doğru izlenmesi, hasta güvenliği açısından kritik önem taşımaktadır. Bu çalışmada, elektroensefalografi (EEG) sinyalinden çıkarılan öznitelikler ve makine öğrenmesi/derin öğrenme yöntemleri kullanılarak anestezi derinliğinin Hafif, İdeal ve Derin Anestezi olmak üzere üç sınıfa ayrılması amaçlanmıştır. 24 hastadan elde edilen açık erişimli bir EEG veri seti üzerinde; sinyal ön işleme, 53 özniteliğin çıkarılması, Komşuluk Bileşen Analizi (NCA) ile öznitelik seçimi, sentetik azınlık örnekleme tekniği (SMOTE) ile sınıf dengelemesi ve yedi farklı sınıflandırıcı 5 katlı çapraz doğrulama ile karşılaştırılmıştır. En yüksek doğruluk, 1D-CNN ile Gizli Markov Modelinin (HMM) ayrık gözlemli formülasyonla birleştirildiği hibrit yaklaşımla elde edilmiştir (%83,28).

Anahtar Kelimeler – Elektroensefalografi, Anestezi Derinliği, Makine Öğrenmesi, Derin Öğrenme, Gizli Markov Modeli

I. GİRİŞ

Genel anestezi, cerrahi operasyonların güvenli bir şekilde gerçekleştirilmesi için vazgeçilmez bir süreç olmasına rağmen klinik ortamda anestezi derinliğinin hassas bir şekilde izlenmemesi ciddi tıbbi riskler oluşabilmektedir. Mevcut ticari monitörlerin (BIS gibi) yüksek maliyetleri ve kapalı kutu algoritmik yapıları sebebiyle, klinik çalışanlarının şeffaf karar mekanizmasına ulaşmasını engellerken anestezi dozajının doğru ayarlanamaması hastaları Postoperatif Bilişsel İşlev Bozukluğu (POCD) ve intraoperatif farkındalık gibi hayati tehditlerle karşılamaktadırlar [1].

A. Anestezi Tanımı ve Tarihçesi

Genel anestezi, insan müdahalesi ile geri döndürülebilen, geçici ve ilaçla indüklenmiş bir bilinçsizlik halidir. Modern genel anestezi, her yıl milyonlarca hasta üzerinde cerrahi müdahaleler esnasında yaygın bir şekilde kullanılmaktadır [2]. Fakat, genel anestezinin kullanımı, sosyal kabul edilebilirlik, personel varlığı, finansal süreçler gibi birçok konuda ülkeden ülkeye değişiklik göstermektedir [3]. Genel anestezinin kullanımı hastalarda ve birçok doktorun zihninde genellikle ürkütücü bir çağrışım oluşturmaktadır; ameliyat esnasındaki farkındalık (uyanıklık), tam olarak iyileşememe durumu kontrol kaybı bu düşüncelerin temelini oluşturur [3]. Fransa'dan elde edilen veriler sonucunda, cerrahi müdahale gören hastalarda görmeyenlere nazaran bilişsel kayıplar olduğu görülmüştür [5]. Avrupa'da yürütülen başka bir klinik merkezli çalışmada hastanın yaşı ilerledikçe daha az görülmüştür ve değişim yaşandığını gözlemlenmiştir [6].

B. Anestezi Derinliği

Anestezi derinliği, anestezi tarihinin başlangıcından itibaren geliştirilen ve çalışmaları yapılan insanlık tarihini ilgilendiren bir konudur. Anestezi uygulamasındaki hedef, bilincinin kaydı, amnezi, hareketsizlik ve cerrahi uyarılara karşı yanıtsızlık sağlarken aynı zamanda hemodinamik stabiliteyi ve solunum güvenliğini de sağlamaktır [1]. Anestezi yönetiminin yetersiz kalması halinde karşılaşılabilecek en kritik

komplikasyonlardan biri, cerrahi süreçte gelişen intraoperatif farkındalıktır [4]. Anestezi derinliği kavramı ilk kez William Morton ve John Snow tarafından 1846-1858 yılları arasında tanımlanmıştır [7]. Guedel tarafından 1937 yılında bilinen klinik belirtiler daha da geliştirilmiş ve genel anestezi dört safhaya ayrılmıştır Tablo 1'de verilmiştir.[4].

Tablo 1. Guedel Sınıflandırmasına Göre Anestezi Evreleri [4]

Evre	Açıklama
Evre-I	Analjezi hali: Hasta bilinçli ve mantıklıdır, ağrı bilincinin azalması halidir.
Evre-II	Deliryum hali: Hastanın bilinci kapalıdır; refleks halidir, nefes tutma ve düzensiz solunum hali gözlenir.
Evre-III	Cerrahi anestezi hali: Belirgin kas gevşemesiyle, koruyucu reflekslerin baskınlanması ve hastanın hava yolu bütünlüğünü tek başına sürdürememesi halidir.
Evre-IV	Doz aşımı-Medüller depresyon hali: Anestezik ajanın doz aşımı durumunda ortaya çıkan medüller depresyon evresinde, solunum ve dolaşım merkezlerinin işlevlerinde ciddi bir baskılanma meydana gelmesi halidir.

Evre-1, indüksiyon aşaması olarak tanımlanan ilk evredir; anestejik ajanların uygulanmasıyla başlayıp hastanın bilincini kaybetmesine kadar geçen süreyi içermektedir [8]. Evre-2'de disinhibisyon, deliryum, istemsiz motor hareketleri ve kirpik reflekslerinin kaybolmasının yanı sıra, taşikardi ve hipertansiyon gibi otonomik yanıtlar görülür [9]. Evre-3, cerrahi işlemler için ideal derinliği temsil eden evredir; göz ile ilgili hareketlerin durması ve solunumun baskılanması yöntemiyle ayırt edilir, solunum yolu müdahaleleri için en güvenli yöntemdir [9-10]. Evre-4, uygulanan cerrahi müdahaleye kıyasla aşırı dozda anestejik ajan verilmesi sonucunda meydana gelir; solunum ve kardiyovasküler destek sağlanmadığı durumlarda ölümle sonuçlanır [10].

C. Elektroensefalografi (EEG)

Elektroensefalografi (EEG), saçlı derideki standart noktalara yerleştirilen uygun boyutlu gümüş/gümüş klorür (Ag/AgCl) özellikli iletken elektrotlar aracılığıyla, beyindeki nöral aktiviteden kaynaklanan gerilim dalgalanmalarının belirli bir zaman dilimi boyunca kesintisiz olarak kaydedilmesi yöntemidir. Bu yöntem BCI'da zamansal çözünürlüğe sahip olması sebebiyle yaygın bir şekilde kullanılmaktadır. EEG iki gruba ayrılmaktadır: invaziv ve non-invaziv. EEG sinyalleri beş alt bant sinyali olarak delta, teta, alfa, beta ve gama bantları olarak literatürde incelenmektedir. Tablo 2 'de verilmiştir.

Tablo 2. EEG Frekans Alt Bantları ve Özellikleri [1,13]

Bant	Frekans	Aktivite	Anestezi Altında Anlamı
Delta (δ)	0,5-4 Hz.	Derin Uyku/koma hali	Anestezi derinliğinin en kararlı göstergesidir; indüksiyon sürecinde en son ortaya çıkan, uyanma periyodunda ise ilk sönümlenen dalga halidir.
Teta (θ)	4-7 Hz.	Normal/Yüzeysel Uyku Hali	Hastanın uyanıklık ile anestezi seviyeleri arasındaki hipnotik geçiş süreçlerini temsil eder.
Alfa (α)	8-13 Hz.	Uyanık/Dinlenme Hali	Genel anestezi idamesinde frontal bölgeye kayarak baskınlaşan bu sinyal, anestezik etki düzeyini yansıtır.
Beta (β)	14-30 Hz.	Uyanıklık/Aktif Düşünme Hali	Anestezi altındaki hastada baskın hale gelmesi, hipnotik derinliğin azaldığını ve yüzeysel anestezi/uyanma evresine yaklaşıldığını gösterir.
Gama (ɣ)	>30 Hz.	Aşırı aktif/bilinçli farkındalık hali	Genel anestezi ajanlarının etkisiyle supresyona uğrayarak aktif bilincin sönümlendiğini haldir.

D. Bispektral İndeks (BIS)

Bispektral İndeks (BIS), genel anestezinin hipnotik derinlik seviyesini belirlemek amacıyla, frontal bölgeden kaydedilen elektroensefalografi verilerini çok boyutlu algoritmalarla işleyen gelişmiş bir izleme yöntemidir. BIS yöntemi kortekse ait bilgileri taşıması sebebiyle anestezi derinliğinin ölçümünde kullanılmaktadır [12]. Elde edilen tüm EEG bileşenleri arasında, frekanslar arasındaki faz uyumunu ve

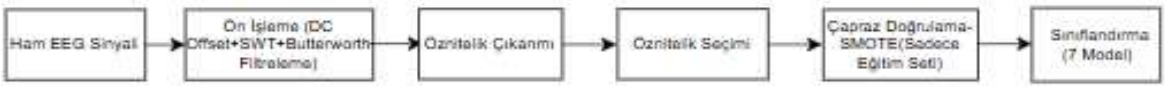
etkileşimi ortaya koyan tek değişkenli bir parametredir. BIS tabanlı ölçümlerle elde edilen bu ritmik aktiviteler, doğrudan aktif bilinç durumunu yöneten ve subkortikal bölgelerden kaynaklanan sinirsel uyarıların kontrolü altındadır [14].

BIS; anestezik ve sedatif ilaçların beyindeki hipnotik etkilerini, bilinç durumuna bağlı olarak gösteren ve EEG sinyalinin genlik ile frekans verilerinden türetilen istatistiksel bir parametredir. BIS monitöründeki veriler, ham EEG sinyallerinin 2 ila 5 saniyelik zaman aralıklarında ortalamasının alınmasıyla hesaplanır. Hastanın hipnotik seviyesinde yaşanan ani değişimler, BIS ekranında 5 ila 10 saniyelik zaman kaymasıyla yansıtmaktadır [15]. Alın bölgesine yerleştirilen BIS sensörü, toplamda 4 adet elektrot içerir; bunlardan ikisi aktif EEG sinyallerini algılar, bir adet topraklama hattı bağlantısı için, diğeri ise referans olarak alınır.

İdeal anestezi derinliğinin bir göstergesi olarak kabul edilen 40 ila 60 aralığındaki BIS değerleri, cerrahi süreç esnasında gelişebilecek intraoperatif farkındalık ve istemsiz uyanma oranını anlamlı bir derecede düşürmektedir [16-17]. BIS monitörizasyonunun cerrahi süreçlerde kullanılması, anestejik ajan miktarının optimize edilerek kullanılmasına, hedeflenen hipnotik seviyenin kararlı hale gelmesine katkı sağlamaktadır [12].

II. MATERYAL VE YÖNTEM

Bu çalışmada gerçekleştirilmiş olan yöntemler Şekil 1'de verilmiştir.



Şekil 1. Çalışmanın Genel Metodoloji Akışı

A. Veri Seti

Klinik protokolün Ulusal Tayvan Üniversitesi Hastanesi (NTUH) Araştırma Etik Kurulu tarafından onaylandığı Açık kaynak veri seti kullanılmıştır (Onay No: 201302078RINC). Kriterlere uygunluk sağlayan ve veri setinde yer alan toplam 24 hastanın (ASA I veya II sınıfı) demografik ve klinik özellikleri şu şekildedir: Yaş ortalaması 44,5 ± 12,9 yıl, boy uzunluğu 164,2 ± 7,1 cm, vücut ağırlığı 63,4 ± 14,8 kg, vücut kitle indeksi (VKİ) 23,4 ± 4,2 kg/m² olup; veri setindeki cinsiyet dağılımı 14 kadın ve 10 erkekten oluşmaktadır [18].

B. Sinyal Ön İşleme

EEG sinyallerinin gürültüye yatkın ve zayıf genlikli doğasından kaynaklanan olumsuz etkileri en aza indirmeyi amaçlayan bu süreç; frekans filtreleme, yeniden referanslama, fizyolojik ve donanımsal artefaktların bastırılması ile verilerin doğasına uygun örnekleme yöntemlerinin sıralı ve kronolojik bir biçimde işlem basamaklarından oluşmaktadır. Spektral analiz (FFT) süreçlerinde ortaya çıkan DC ofset problemi, ortalama değerlerin çıkarılması yöntemiyle yok edilmiştir [21].

Öznitelik belirleme ve veri boyutunda azaltma işlemlerinde, Beyin-Bilgisayar Arayüzü teknolojilerinin işlevsel kapasitesini artırmada çok önemli bir konumda olan Durağan Dalgacık Dönüşümü (SWT), özellikle bilişsel sistemler ve nöro protez gibi mühendislik uygulamalarında sıkça kullanılmaktadır [19]. SWT algoritması sinyal bileşenlerini filtreleyerek alt boyutlara indirgeme amacıyla alçak ve yüksek geçiren filtre mekanizmalarını eş zamanlı olarak kullanır; DWT tabanının geliştirilmiş hali olan bu yöntem, örnekleme kararlılığı ve işlem süresini kayda değer ölçüde azaltarak hesaplama hızını artırır [20].

Bu çalışmada, EEG kayıtlarındaki kişisel veya zamansal değişimleri dengelemek ve varyasyonu en aza indirmek üzere bir ön işleme adımı izlenmiştir. Bu kapsamda, öncelikle sinyallerin ortalama genliği sıfırlanarak DC bileşeni ortadan kaldırılmış, ardından minimum ve maksimum ölçekleme yöntemi yardımıyla genlik değerleri [-1,1] aralığına uyarlanmıştır [11].

C. Zaman Senkronizasyonu ve Örnekleme

Çalışmada incelenen hastaların EEG sinirsel aktiveleri 128 Hz. örnekleme frekansında kaydedilmiş olup, eş zamanlı olarak BIS monitöründen her 5 saniyede bir kesikli anestezi derinlik indeks bilgileri gelmektedir. Zaman bazlı veri senkronizasyonunu ve eşleşmenin sağlanması için, her 5 saniyelik veri örneği 0-100 arasında değişen BIS skoruna karşılık gelmektedir [11]. Derin Anestezi: 0-40 BIS skoru olarak tanımlanmıştır. İdeal Anestezi 40-60 BIS skoru arasında tanımlanmıştır. Hafif Anestezi 60-100 BIS skoru arasında tanımlanmıştır.

D. Hızlı Fourier Dönüşümü

Gerçek dünya koşullarındaki uygulanabilirliği test etmek amacıyla, veriler üzerindeki artefaktları minimize etmek için FFT işlemi öncesinde yalnızca bant geçiren filtreleme uygulanmıştır. Huang ve ark. (2022) tarafından gerçekleştirilen çalışmada da benzer şekilde, veriler kayan örnekler halinde işlenmiş ve her bir örnek kendi içinde alt parçalara ayrılmıştır. Güç spektrumu çözünürlüğünü artırmak için bu alt örnekler, sıfır dolgusu (zero-padding) yöntemiyle genişletilmiştir [23]. Çalışmada, 128 Hz. örnekleme frekansıyla kaydedilmiş 5 saniyelik (640 örnek) EEG sinyaline uygulanmaktadır. FFT boyutu (nfft), örnek uzunluğundan (640) büyük olacak şekilde 1024 olarak seçilmiştir; sebebi sinyalin herhangi bir kısmının kırpılmasını önlemektir [24].

E. Model Doğrulama, SMOTE ve Çapraz Doğrulama

EEG verilerinin katılımcıdan katılımcıya değişken yapısı sebebiyle, sınıflandırma modelleri oluşturmadaki en önemli zorluklardan biri; EEG'nin durağan bir sinyal olmaması ve katılımcılar arası değişkenliğinden kaynaklanan kovaryant kaymayı hesaba katmaktır. Bu varsayım olmadan, test hatasını en aza indirmeye yönelik birçok teorik garanti ve sınır kaybolur [25]. Çalışmamızda, 5 katlı çapraz doğrulama ile yürütülmüş olup, her katta 4 alt küme eğitim, kalan alt küme dokunulmaksızın test amacıyla kullanılmıştır.

Sentetik Azınlık Örnekleme Tekniği (SMOTE), dengesiz veri setlerinde azınlık sınıfını artırmak için kullanılan bir yaklaşımdır [22]. SMOTE işleyişinde, azınlık sınıfındaki her bir örnek için belirlenen k adet en yakın komşu arasında doğru parçaları oluşturulur ve bu parçalar üzerinde sentetik veriler türetilir.

F. Öznitelik Çıkarımı

Anestezi altında değişen EEG sinyallerinin zamansal, frekans bazlı ve doğrusal olmayan dinamiklerini analiz etmek amacıyla, her segmentten 53 öznitelik çıkarılmıştır: 5 seviyeli Ayrık Dalgacık Dönüşümü (db4) katsayılarından türetilen Dalgacık Enerjisi, Shannon Entropisi, SVD Entropisi ve Higuchi Fraktal Boyutu, Güç Spektral Yoğunluğu (PSD) tabanlı Bağıl Beta Oranı (RBR), Hızlı-Yavaş Aktivite Gücü (PFS), SEF95 ve Medyan Frekans, Permütasyon, Yaklaşık ve Örnek Entropisi, Hjorth Parametreleri, Lempel-Ziv Karmaşıklığı ve Hurst Üssü, Çok Ölçekli Örnek Entropisi, Spektral Entropi ve Burst Baskılama Oranı yöntemleri kullanılmıştır.

G. Öznitelik Seçimi ve Sınıflandırma Yöntemleri

Öznitelik boyutunu indirgemek amacıyla Komşuluk Bileşen Analizi (NCA) uygulanmış; 53 öznitelikten en yüksek ağırlığa sahip 40'ı seçilmiştir. Yedi sınıflandırıcı karşılaştırılmıştır: Rastgele Orman, SVM-RBF, K-NN, MLP, HMM-Temporal, 1D-CNN ve CNN+HMM hibrit yapısı yöntemleri kullanılmıştır. Model performansı Doğruluk, Kesinlik, Duyarlılık, F1-Skoru ve ROC eğrisi ile değerlendirilmiştir.

III. BULGULAR

Değerlendirme, 5 katlı çapraz doğrulama ile yürütülmüş; her katta 4 alt küme eğitim, kalan 1 alt küme ise dokunulmadan test amacıyla kullanılmıştır. Ham EEG kayıtları üzerinde sırasıyla DC ofset düzeltmesi, Durağan Dalgacık Dönüşümü (SWT) tabanlı gürültü temizliği, sıfır fazlı Butterworth bant geçiren (0,5-50 Hz.) filtreleme ve [-1,1] aralığında normalizasyon uygulanmıştır.

BIS senkronizasyonu tamamlanan veri setinde toplam 32.760 örnek elde edilmiştir. Bu örneklerin %49,9'u (16.348 adet) derin anestezi, %33,3'ü (10.899 adet) ideal anestezi ve %16,8'i (5.513 adet) hafif anestezi sınıfına aittir.

Tablo 3. Yedi Sınıflandırma Modelinin Karşılaştırmalı Doğruluğu

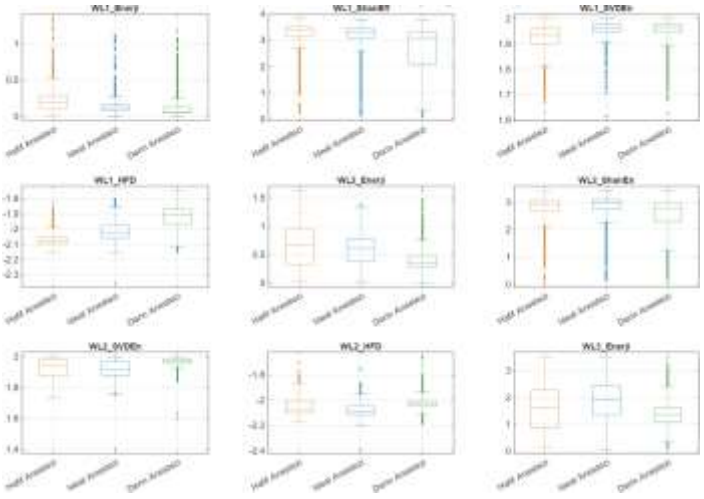
Model	Ort. Doğruluk (%)	Std (±)
CNN+HMM (Ayrık)	83,28	0,69
HMM-Temporal	82,75	0,46
1D-CNN	79,63	0,44
Random Forest	79,21	0,39
MLP	77,98	0,22
SVM-RBF	76,01	0,55
WKNN	72,44	0,69

Zamansal geçiş bilgisini kullanan iki yaklaşımın (CNN+HMM ve HMM-Temporal) diğer beş modele kıyasla en yüksek doğruluk değerlerine ulaştığı görülmektedir. Bu iki modelin ortak paydası, temel sınıflandırıcının çıktısının bir Gizli Markov Modeli ile zamansal olarak düzeltilmesidir; nitekim 1D-CNN'in HMM'siz hâliyle elde ettiği doğruluk (%79,63), HMM ile birleştirildiğinde 3,65 puanlık bir artış göstermektedir. Klasik makine öğrenmesi modelleri kendi aralarında karşılaştırıldığında, topluluk tabanlı Random Forest (%79,21) ve MLP (%77,98), mesafe/çekirdek tabanlı yöntemlere (SVM-RBF: %76,01; WKNN: %72,44) kıyasla daha yüksek doğruluk sergilemiştir. Çalışma neticesinde elde edilen bulgular Tablo 3'te verilmiştir.

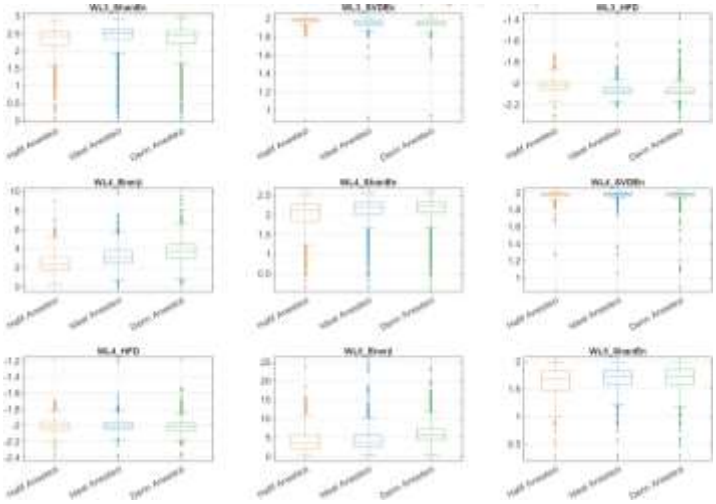
IV. TARTIŞMA

Hesaplanan 53 özniteliğin sınıf bazlı dağılımı kutu grafikleri ile incelenmiştir. Özniteliklerin sınıf bazlı dağılım grafikleri şekil 2'de verilmiştir. Bu grafiklerin yorumlanmasında kutu içindeki orta çizgi ilgili sınıfa ait medyan değerini, kutunun alt ve üst sınırları verilerin birinci ve üçüncü çeyreklik dilimlerini, kutudan dışarı uzanan çizgiler aykırı sayılmayan gözlem aralığını, kutunun tamamen dışında kalan noktalar ise istatistiksel aykırı değerleri göstermektedir. Bu doğrultuda her grafik dört unsur açısından değerlendirilmektedir: (i) üç sınıfa ait kutuların dikey eksendeki konumu ve birbirleriyle örtüşme düzeyi, (ii) medyan değerlerinin sınıflar arasındaki sıralaması, (iii) her kutunun genişliği, (iv) kutu dışında kalan aykırı değerlerin sayısı ve dağılımı.

Bu değerlendirme çerçevesi uygulandığında, bazı özniteliklerin (karmaşıklık ve entropi tabanlı ölçütler, spektral kenar frekansı, bant oranları) sınıflara ait kutularının birbirinden belirgin biçimde ayrıştığı; buna karşılık bir kısmının kutularının yoğun biçimde örtüştüğü görülmektedir. Bu gözlem, düşük bilgi taşıma kapasitesine sahip özniteliklerin ayıklanmasına yönelik NCA tabanlı seçim adımının gerekliliğini ortaya koymaktadır.



Şekil 2. a) Grup 1



b) Grup 2

Tablo 4. Sınıf Bazlı Kesinlik, Duyarlılık ve F1-Skor Değerleri

Model	Sınıf	Kesinlik (%)	Duyarlılık (%)	F1 (%)
RF	Hafif	80,86	81,39	81,12
RF	İdeal	70,63	70,37	70,50
RF	Derin	84,36	84,37	84,36
SVM-RBF	Hafif	79,03	78,03	78,52
SVM-RBF	İdeal	66,98	64,23	65,57
SVM-RBF	Derin	80,65	83,18	81,89
WKNN	Hafif	68,81	80,79	74,32
WKNN	İdeal	60,41	65,78	62,98
WKNN	Derin	83,98	74,07	78,71
MLP	Hafif	77,78	82,75	80,17
MLP	İdeal	67,57	71,61	69,52
MLP	Derin	85,94	80,62	83,19
HMM-T.	Hafif	86,96	83,93	85,40
HMM-T.	İdeal	73,10	77,68	75,31
HMM-T.	Derin	88,42	85,74	87,06
1D-CNN	Hafif	85,94	77,76	81,64
1D-CNN	İdeal	73,05	67,02	69,84
1D-CNN	Derin	81,66	88,67	85,00
CNN+HMM	Hafif	88,26	84,18	86,17
CNN+HMM	İdeal	75,37	75,03	75,19
CNN+HMM	Derin	86,89	88,47	87,67

Tüm modellerde uygulanan kesinlik, duyarlılık ve F1-Skor Değerleri Tablo 4'te verilmiştir. Modellerde ideal Anestezi sınıfının, komşu sınıflarla kısmi örtüşmesi nedeniyle en düşük F1-skoruna sahip olduğu görülmektedir; bu durum, ideal anestezi bandının klinik olarak keskin bir eşikle (BIS 40-60) tanımlanmasına karşın, altta yatan nörofizyolojik sürecin sürekli bir spektrum oluşturmasından kaynaklanabilir. Buna karşılık Derin Anestezi sınıfı, tüm modellerde en yüksek ve en istikrarlı F1-skoruna ulaşmıştır. En büyük fark, İdeal Anestezi sınıfında zamansal modellerin (CNN+HMM F1: %75,19; HMM-Temporal F1: %75,31) klasik SVM-RBF'ye (F1: %65,57) kıyasla yaklaşık 10 puanlık bir üstünlük sağlamasıdır. Bu çalışmanın veri setinde diğer çalışmalardan farklı olarak 1D-CNN + HMM Hibrit sınıflandırma modeli kullanılmıştır.

V. SONUÇLAR

Bu çalışmada, EEG tabanlı anestezi derinliği sınıflandırmasında yedi farklı makine öğrenmesi ve derin öğrenme yaklaşımı karşılaştırmalı olarak değerlendirilmiştir. Sonuçlar, zamansal geçiş bilgisini modele

dahil eden hibrit yaklaşımların (CNN+HMM: %83,28; HMM-Temporal: %82,75), bu bilgiyi kullanmayan klasik sınıflandırıcılara kıyasla tutarlı biçimde daha yüksek performans sağladığını göstermektedir.

TEŞEKKÜR

Tez çalışmam boyunca bilgi, birikim ve sabrıyla bana yol gösteren, her aşamada değerli katkılarını esirgemeyen danışman hocam Prof. Dr. Yücel Koçyiğit'e en içten teşekkürlerimi sunarım. Manevi desteklerini ve her zaman yanımda olduklarını hissettiren sevgili aileme çok teşekkür ederim.

KAYNAKLAR

[1] Rampil, I. J. (1998). A primer for EEG signal processing in anesthesia. Anesthesiology, 89(4), 980–1002.
[2] L. Vutskits and Z. Xie, "Lasting impact of general anaesthesia on the brain: Mechanisms and relevance," Nature Reviews Neuroscience, vol. 17, pp. 705-717, 2016.
[3] Dodds, C. (1999). General anaesthesia: Practical recommendations and recent advances. Drugs, 58(3), 453–467.
[4] T. E. Keys, "Historical vignettes: Dr. Arthur Ernest Guedel 1883-1956," Anesthesia & Analgesia, vol. 54, no. 4, pp. 442-443, 1975.
[5] Dartigues, J. F., Commenges, D., Letenneur, L., Tavernier, B., Garner, C., Barberger-Gateau, P., & Salamon, R. (1997). Cognitive predictors of dementia in elderly community residents. Neuroepidemiology, 16(1), 29–39.
[6] Moller, J. T., Cluitmans, P., Houx, P., Rasmussen, L. S., Canet, J., Rabbitt, P., ... & Hanning, C. D. (1998). Prolonged postoperative cognitive dysfunction in the elderly. The Lancet, 351(9106), 857–861.
[7] Ramsay, M. A. E. (2006). John Snow, MD: Anaesthetist to the Queen of England and pioneer epidemiologist. Proceedings (Baylor University Medical Center), 19(1), 24–28.
[8] Winterberg, A. V., Colella, C. L., Weber, K. A., & Varughese, A. M. (2018). The child induction behavioral assessment tool: A tool to facilitate the electronic documentation of behavioral responses to anesthesia inductions. Journal of PeriAnesthesia Nursing, 33(3), 296–303.
[9] Douglas, B. L. (1958). A re-evaluation of Guedel's stages of anesthesia: With particular reference to the ambulatory dental general anesthetic patient. The Journal of the American Dental Association, 57(1), 11–14.
[10] Hedenstierna, G., & Edmark, L. (2015). Effects of anesthesia on the respiratory system. Best Practice & Research Clinical Anaesthesiology, 29(3), 273–284.
[11] J. W. Bang, J.-S. Choi, and K. R. Park, "Noise reduction in brainwaves by using both EEG signals and frontal viewing camera images," Sensors, vol. 13, no. 5, pp. 6272-6294, 2013.
[12] J. Liu, H. Singh, and P. F. White, "Electroencephalographic bispectral index correlates with intraoperative recall and depth of propofol-induced sedation," Anesthesia & Analgesia, vol. 84, no. 1, pp. 185-189, 1997
[13] M. Teplan, "Fundamentals of EEG measurement," Measurement Science Review, vol. 2, no. 2, pp. 1-11, 2002.
[14] Kearse, L. A., Rosow, C., Zaslavsky, A., Connors, P., Dershwitz, M., & Denman, W. (1998). Bispectral analysis of the electroencephalogram predicts conscious processing of information during propofol sedation and hypnosis. Anesthesiology, 88(1), 25–34.
[15] Musizza, B., & Ribarič, S. (2010). Monitoring the depth of anaesthesia. Sensors, 10(12), 10896–10935.
[16] H. L. Kaul and N. Bharti, "Monitoring depth of anaesthesia," Indian Journal of Anaesthesia, vol. 46, no. 4, pp. 323-332, 2002.
[17] Petersen, J., Stockmanns, G., & Nahm, W. (2000). EEG analysis for assessment of depth of anaesthesia. In P. S. Szczepaniak, P. J. G. Lisboa, & J. Kacprzyk (Eds.), Fuzzy systems in medicine (Vol. 41, pp. 261–277). Heidelberg: Physica-Verlag HD.
[18] Liu, Q., Ma, L., Fan, S. Z., Abbod, M. F., & Shieh, J. S. (2018). Sample entropy analysis for the estimating depth of anaesthesia through human EEG signal at different levels of unconsciousness during surgeries. PeerJ, 6, e4817.
[19] Talha, A. Z., Eissa, N. S., & Shapiai, M. I. (2024). Applications of brain computer interface for motor imagery using deep learning: Review on recent trends. Journal of Advanced Research in Applied Sciences and Engineering Technology, 40(1), 96–116
[20] Elshekhidris, I. H., MohamedAmien, M. B., & Fragoon, A. (2023). Wavelet transforms for EEG signal denoising and decomposition. International Journal of Advanced Signal and Image Sciences, 9(1), 11–28.
[21] N. Phutela, P. Abhilash, K. Sreevathsan, and B. N. Krupa, "Intelligent analysis of EEG signals to assess consumer decisions: A study on neuromarketing," in Proc. INCET, 2022, pp. 1-7.
[22] T. M. Ha and H. Bunke, "Off-line, handwritten numeral recognition by perturbation method," IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 5, pp. 535-539, 1997.
[23] Huang, Y., Wen, P., Song, B., & Li, Y. (2022). Real-time depth of anaesthesia assessment based on hybrid statistical features of EEG. Sensors, 22(16), 6099.
[24] Miller, A., Sleigh, J. W., Barnard, J., & Steyn-Ross, D. A. (2004). Does bispectral analysis of the electroencephalogram add anything but complexity? British Journal of Anaesthesia, 92(1), 8-13.
[25] S. N. R. Gowda, N. C. Dvornek, V. Subbaraju, and M. Sahin, "The effects of individual differences, non-stationarity, and the importance of data partitioning decisions for training and testing of EEG cross-participant models," Sensors, vol. 21, no. 9, 3225, 2021.

8th International Conference on Contemporary Academic Research

August 05-06, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/iccar/home

Agreement Without Convergence? Constructive Ambiguity, Transatlantic Coherence, and Implementation in the Kosovo–Serbia Dialogue

Leonora Bajrami*

South East European University, North Macedonia

**(lb32351@seeu.edu.mk) Corresponding author*

Abstract – This paper examines why agreements in the Kosovo–Serbia dialogue can be reached even when the parties do not share the same understanding of what those agreements ultimately mean. It focuses on the relationship between constructive ambiguity and United States–European Union coordination in two negotiating episodes: the 2013 Brussels Agreement and the 2023 Agreement on the Path to Normalisation together with its Implementation Annex. The study uses qualitative document analysis and compares the political function of flexible language at the moment of agreement with the difficulties that emerge when that language has to be implemented. The argument is deliberately limited. It does not treat ambiguity as the sole explanation for implementation problems, nor does it assume that stronger transatlantic coordination automatically produces clearer outcomes. Instead, the comparison shows that coordinated external diplomacy can narrow the space for competing negotiating tracks and make compromise easier to sustain, while still leaving Kosovo and Serbia with different expectations about status, sequencing, and compliance. The two cases suggest that ambiguity is most useful when it creates temporary room for agreement and is later followed by clearer procedures, benchmarks, and monitoring. When clarification does not follow, disagreement tends to reappear during implementation. The paper therefore treats constructive ambiguity as a negotiating instrument whose value changes over time rather than as either an inherently successful or inherently problematic diplomatic technique.

Keywords – constructive ambiguity; Kosovo–Serbia dialogue; transatlantic coherence; European Union; United States; implementation; normalization

I. INTRODUCTION

The Kosovo–Serbia dialogue has produced a series of political and technical arrangements even though the parties continue to disagree over Kosovo's status and over the destination of the normalization process. This is not simply a contradiction. It is also a recurring feature of negotiations in which a more precise formula may be politically unacceptable, while a more open formula allows both sides to remain at the table. The difficulty is that the room for interpretation that helps an agreement to be accepted can later become the basis for disputes about what was actually accepted.

Research on the dialogue has already drawn attention to this problem. Bieber links part of the negotiating progress to the use of creative ambiguity in agreements between Serbia and Kosovo [3]. Gashi, Musliu, and Orbie go further by showing that ambiguity also concerns the way the dialogue itself is represented and recontextualized by the European Union [7]. In diplomatic studies, constructive ambiguity generally refers to the deliberate use of wording that does not settle every contested point when greater precision would make agreement more difficult [2], [4]. The concept is therefore useful only if the analytical distinction between reaching an agreement and implementing it is kept visible.

This paper asks how constructive ambiguity interacts with US–EU coherence in enabling agreements and shaping later implementation problems in the Kosovo–Serbia dialogue. The focus is narrower than a general

comparison of American administrations or a full assessment of transatlantic policy toward the Western Balkans. Instead, two episodes are examined: the 2013 Brussels Agreement and the 2023 normalization framework. They were selected because both created a common political text while leaving important questions open to competing interpretations.

The comparison starts from two expectations. Greater coordination between Washington and Brussels should make it more difficult for the parties to rely on competing external messages or alternative diplomatic channels, which may increase the chances of reaching a compromise. However, the situation changes once an agreement has been accepted. At that point, general political support has to be translated into decisions about sequencing, institutional responsibilities, benchmarks, and compliance. US–EU coordination can facilitate this process, but it cannot remove the different interpretations that Kosovo and Serbia themselves may attach to the agreement.

The paper brings constructive ambiguity and transatlantic coherence into the same analytical framework while keeping them conceptually distinct. Constructive ambiguity relates primarily to the openness of the negotiated text and the possibility of more than one interpretation. Transatlantic coherence, by contrast, concerns the degree to which the external actors supporting the dialogue pursue compatible objectives, messages, and diplomatic practices. These dimensions do not necessarily develop in parallel. Washington and Brussels may be closely coordinated while still accepting intentionally flexible wording, and even relatively precise provisions may become difficult to implement when external actors disagree about priorities or sequencing. The analysis therefore considers not only whether US–EU coordination existed, but also what it actually contributed at different stages of the negotiating process.

This temporal distinction also prevents the analysis from judging an agreement only by the moment at which it is announced. A flexible formula may be diplomatically successful in the short term if it closes a negotiation that would otherwise fail, yet that same formula can generate additional bargaining later when institutions have to translate it into practice. The relevant question is therefore not whether ambiguity is good or bad in the abstract. It is whether the type of ambiguity used at the negotiating stage is subsequently accompanied by mechanisms capable of narrowing competing interpretations. This is the standard used throughout the comparison of the 2013 and 2023 episodes.

II. Constructive Ambiguity and Multi-Actor Mediation

A. Constructive ambiguity as a negotiating device

Constructive ambiguity is best understood here as a bargaining technique rather than as a quality of language in the abstract. Berridge, James, and Lloyd describe it as a diplomatic device that leaves room for compromise on sensitive issues [2]. Byers likewise emphasizes the deliberate use of ambiguous language in negotiated legal texts where precise wording could prevent consensus [4]. The practical attraction is clear: parties can agree to the same formulation without first resolving every difference that lies behind it.

That possibility matters in the Kosovo–Serbia case because the dispute is not limited to individual policy questions. The parties approach the dialogue with different understandings of sovereignty, recognition, institutional authority, and normalization. From a constructivist perspective, political language is interpreted through prior identities and expectations rather than acquiring one uncontested meaning simply because it appears in an agreed text [13]. A provision can therefore be incorporated into different domestic narratives without those narratives converging. In such circumstances, ambiguity may temporarily contain disagreement rather than eliminate it.

A further distinction is useful between ambiguity about political end states and ambiguity about operational obligations. The first allows actors to avoid settling the ultimate status question; the second concerns what specific commitments require in practice. These forms can overlap, but they do not have identical consequences. Political ambiguity may be difficult to remove because it is part of the compromise itself, whereas operational ambiguity can in principle be reduced through implementation plans, sequencing, deadlines, and monitoring. For this reason, the paper does not treat every open formulation as equally problematic. The analytical concern is strongest where flexibility that was necessary for political acceptance also leaves the implementation process without a sufficiently shared understanding of the steps that follow.

This point also helps explain why different domestic narratives can coexist with the same negotiated formula. Agreement does not erase the political need of each side to present the outcome in a manner compatible with its established position. The text can therefore function as a meeting point between narratives without becoming a single narrative shared by both parties. In the short run this may be precisely what allows compromise. Over time, however, the gap between those narratives matters if officials rely on them to justify different priorities or different readings of the obligations. The practical durability of constructive ambiguity thus depends on whether later procedures can manage that interpretive gap without reopening the entire negotiation.

B. Hopeful ambiguity and divergent expectations

Troitskiy's notion of 'hopeful ambiguity' adds an important temporal element to this logic [11]. A party may accept an unclear settlement because it expects later political, legal, or institutional developments to favor its preferred interpretation. The agreement is then acceptable not because the disagreement has disappeared, but because the final distribution of gains and losses is postponed. This dynamic is particularly relevant when implementation unfolds in stages and when later decisions determine the practical value of the original text.

For that reason, the analysis distinguishes three connected moments: negotiated ambiguity, interpretive divergence, and implementation management. Ambiguity is constructive at the first moment when it widens the space for agreement without making later action impossible. Its costs increase when incompatible interpretations persist into implementation and there is no accepted procedure for clarifying priorities, sequencing obligations, or resolving disagreements about compliance. This distinction avoids treating ambiguous language as either inherently positive or inherently negative.

The three moments should not be understood as rigid chronological boxes. Interpretive divergence can already be visible during negotiation, and implementation questions may shape the wording chosen before an agreement is finalized. The distinction is instead analytical: it identifies the dominant function of ambiguity at each stage. During bargaining, ambiguity can lower the political cost of consent. During interpretation, the parties assign meaning to the common formula through their existing expectations. During implementation, those meanings encounter concrete institutional choices. The value of the concept lies precisely in following the same contested wording across these stages rather than assuming that signature marks the end of the dispute.

Hopeful ambiguity also draws attention to expectations about time. If each party believes that later developments will strengthen its own preferred interpretation, postponement can be rational even where disagreement is obvious. This can make an agreement more stable at the moment of closure but more fragile when the anticipated future begins to arrive. The implementation stage then becomes a test of whether the compromise was genuinely manageable or whether the parties merely deferred a distributive conflict. In this sense, the concept connects the political attraction of an open formula with the later incentives to resist a clarification that appears to favor the other side.

C. The conditioning role of US–EU coherence

US–EU coherence is treated as a conditioning factor, not as the outcome that the paper seeks to explain. In this context, coherence concerns the compatibility of American and European objectives, diplomatic contacts, initiatives, and policy instruments. Existing work on policy coherence and transatlantic cooperation shows that coordination can be examined at several levels, including declared objectives, institutional relations, operational behavior, and the use of instruments [8], [9]. Applying that distinction to the dialogue helps separate joint political messaging from the more demanding question of whether external actors also coordinate the practical management of implementation.

The relationship between coherence and ambiguity is therefore not linear. Coordinated mediators may knowingly tolerate flexible language because it is the price of obtaining agreement. The decisive question comes afterward: do they also converge on how obligations should be sequenced, monitored, and interpreted? If they do not, unity at the moment of agreement may coexist with continuing uncertainty about the agreement's operational meaning. In other words, coherence can strengthen leverage without supplying the substantive clarity that the parties themselves have avoided.

For the purposes of this paper, declarative coherence refers to the compatibility of public political messages; institutional coherence concerns the extent to which the relevant actors work through mutually supportive channels; operational coherence concerns the practical alignment of diplomatic activity; and instrument-level coherence concerns the compatibility of incentives, pressure, and other policy tools. These dimensions are analytically separate. External actors may speak in similar terms while still differing over timing or implementation priorities, or they may use different institutional channels while still pushing toward the same political objective. The framework is therefore used as a way of specifying where coordination matters rather than assigning a single overall score to the transatlantic relationship [8], [9].

This distinction is particularly important after an agreement has been reached. At that point, joint statements of support may have less practical effect than coordination over sequence, benchmarks, and responses to non-compliance. A common diplomatic message can reinforce the legitimacy of the framework, but implementation requires decisions about what happens first, how progress is assessed, and how disagreement over performance is handled. The paper therefore treats implementation coherence as a more demanding form of coordination than agreement-stage coherence. The latter can help create a common text; the former must help turn that text into a common process.

III. Materials and Method

The study uses a qualitative, focused comparison of two negotiating episodes: the 2013 First Agreement of Principles Governing the Normalization of Relations and the 2023 Agreement on the Path to Normalisation together with the Implementation Annex agreed in Ohrid. The cases are not treated as identical. They are compared because each represents a major externally facilitated agreement and because each makes it possible to observe the movement from common wording to disputes over meaning and implementation.

The empirical material is documentary. It consists of the texts and official material connected to the dialogue, European Union and United States statements, implementation-related documents, institutional reports, and selected academic literature [3], [5]–[7], [10]. No interview or survey evidence is presented as a completed finding in this conference paper. Each case is read through the same set of questions: Which issue was left open or flexible? What political function did that flexibility serve when the agreement was reached? Did US and EU diplomacy reinforce a common interpretation, or mainly a common acceptance of the text? What mechanism was available for implementation? How did competing interpretations affect the move from political commitment to practice?

The method has clear limits. Documentary comparison can identify patterns in language, official positioning, and implementation disputes, but it cannot establish that ambiguity alone caused non-implementation. Domestic political incentives, leadership strategies, security developments, and the unresolved status dispute remain plausible alternative explanations. The purpose is therefore analytical rather than monocausal: to identify where ambiguity and external coordination intersect and where they cease to solve the same problem.

The comparison proceeds in two steps. First, each episode is examined within its own political and textual setting in order to identify the issue on which flexibility was useful and the way external coordination supported acceptance. Second, the cases are compared across the same dimensions: the political purpose of ambiguity, the persistence of divergent interpretation, the degree and type of US–EU coherence, and the mechanisms available to manage implementation. This structured comparison is intended to keep the analysis focused on the relationship between the two concepts rather than turning the paper into a general history of the dialogue.

The documentary approach also requires caution in interpreting official language. Public statements are evidence of stated positions and diplomatic framing, not direct evidence of private motives. The paper therefore uses them to reconstruct how agreements were presented, what kinds of commitments were emphasized, and where interpretive differences became visible. Claims about causal intent are kept limited. Where the analysis speaks of the function of ambiguity, it refers to the political effect that flexible wording can have within the negotiation rather than claiming access to the private intentions of every participant.

IV. The 2013 Brussels Agreement: Compromise Through Flexible Meaning

The 2013 Brussels Agreement marked a shift from a largely technical process toward a more explicitly political framework of normalization. It opened the way for changes involving security, justice, local governance, and the integration of structures in northern Kosovo, while avoiding a definitive settlement of the status dispute [3], [5]. That combination is central to its political significance. The agreement addressed concrete institutional questions without requiring either side to adopt the other side's overall understanding of sovereignty.

The different readings of the arrangement were visible in the way its institutional provisions could be presented domestically. For Kosovo, integration could be framed as an extension and consolidation of its legal and administrative order. For Serbia, the provisions concerning the Association/Community of Serb-majority municipalities could be framed as preserving a collective institutional mechanism for the Serb community. These readings were not equivalent, but the wording gave both governments room to defend the compromise without declaring that their core political position had been abandoned. Bieber's analysis of the agreements captures this role of creative ambiguity [3].

External coordination strengthened the negotiating environment in which this compromise was reached. The European Union retained the formal facilitation role and linked the process to the broader perspective of European integration, while the United States offered political and strategic support to the EU-led framework. The significance of this complementarity was not that Washington and Brussels produced one detailed interpretation of every provision. Rather, they reduced the likelihood that the parties could rely on competing external frameworks. That made the negotiating process more credible and increased the political cost of walking away from compromise.

The limits of that success became clearer once broad political formulations had to be converted into institutional arrangements. Questions concerning institutional design, legal hierarchy, and the practical content of specific obligations remained contested [3]. The 2013 case therefore separates two capacities that are often treated as one: the capacity to help the parties close a negotiation and the capacity to help them implement what they have signed. External coherence contributed more directly to the first. The second required clarification that the agreement itself did not fully provide.

Seen in this way, the 2013 agreement illustrates why ambiguity can be simultaneously enabling and incomplete. Its political usefulness came from allowing the parties to move forward without requiring them to adopt a shared statement on sovereignty. The common text was therefore capable of producing institutional movement even while the underlying status dispute remained unresolved. That is a significant negotiating achievement. At the same time, the very flexibility that protected the agreement from the status question increased the importance of later clarification, because institutional provisions had to be located within legal and political frameworks that the parties understood differently [3].

The implementation problem was consequently not a simple reversal of the negotiating success. Rather, it was the next stage of the same compromise. Once broad formulas had to be converted into administrative and legal arrangements, questions that could be left open at the negotiating table became harder to avoid. This is why the case is analytically useful: it shows that an agreement can be effective as a device for political closure while remaining under-specified as a guide to implementation. Assessing the agreement only by whether it was signed, or only by later implementation disputes, would miss this change in function over time.

V. The 2023 Normalization Framework: Agreement Without a Shared End State

The 2023 Agreement on the Path to Normalisation and its Implementation Annex returned to the same structural dilemma under different political conditions. The framework sought to advance normalization without requiring an immediate act of formal mutual recognition [6], [10]. This made movement possible, but it also allowed the parties to preserve different views about status and about the legal and political destination of the process.

The interpretive gap was evident in official reactions. Kosovo could read the framework as movement toward the treatment of the two sides as equal parties and toward broader recognition of its statehood. Serbia continued to reject interpretations that would amount to formal recognition and referred to its constitutional

and international-law position. Stanicek documents these divergent reactions to the 2023 process [10]. A common text therefore did not amount to a common account of where normalization was supposed to lead.

The external setting nevertheless differed from periods in which American and European initiatives appeared more parallel than integrated. In 2023, US representatives broadly supported the EU-facilitated normalization framework and its implementation [6], [10]. This produced a relatively coherent external message around the need to move forward with the agreement. Yet this coherence did not settle disagreements over the order in which obligations should be implemented, which obligations should receive priority, or what political consequences should follow from different forms of compliance.

This case makes the distinction between sponsorship and implementation especially visible. External actors can agree that a text should be accepted and implemented while the parties continue to expect different future outcomes from the same text. Under those conditions, hopeful ambiguity becomes more than a theoretical possibility: the value of the agreement depends partly on what each side expects later steps to produce [11]. Repeated calls for implementation are then insufficient unless they are accompanied by agreed sequencing, benchmarks, monitoring, and a reasonably consistent external interpretation of compliance.

The 2023 framework also sharpens the problem of sequence. When parties attach different political meanings to the same agreement, the order of implementation is not merely technical. Priority can be interpreted as evidence of what the agreement is fundamentally about and which side is expected to move first. A disagreement over sequence can therefore reproduce the larger disagreement over the agreement's political direction. This helps explain why implementation cannot be reduced to repeating that all obligations must be fulfilled. Even where that principle is accepted, the practical process still requires a sufficiently common understanding of order, benchmarks, and the consequences attached to delay or partial compliance.

Compared with 2013, the 2023 episode therefore does not suggest that constructive ambiguity disappeared; rather, it shows its continued relevance under a framework supported by relatively coordinated external diplomacy. The important comparison is not that the two agreements are identical, but that both reveal the same analytical separation between acceptance and convergence. In each case, external support helped sustain a common framework, while divergent expectations survived inside that framework. The persistence of this gap is what makes post-agreement management central to the argument.

VI. Comparative Discussion

The comparison reveals a recurring pattern, but not a mechanical one. In both 2013 and 2023, flexible formulations helped bridge positions that could not be reconciled through explicit status language. In both episodes, coordinated external diplomacy strengthened the pressure to accept a common framework. Yet neither agreement erased the difference between accepting a text and sharing its meaning. Once negotiations moved into implementation, the central dispute shifted from whether there would be an agreement to what the agreement required in practice.

This shift explains why constructive ambiguity should be evaluated over time. At the negotiating stage, ambiguity can widen the zone of possible agreement by postponing questions that would otherwise produce deadlock. During implementation, the same openness gives political actors room to prioritize different provisions, attach different legal meanings to the text, or argue that implementation must begin with the other side's obligations. The technique can therefore succeed in one phase and generate costs in another without this being a logical contradiction.

Transatlantic coherence modifies this process but does not remove it. A coordinated US–EU approach can reduce opportunities for forum shopping, reinforce incentives, and make it harder for the parties to avoid the main negotiating framework by appealing to competing external sponsors. What it cannot do automatically is convert a politically flexible compromise into a detailed implementation regime. Coherence in messaging is not the same as coherence in implementation design. If Washington and Brussels support the same agreement but do not also reinforce a shared sequence, monitoring framework, and consequences for non-implementation, their unity may help produce an agreement whose operational meaning remains disputed.

The two cases also help distinguish ambiguity inside the agreement from ambiguity in the external division of labor. In 2013, the external roles were comparatively complementary even though important provisions remained open to competing interpretations. In 2023, the normalization framework also received broad US–EU backing, but the parties still attached different political trajectories to the same text. The main difficulty was therefore not necessarily disagreement between Washington and Brussels. It was the inability of external coherence, by itself, to generate an internal convergence of meaning between Kosovo and Serbia.

This finding narrows the claim that stronger US–EU coordination necessarily produces progress. The claim is most persuasive at the stage of negotiation and agreement-making, where coordinated pressure and incentives can influence whether a common text is accepted. It is less complete as an explanation of implementation. Implementation depends on an additional mechanism: the conversion of political flexibility into sufficiently specific, sequenced, and monitored obligations. That mechanism is related to transatlantic coherence, but it is analytically distinct from it.

The comparison can therefore be summarized as a two-stage mechanism. In the first stage, external coherence can increase the feasibility of agreement by reducing contradictory signals and strengthening the incentives to remain within one negotiating framework. Constructive ambiguity complements that effect by lowering the substantive cost of consent where precise wording would expose irreconcilable positions. In the second stage, however, the same combination is insufficient. The compromise has to be operationalized, and this requires a further layer of coordination around interpretation, sequence, monitoring, and consequences. Without that layer, the mechanisms that facilitated agreement may merely transfer part of the bargaining conflict into implementation.

This also clarifies the relationship between ambiguity and non-implementation. The paper does not argue that open wording mechanically causes delay or non-compliance. Rather, ambiguity creates a space in which different interpretations can remain politically defensible. Whether that space becomes obstructive depends on how actors use it and on whether institutions and mediators provide procedures that narrow the range of plausible interpretations. The causal claim is therefore conditional: ambiguity raises the management burden after agreement, and the effect of that burden depends on the quality of post-agreement coordination and clarification.

A second implication concerns the level at which clarification should occur. The comparison suggests that not every political disagreement has to be resolved before implementation can begin. What matters is whether the parties and mediators can separate questions that must remain politically open from questions that require operational precision. An agreement may continue to leave the final status issue unresolved while still specifying procedures for timing, institutional responsibility, reporting, and review. In that sense, post-agreement clarification does not necessarily undo constructive ambiguity. It can preserve the political compromise while reducing uncertainty in the parts of the agreement that have to guide concrete action.

A third implication concerns the role of external actors. If mediators focus primarily on obtaining formal acceptance, they may overestimate the durability of the result. Agreement-stage leverage and implementation-stage management are related but distinct tasks. The latter requires sustained attention to the meanings that the parties attach to the text and to the way those meanings influence sequencing and compliance. This does not mean that external actors can impose a single interpretation in every contested area. It means that coordinated mediation is more likely to remain effective when it identifies where interpretive flexibility can safely continue and where ambiguity has to be narrowed for implementation to proceed. The 2013 and 2023 cases both point to this distinction.

VII. CONCLUSION

The comparison of the 2013 and 2023 agreements indicates that constructive ambiguity and transatlantic coherence contribute to the dialogue in different ways. Flexible wording can make agreement possible when Kosovo and Serbia are unwilling to resolve their underlying disagreement over status, while coordinated US–EU diplomacy can strengthen the negotiating framework and reduce contradictory external signals. Neither factor, however, ensures that the parties will later interpret their commitments in the same way. Once implementation begins, unresolved differences may re-emerge in debates over sequencing,

institutional arrangements, and compliance. The findings therefore suggest that ambiguity is most useful when political flexibility is followed by greater operational clarity. Procedures, deadlines, benchmarks, monitoring arrangements, and credible responses to non-implementation can help preserve the compromise without requiring every political disagreement to be settled at the moment of agreement.

The conclusions remain limited by the design of the study. This paper compares two episodes through documentary evidence and does not claim a final causal explanation of implementation failure. A broader research project could test the argument through interviews and perception data, especially by examining how officials, experts, and other political actors understand the same provisions and the role of external sponsors. Such evidence would make it possible to identify more precisely when ambiguity continues to facilitate cooperation, when it becomes obstructive, and whether coordinated mediators can successfully manage the transition from negotiated flexibility to operational clarity.

REFERENCES

[1] E. Adler and V. Pouliot, Eds., International Practices. Cambridge: Cambridge University Press, 2011.

[2] G. R. Berridge, A. James, and L. Lloyd, The Palgrave Macmillan Dictionary of Diplomacy, 3rd ed. Basingstoke: Palgrave Macmillan, 2012.

[3] F. Bieber, "The Serbia–Kosovo agreements: An EU success story?" Review of Central and East European Law, vol. 40, no. 3–4, pp. 285–319, 2015, doi: 10.1163/15730352-04003008.

[4] M. Byers, "Still agreeing to disagree: International security and constructive ambiguity," Journal on the Use of Force and International Law, vol. 8, no. 1, pp. 91–114, 2021, doi: 10.1080/20531702.2020.1761656.

[5] European External Action Service, "Belgrade–Pristina Dialogue," 2021.

[6] European External Action Service, "Agreement on the path to normalisation between Kosovo and Serbia," Feb. 27, 2023.

[7] K. Gashi, V. Musliu, and J. Orbie, "Mediation through recontextualization: The European Union and the dialogue between Kosovo and Serbia," European Foreign Affairs Review, vol. 22, no. 4, pp. 533–550, 2017, doi: 10.54648/EERR2017039.

[8] M. Nilsson, T. Zamparutti, J. E. Petersen, B. Nykvist, P. Rudberg, and J. McGuinn, "Understanding policy coherence: Analytical framework and examples of sector–environment policy interactions in the EU," Environmental Policy and Governance, vol. 22, no. 6, pp. 395–423, 2012, doi: 10.1002/eet.1589.

[9] K. Raube and R. Vega Rubio, "Coherence at last? Transatlantic cooperation in response to the geostrategic challenge of China," Politics and Governance, vol. 10, no. 2, pp. 176–185, 2022, doi: 10.17645/pag.v10i2.5022.

[10] B. Stanicek, Belgrade–Pristina dialogue: The rocky road towards a comprehensive normalisation agreement, PE 689.371. Brussels: European Parliamentary Research Service, 2023.

[11] M. Troitskiy, "'Hopeful ambiguity' in negotiation closure," Negotiation Journal, vol. 32, no. 4, pp. 399–414, 2016.

[12] United Nations General Assembly, Resolution 64/298, 2010.

[13] A. Wendt, "Anarchy is what states make of it: The social construction of power politics," International Organization, vol. 46, no. 2, pp. 391–425, 1992, doi: 10.1017/S0020818300027764.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Enhanced Speed Control of DTC based PMSM using Super Twisting Sliding Mode Controller

Mohamed Bashir BIHI [1*], Hilmi AYGÜN [2] and Yunus Emre KARASU [3]

[1] *Department of Electrical and Electronics Engineering, Karabük University, Karabük, Türkiye*

[2] *Department of Mechatronics Engineering, Karabük University, Karabük, Türkiye*

[3] *Department of Electrical and Electronics Engineering, Karabük University, Karabük, Türkiye*

[*] *mohamedbashirbihi@gmail.com*

Abstract – This paper presents a comparative investigation of conventional sliding mode control (CSMC) and super twisting sliding mode control (STSMC) for speed regulation of a direct torque control (DTC) based permanent magnet synchronous motor (PMSM) drive. The two nonlinear control strategies are implemented in MATLAB/Simulink and evaluated under constant-speed references, variable-speed commands, and external load disturbances. The comparative analysis focuses on settling time, speed-tracking accuracy, transient response, and disturbance rejection capability. Under constant-speed operation at 500 rpm and 100 rpm, STSMC achieves settling times of 0.0019 s at both reference speeds, compared with 0.0065 s and 0.0094 s for CSMC, respectively. Under variable reference speed conditions, STSMC provides faster convergence with shorter settling times and no overshoot compared with CSMC. Furthermore, under load disturbances of 1 Nm and 2 Nm at a 500-rpm reference speed, STSMC exhibits maximum speed deviations of 11.89 rpm and 10.72 rpm, respectively, compared with 12.86 rpm and 49.73 rpm for CSMC. These results demonstrate that STSMC provides faster transient convergence and superior disturbance rejection while maintaining accurate speed tracking. Therefore, STSMC represents a promising nonlinear control strategy for high-performance DTC-based PMSM drives requiring fast dynamic response and strong robustness against operating condition variations.

Keywords – PMSM, DTC, nonlinear control, CSMC, STSMC

I. INTRODUCTION

Permanent magnet synchronous motors (PMSMs) have become an important class of electric machines for high-performance motion-control applications because of their high efficiency, high torque-to-volume ratio, and compact construction [1]. These characteristics have supported their increasing use in applications such as electric vehicles, robotics, aerospace actuation, servo systems, and advanced industrial drives, where accurate speed regulation and rapid transient response are essential [2–5]. In such applications, the quality of the control strategy has a direct influence on the dynamic and steady-state behaviour of the drive, particularly during rapid speed changes and external load disturbances.

Direct Torque Control (DTC) is one of the established control techniques for high-performance AC drives. Since its early development by Takahashi [6] and Depenbrock [7], DTC has attracted considerable attention because it regulates the stator flux and electromagnetic torque directly through inverter voltage-vector selection rather than relying on the conventional current-control structure used in field-oriented control (FOC) [8,9]. This direct control principle provides a relatively simple control structure and very rapid torque response, making DTC particularly attractive for applications requiring fast dynamic behaviour. Nevertheless, the conventional hysteresis-based DTC structure presents well-known limitations, including torque and flux ripple, variable switching frequency, and degraded behaviour in some operating regions [8–11]. In DTC-based PMSM drives, the outer speed loop typically uses a PI controller, valued for

its simplicity and good performance near nominal conditions. However, PMSM drives exhibit inherently nonlinear behaviour arising from coupled electromagnetic and mechanical dynamics, inverter switching, and parameter variations. Fixed gain PI controllers are thus sensitive to operating point changes, parameter uncertainty, and load disturbances, and poor gain tuning can cause overshoot or slow transient response motivating research into nonlinear, adaptive, intelligent, and optimization-based control research for PMSM drives.

Several approaches have been investigated to improve the performance of PMSM control systems. Adaptive control has been employed to compensate for disturbances and variations in the drive system, while fuzzy-logic-based techniques have been investigated to improve the ability of the controller to cope with nonlinear behaviour and uncertainty [12,13]. Optimization algorithms have also been increasingly incorporated into PMSM control, mainly for the adjustment of controller parameters. For example, genetic-algorithm (GA) based tuning has been applied to PI-controlled DTC systems, while grey wolf optimization (GWO) and particle swarm optimization (PSO) have been employed to improve controller performance and parameter selection [14–16]. These approaches can improve transient response, steady-state accuracy, and disturbance rejection; however, many optimization-based studies retain a fundamentally linear PI structure. Consequently, the resulting controller may still exhibit limited robustness when the system is subject to significant nonlinearities or disturbances.

To overcome the robustness limitations inherent in PI-based control, nonlinear strategies, most notably Sliding Mode Control (SMC), have been extensively investigated for PMSM drive applications [17–19]. By forcing the system trajectories onto a predefined sliding surface, SMC ensures strong robustness against parameter uncertainties and external disturbances. Nevertheless, its discontinuous switching law inherently induces chattering, which contributes to elevated mechanical stress, acoustic noise, and switching losses [20,21]. To address this drawback while retaining robustness and finite-time convergence, higher-order SMC variants including integral terminal, super-twisting, smooth super-twisting, and fast-terminal schemes have subsequently been proposed [22–25]. The conventional SMC (CSMC) offers strong robustness against parameter variations and external disturbances, making it suitable for the nonlinear dynamics of PMSM. Compared with conventional PI control, SMC can provide improved speed stability and disturbance rejection under uncertain operating conditions [26]. However, the main limitation of conventional SMC is the high-frequency chattering generated by its discontinuous switching action. In PMSM drives, this phenomenon can cause torque and current ripples and may increase switching losses. Therefore, Super-Twisting Sliding Mode Control (STSMC) has been introduced to preserve the robustness of SMC while reducing chattering and maintaining fast tracking performance [27].

Accordingly, this paper investigates and compares the speed-control performance of CSMC and STSMC integrated into a DTC-based PMSM drive. The comparison focuses on important performance indicators including rise time, settling time, overshoot, steady-state tracking accuracy, transient behaviour during speed-reference changes, and robustness against load disturbances.

II. Materials and Method

Describe in detail the materials and methods used when conducting the study. The citations you make from different sources must be given and referenced in references.

A. Dynamic Model of the PMSM

The PMSM mathematical model is established in the synchronously rotating *d-q* coordinate system, where the stator voltage equations are given in terms of the *d-q* axis flux linkages:

$$v_d = R_s i_d + \frac{d\psi_d}{dt} - \omega_r \psi_q \tag{1}$$

$$v_q = R_s i_q + \frac{d\psi_q}{dt} + \omega_r \psi_d \tag{2}$$

where v_d, v_q, i_d, i_q, and ψ_d, ψ_q denote the *d-q* axis stator voltages, currents, and flux linkages, respectively; ω_r represents the electrical rotor speed, and R_s is the stator resistance. The flux-linkage equations are consequently given by

$$\psi_d = L_d i_d + \psi_m\,, \quad \psi_q = L_q i_q \tag{3}$$

where L_d, L_q, ψ_m are the *d-q* axis stator inductances and permanent magnet flux, respectively. The mechanical dynamics are described by:

$$J\frac{d\omega_r}{dt} = T_e - B\omega_r - T_L \tag{4}$$

where J is the rotor inertia, B denotes the viscous friction coefficient, T_e and T_L represents the electromagnetic torque and load torque, respectively. The stator flux components are subsequently estimated in the stationary ψ_α, ψ_β frame by integrating the stator voltage after subtracting the voltage drop across the stator resistance, yielding:

$$\psi_\alpha(t) = \int (v_\alpha - R_s i_\alpha)\, dt \tag{5}$$

$$\psi_\beta(t) = \int \left(v_\beta - R_s i_\beta\right) dt \tag{6}$$

The stator flux magnitude and the flux angle are then computed as:

$$\psi_s = \sqrt{\psi_\alpha^2 + \psi_\beta^2}, \quad \theta = \tan^{-1}\left(\frac{\psi_\beta}{\psi_\alpha}\right) \tag{7}$$

The electromagnetic torque is expressed as:

$$T_e = \frac{3}{4}P\left(\psi_\alpha i_\beta - \psi_\beta i_\alpha\right) \tag{8}$$

where P is the pole number.

B. *Principles of Direct Torque Control*

DTC is a hysteresis-based control strategy for AC motor drives that provides direct and decoupled regulation of stator flux and electromagnetic torque, offering a simple control structure and fast dynamic response without requiring the coordinate transformations and inner current-control loops typically employed in FOC. As shown in Fig.1, the conventional DTC scheme consists of stator flux and electromagnetic torque estimators, hysteresis comparators for the corresponding flux and torque errors, and a switching table that directly selects the appropriate inverter voltage vector according to the torque, flux demands and the stator-flux sector to regulate the electromagnetic torque and stator-flux magnitude. The corresponding switching rules are presented in Table 1, where $d\psi$ and dT_e denote the outputs of the stator-flux and torque hysteresis comparators, respectively, while S_i represents the *i*-th stator-flux sector and V_i denotes the corresponding inverter voltage vector.

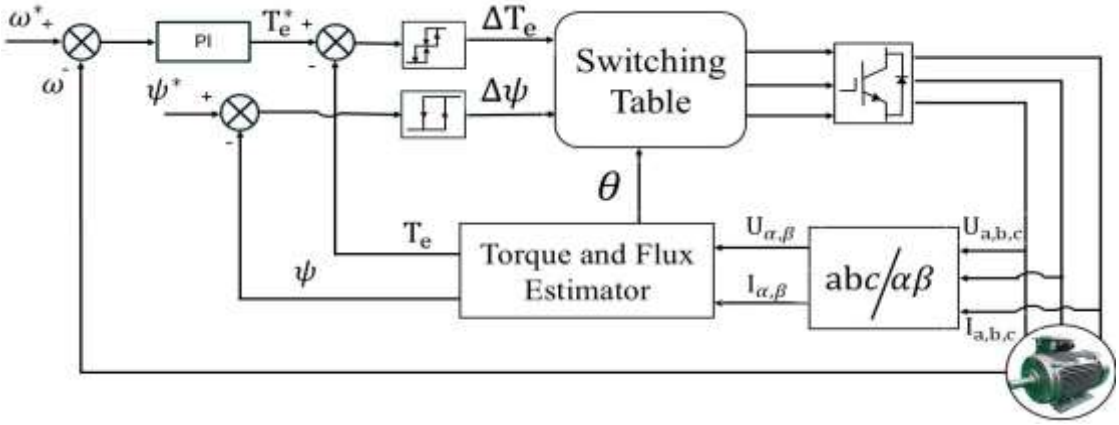


Fig. 1 Block diagram of the DTC-based PMSM drive

Table 1. Switching table for DTC

$d\psi$	dT_e	S_1	S_2	S_3	S_4	S_5	S_6
	1	V_2	V_3	V_4	V_5	V_6	V_1
1	0	V_7	V_0	V_7	V_0	V_7	V_0
	-1	V_6	V_1	V_2	V_3	V_4	V_5
	1	V_3	V_4	V_5	V_6	V_1	V_2
0	0	V_0	V_7	V_0	V_7	V_0	V_7
	-1	V_5	V_6	V_1	V_2	V_3	V_4

C. *Sliding Mode Controllers*

1. Conventional Sliding Mode Controller

The CSMC employs a variable-structure control law comprising an equivalent term that maintains the trajectory on the sliding surface and a discontinuous switching term that drives it toward the surface and rejects bounded matched disturbances. The sign-based switching action provides robustness against uncertainties. The sliding surface for the DTC-based PMSM is defined as:

$$s = Ce + \dot{e} \tag{9}$$

where C is a positive design constant, and e denotes the speed tracking error as given follows.

$$e = \omega_r - \omega_{ref} \tag{10}$$

By substituting Eq. (10) into Eq. (9) and assuming a constant speed reference, differentiating the speed tracking error in Eq. (9) with respect to time gives:

$$\dot{s} = C\dot{\omega}_r + \ddot{\omega}_r = 0 \tag{11}$$

Substituting the mechanical equation of motion of the PMSM into Eq. (11) gives:

$$\dot{\omega}_r \left(C - \frac{B}{J}\right) + \frac{\dot{T}_e}{J} - \frac{\dot{T}_L}{J} = 0 \tag{12}$$

Rearranging Eq. (12) to express the equivalent control component in terms of the torque yields:

$$u_{eq} = T_L - J\omega_r \left(C - \frac{B}{J}\right) \tag{13}$$

To satisfy the reaching condition and ensure that the system trajectories are driven toward the sliding surface, a discontinuous switching control component is introduced as follows:

$$u_{sw} = -K\,\mathrm{sign}(s) \tag{14}$$

where K is the switching gain that determines the aggressiveness of the discontinuous control action in suppressing disturbances and parameter uncertainties while maintaining the system on the sliding surface.

The total CSMC control signal is then obtained by combining the equivalent and switching control components as follows:

$$u = u_{eq} + u_{sw} \tag{15}$$

By substituting Eqs. (13) and (14) into Eq. (15), the torque reference generated by the CSMC is expressed as:

$$u = T_{e_{ref}} = T_L - J\omega_r\left(C - \frac{B}{J}\right) - K\,\mathrm{sign}(s) \tag{16}$$

The discontinuous switching action in Eq. (16) is generated using the signum function, which determines the direction of the control effort according to the sign of the sliding variable s. The signum function is defined as:

$$\mathrm{sign}(s) = \begin{cases} 1, & s > 0 \\ 0, & s = 0 \\ -1, & s < 0 \end{cases} \tag{17}$$

2. Super Twisting Sliding Mode Controller

To enhance convergence, particularly when the system operates far from the equilibrium point, the STSMC introduces additional linear terms and a tuneable homogeneity degree into the CSMC formulation. This modification improves transient performance while preserving finite-time convergence and continuous control action, thereby reducing chattering. The equivalent control component u_{eq} remains unchanged, while only the switching component u_{sw} is redesigned as follows;

$$u_{sw} = -K_1|s|^{\frac{p-1}{p}}\mathrm{sign}(s) + y \tag{18}$$

$$\dot{y} = -K_2|s|^{\frac{p-2}{p}}\mathrm{sign}(s) \tag{19}$$

where K_1, K_2 are positive controller gains and $p = 2$. The complete STSMC control law is obtained by combining the equivalent control term of Eq. (13) with the switching control component of Eq. (18).

$$u = T_{e_{ref}} = T_L - J\omega_r\left(C - \frac{B}{J}\right) - K_1|s|^{\frac{p-1}{p}}\mathrm{sign}(s) - \int K_2|s|^{\frac{p-2}{p}}\mathrm{sign}(s)\,dt \tag{20}$$

III. RESULTS AND DISCUSSION

This section presents the simulation results obtained for the considered control strategies. The controllers were implemented and evaluated in MATLAB/Simulink using the PMSM parameters listed in Table 2. Their performance was comparatively assessed in terms of speed-tracking accuracy and disturbance rejection under different reference-speed and load conditions.

Table 2. Motor Parameters

Symbols	Value
R_s	2.98 Ω
L_d, L_q	7 mH
P	4
ψ_m	0.125 Wb
J	0.47×10^{-4} kg·m²
B	11×10^{-6} N·m·s/rad

A. Case 1: Constant Reference Speed

The constant-speed tracking performance of the CSMC and STSMC controllers was evaluated at reference speeds of 500 and 100 rpm under no-load conditions, as shown in Fig. 2(a) and (b), respectively. At 500 rpm, the CSMC controller achieves a settling time of 0.0065 s, whereas the STSMC controller reduces the settling time to 0.0019 s. Similarly, at 100 rpm, CSMC exhibits a settling time of 0.0094 s, while STSMC achieves a shorter settling time of 0.0019 s. Overall, STSMC demonstrates superior settling-time performance compared with CSMC at both reference speeds. It is also observed that both controllers exhibit no overshoot in their responses.

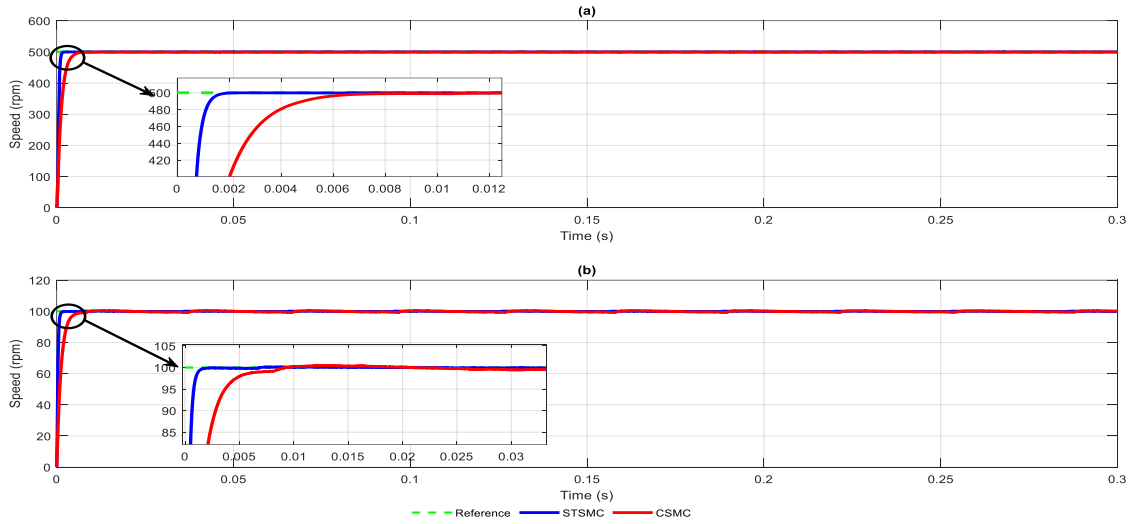


Fig 2. Transient response of the DTC-based PMSM under constant speed references as (a) 500, (b) 100 rpm.

B. Case 2: Variable Reference Speed

The variable-speed tracking performance of the CSMC and STSMC controllers was evaluated under two reference-speed profiles: a step change from 500 rpm to 100 rpm followed by an increase to 200 rpm, as shown in Fig. 3(a), and a step change from 500 rpm to 400 rpm followed by an increase to 750 rpm, as shown in Fig. 3(b). For the speed reduction from 500 to 100 rpm, STSMC achieves a settling time of 0.00168 s, compared with 0.00672 s for CSMC, while the subsequent increase from 100 to 200 rpm results in settling times of 0.00141 s and 0.00594 s for STSMC and CSMC, respectively. Similarly, for the 500–400 rpm transition in Fig. 3(b), STSMC settles within 0.00118 s, whereas CSMC requires 0.00481 s. For the subsequent increase from 400 to 750 rpm, the settling times are 0.00187 s for STSMC and 0.00760 s for CSMC. Overall, STSMC consistently achieves shorter settling times than CSMC during both speed reductions and increases, demonstrating faster transient convergence under variable-speed operation.

A. Case 3: Varying load in Constant Reference

The variable-load disturbance rejection performance of the STSMC and CSMC controllers was evaluated at a constant reference speed of 500 rpm by applying load torques of 1 Nm and 2 Nm at 0.1 s, as shown in

Fig. 4(a) and (b), respectively. At 1 Nm, the STSMC controller exhibits a speed reduction from 500 rpm to approximately 488.11 rpm, corresponding to a maximum speed deviation of 11.89 rpm, while CSMC decreases to approximately 487.14 rpm, resulting in a larger deviation of 12.86 rpm. The STSMC response recovers toward the 500-rpm reference more rapidly than CSMC. Under the increased load torque of 2 Nm, the difference becomes more pronounced: STSMC reaches a minimum speed of approximately 489.28 rpm, corresponding to a deviation of only 10.72 rpm, whereas CSMC drops to approximately 450.27 rpm, resulting in a substantially larger deviation of 49.73 rpm. Furthermore, STSMC returns to the reference speed considerably faster than CSMC. Overall, the results demonstrate that STSMC provides superior disturbance rejection, with smaller speed deviations and faster recovery under increasing load torque while maintaining accurate reference-speed tracking.

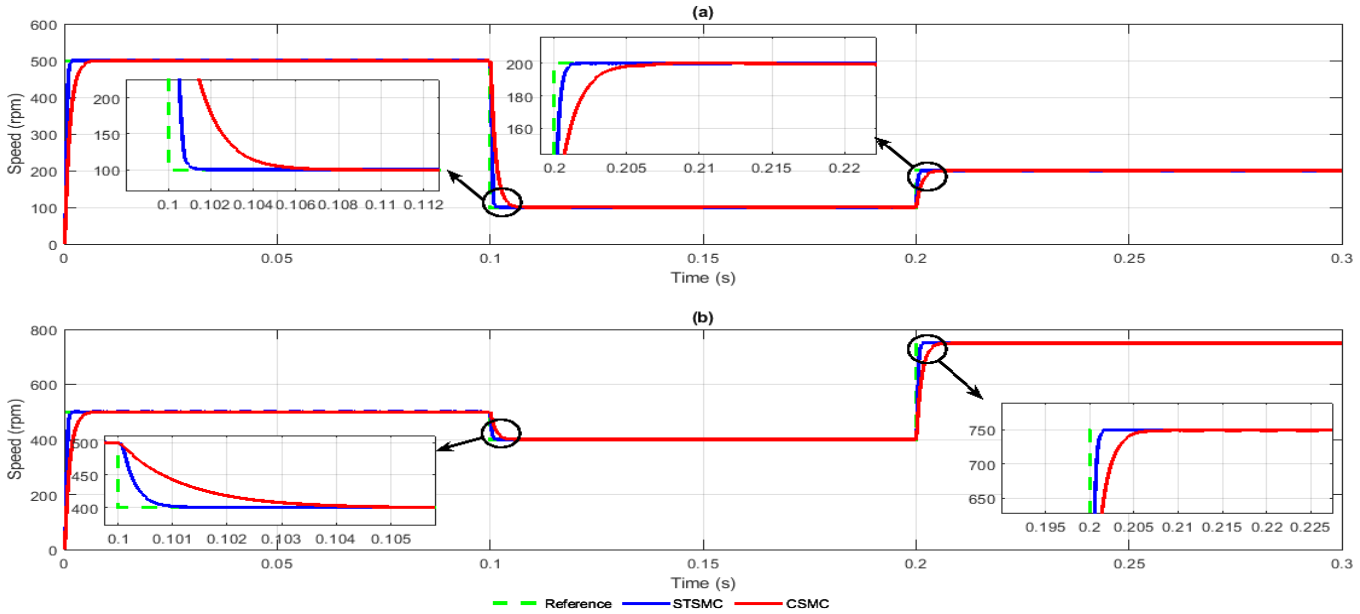


Fig 3. Dynamic response of the DTC based PMSM under varying speed from 500 to (a) 100 to 200 rpm, (b) 400 to 750 rpm.

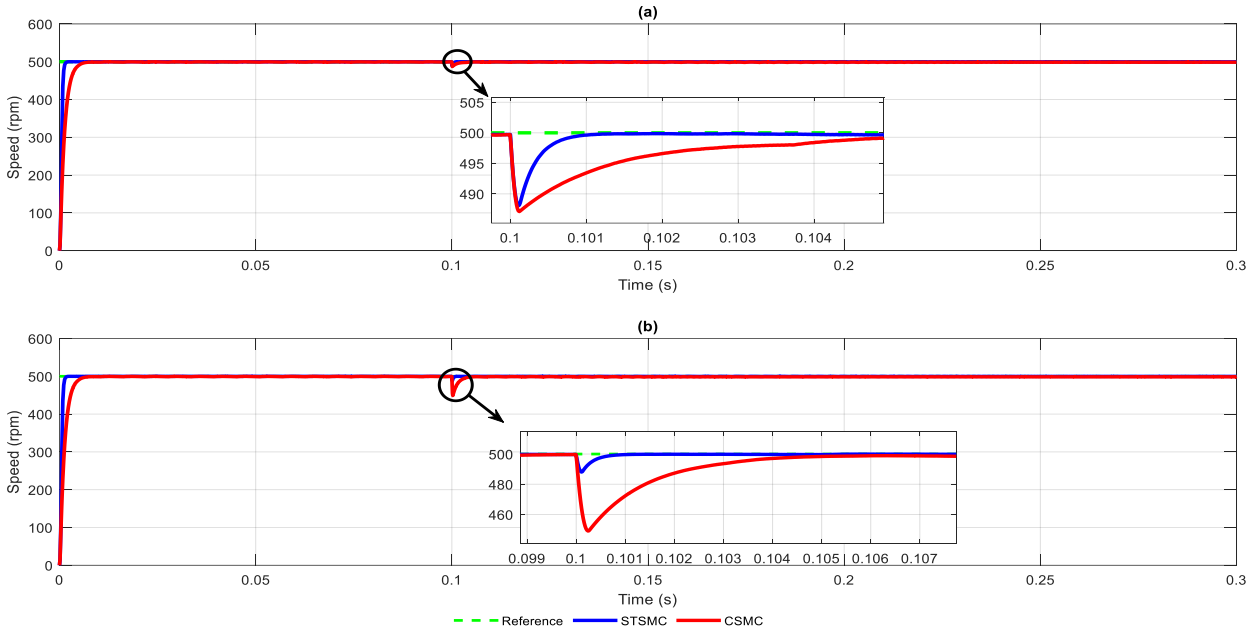


Fig 4. Speed Response of STSMC and CSMC Under Variable Load Torque at 500 rpm (a) 1 Nm, (b) 2 Nm.

IV. CONCLUSION

This paper presents a comparative investigation of CSMC and STSMC for speed control of a DTC-based PMSM drive. The simulation study demonstrates that both nonlinear control strategies provide effective

speed regulation under different operating conditions, while STSMC offers a more favorable overall dynamic response. In particular, the super-twisting formulation enhances convergence behavior and ensures reliable tracking performance under varying operating conditions. The comparative results indicate that STSMC provides enhanced robustness against variations in reference speed and external load disturbances without compromising tracking capability. These characteristics make STSMC a promising alternative to CSMC for high-performance PMSM drives, particularly in applications requiring fast dynamic response and strong disturbance rejection.

REFERENCES

[1] J. Chukwunweike, "Design and Optimization of Energy-Efficient Electric Machines for Industrial Automation and Renewable Power Conversion Applications," 2019. [Online]. Available: www.ijcat.com548

[2] An, Q., Lu, Y., & Zhao, M. (2024). Review of key technologies of the high-speed permanent magnet motor drive. Energies, 17(21), 5252.

[3] G. Anbalagan and A. Senthilnathan, "Comprehensive Review on Intelligent controller for Permanent Magnet Synchronous motor," Journal of Electrical Systems, vol. 21, no. 1, pp. 321–328, Apr. 2025.

[4] P. Pietrzak and M. Wolkiewicz, "Application of continuous wavelet transform and convolutional neural networks in fault diagnosis of PMSM stator windings," Bulletin of the Polish Academy of Sciences: Technical Sciences, vol. 72, no. 5, Sep. 2024, doi: 10.24425/bpasts.2024.150202.

[5] J. Tang, C. Liang, Y. Wang, S. Lu, and J. Zhou, "A stator fault diagnosis method based on the offline motor parameter measurement for PMSM," World Electric Vehicle Journal, vol. 12, no. 4, Dec. 2021, doi: 10.3390/wevj12040248.

[6] Takahashi I, Noguchi TA. New quick-response and high-efficiency control strategy of an induction motor. IEEE Trans Ind Appl 1986;IA-22(5):820–7.

[7] Depenbrock M. Direct self control (DSC) of inverter-fed induction machine. IEEE Trans Power Electron 1988;3(4):420–9.

[8] A. U. Rehman, H. H. Choi, and J. W. Jung, "An Optimal Direct Torque Control Strategy for Surface-Mounted Permanent Magnet Synchronous Motor Drives," IEEE Trans. Industr. Inform., vol. 17, no. 11, pp. 7390–7400, Nov. 2021, doi: 10.1109/TII.2021.3053107.

[9] A. H. Abosh, Z. Q. Zhu, and Y. Ren, "Cascaded Direct Torque Control of Unbalanced PMSM with Low Torque and Flux Ripples," IEEE Trans. Power Electron., vol. 33, no. 2, pp. 1740–1749, Feb. 2018, doi: 10.1109/TPEL.2017.2683460.

[10] De Klerk, M. L., & Saha, A. K. (2021). A comprehensive review of advanced traction motor control techniques suitable for electric vehicle applications. IEEE Access, 9, 125080-125108.

[11] I. M. Alsofyani, K. Y. Kim, S. S. Lee, and K. B. Lee, A modified flux regulation method to minimize switching frequency and improve DTC-hysteresis-based induction machines in low-speed regions, IEEE J. Emerg. Sel. Topics Power Electron. 7 (2019), no. 4, 2346–2355.

[12] J. Wu, J. Zhang, B. Nie, Y. Liu, and X. He, "Adaptive Control of PMSM Servo System for Steering-by-Wire System With Disturbances Observation," IEEE Transactions on Transportation Electrification, vol. 8, no. 2, pp. 2015–2028, Jun. 2022, doi: 10.1109/TTE.2021.3128429.

[13] A. E. Rabhi, K. Hartani, and Y. Guettaf, "Direct torque control improvement for PMSM drive based on FLC," 2022. [Online]. Available: https://www.scienceliterature.com

[14] O. Ouledali, A. Meroufel, P. Wira, and S. Bentouba, "Genetic Algorithm Tuned PI Controller on PMSM Direct Torque Control," Algerian Journal of Renewable Energy and Sustainable Development, vol. 01, no. 02, pp. 204–211, Dec. 2019, doi: 10.46657/ajresd.2019.1.2.10.

[15] O. M. Arafa, S. A. Wahsh, M. Badr, and A. Yassin, "Grey wolf optimizer algorithm based real time implementation of PIDDTC and FDTC of PMSM," International Journal of Power Electronics and Drive Systems, vol. 11, no. 3, pp. 1640–1652, 2020, doi: 10.11591/ijpeds.v11.i3.pp1640-1652.

[16] H. Mesloub, M. T. Benchouia, R. Boumaaraf, A. Goléa, N. Goléa, and M. Becherif, "Design and implementation of DTC based on AFLC and PSO of a PMSM," Math. Comput. Simul., vol. 167, pp. 340–355, Jan. 2020, doi: 10.1016/j.matcom.2018.04.010.

[17] B. Çavuş and M. Aktaş, "International Journal of Automotive Engineering and Technologies Direct torque control of permanent magnet synchronous motor for electric vehicles." [Online]. Available: https://dergipark.org.tr/en/pub/ijaet

[18] Y. Ma, D. Li, Y. Li, and L. Yang, "A Novel Discrete Compound Integral Terminal Sliding Mode Control with Disturbance Compensation for PMSM Speed System," IEEE/ASME Transactions on Mechatronics, vol. 27, no. 1, pp. 549–560, Feb. 2022, doi: 10.1109/TMECH.2021.3068192.

[19] A. Pal, P. Roy, and S. Ray, "Sliding Mode Speed Control of a Permanent Magnet Synchronous Motor using Twelve Sectors Modified DTC," in 2022 IEEE Calcutta Conference, CALCON 2022 - Proceedings, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 128–132. doi: 10.1109/CALCON56258.2022.10060752.

[20] Z. Li, F. Wang, D. Ke, J. Li, and W. Zhang, "Robust Continuous Model Predictive Speed and Current Control for PMSM with Adaptive Integral Sliding-Mode Approach," IEEE Trans. Power Electron., vol. 36, no. 12, pp. 14398–14408, Dec. 2021, doi: 10.1109/TPEL.2021.3086636.

[21] F. M. Zaihidee, S. Mekhilef, and M. Mubin, “Robust speed control of pmsm using sliding mode control (smc)-a review,” 2019, MDPI AG. doi: 10.3390/en12091669.
[22] Li, Y., & Zhang, Z. (2024). High-order nonsingular fast integral terminal sliding mode control with perturbation estimation for a class of nonlinear hysteresis systems. IEEE Transactions on Industrial Electronics, 72(2), 2045-2055.
[23] Mahjoub S, Mnif F, Derbel N. Second-order sliding mode approaches for the control of a class of underactuated systems. Int J Autom Comput 2015;12:134–41.
[24] Ao, D., Huang, W., Wong, P. K., & Li, J. (2021). Robust backstepping super-twisting sliding mode control for autonomous vehicle path following. IEEE Access, 9, 123165-123177.
[25] Pan, H., Zhang, G., Ouyang, H., & Mei, L. (2020). A novel global fast terminal sliding mode control scheme for second-order systems. IEEE Access, 8, 22758-22769.
[26] Matawong, S., Audomsi, S., Photong, C., Thongsan, T., & Sa-Ngiamvibool, W. (2026). Enhanced Speed Control of PMSM Using Sliding Mode Controller Optimized by Chess Optimization Algorithm. Applied Sciences, 16(14), 7143.
[27] Yuan, L., Chen, H., Wang, P., Liu, Y., & Xu, A. (2023). Design of single neuron super-twisting sliding mode controller for permanent magnet synchronous servo motor. Frontiers in Energy Research, 11, 1130646.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Osman Nuri Ergin as a Bureaucrat of Transition and the Rationalization of Local Government in Turkey

Mustafa Demirkol[1,*]

[1]Faculty of Economics and Administrative Sciences, Tekirdağ Namık Kemal University, Tekirdağ, Türkiye
ORCID: 0000-0002-9860-8836

[*](mustafademirkol@nku.edu.tr) Corresponding author

Abstract – This study analyzes the institutionalization process of the Turkish local government system through the lens of Osman Nuri Ergin (1883–1961), a pivotal technocrat in the transition from the Ottoman Empire to the Republic of Türkiye. The aim is to reveal Ergin's role in constructing modern Turkish municipalism by systematizing administrative knowledge and rationalizing urban space. Adopting a qualitative approach based on historical analysis and document review, the study employs content analysis to examine Ergin's primary works — chiefly Mecelle-i Umûr-ı Belediyye and İstanbul Şehri Rehberi — together with the administrative legislation of the period, interpreting them within a theoretical framework combining historical institutionalism, governmentality, and the "legible city" thesis. Ergin's career is evaluated through Maurice Hauriou's concept of administrative reason, James C. Scott's notion of legibility, Michel Foucault's governmentality, Pierre Bourdieu's cultural capital, and Paul Pierson's path dependency and increasing returns. The central argument is that Ergin followed a "continuity within change" paradigm rather than a radical break, integrating Ottoman administrative heritage — particularly the waqf and imaret systems — into the rational bureaucracy of the new nation-state. The study finds that by constructing administrative reason through Mecelle-i Umûr-ı Belediyye and by rendering urban space "legible" through the street-naming, house-numbering, and city-guide projects carried out for the 1927 census, Ergin acted as one of the founding architects of the Turkish municipal tradition. The study contributes to the international literature on administrative modernization by challenging the "radical break" narrative of Turkish modernization and by showing that traditional institutions were functionally converted into, rather than abolished by, Republican-era social municipalism.

Keywords – Osman Nuri Ergin; Local Government; Continuity within Change; Path Dependency; Governmentality; Historical Institutionalism

I. INTRODUCTION

The last century of the Ottoman Empire and the founding years of the Republic of Türkiye were difficult periods for Turkish public administration: a time of structural transformation that still had to carry forward a deep administrative heritage. It is not only about administration; it maintained his institutional memory along with the continuity of the bureaucratic tradition. In this case, Osman Nuri Ergin (1883–1961) emerges as an important person. Starting from 1901, for 46 years in a row, he supported the Istanbul Şehremaneti (municipal) bureaucracy [1], [2] during the reign of Abdul Hamid II until the Republic era, passing through the period of the Second Constitutional Era and the Armistice. Not many people had such an experience as Ergin, as his career took place right at the heart of cultural transformation that Turkey experienced. In effect, he was the link between the centuries of accumulated Ottoman administrative expertise and the modernizing visions of the Republic [3].

The study argues that what Ergin did while creating Turkish public administration is not the creation of any revolutionary new thing; on the contrary, he simply enhanced certain things about the origin of public administration. It could be said that his approach to creating public administration was one of "persistence amid change." In fact, Ergin did not take the emergence of a Western-type municipal system in the Ottoman Empire as something alien; he simply reinterpreted the waqf, imaret, and guild system of the Islamic city as the historical equivalents of modern municipalism – which reminds us of Pierson's concept of "increasing returns.". Ergin treated mosques, dervish lodges, and imarets as more than religious buildings. He read them as "cité"-like institutions with worldly, practical functions in urban planning and social services, a reading that both gave the new regime's secularization policies a measure of legitimacy and kept these older institutions alive by "updating" their role.

Accordingly, the study analyzes Ergin's work along two interrelated axes: the codification of local-government knowledge and the regulation of urban space. Through monumental works such as Mecelle-i Umûr-ı Belediyye, Ergin recorded the scattered municipal legislation and practices of his time, effecting a transition from an oral culture to a standardized bureaucratic one, in a manner paralleling Hauriou's ideal of administrative rationalization [5]. Through the 1927 census and the associated street-naming and numbering project, he disciplined and nationalized Istanbul's physical space in what Scott [6] terms the "legible city," rendering the city a manageable administrative database; in Foucauldian terms, this consolidated the knowledge-based power of modern bureaucracy [7].

II. Materials and Method

This study uses a qualitative design built on historical analysis and document review. The primary sources are Ergin's own foundational works, chiefly Mecelle-i Umûr-ı Belediyye, İstanbul Şehri Rehberi [8], Türk Şehirlerinde İmaret Sistemi [9], and Türk İmar Tarihinde Vakıflar, Belediyeler, Patrikhaneler [10], read alongside archival documents and the administrative legislation of the period, notably Law No. 1003 of 1927 on building numbering and street naming and Municipal Law No. 1580 of 1930. Content analysis was applied to these texts. The interpretation draws on historical institutionalism and path dependency [4], Hauriou's account of administrative reason [5], Foucault's governmentality [7], Scott's concept of the "legible city" [6], and Bourdieu's cultural capital [11]. Secondary biographical and historiographical studies of Ergin [1], [2], [3], [12], [13], [14] round out and check the analysis.

III. Results

A. Codification: Mecelle-i Umûr-ı Belediyye

Ergin's magnum opus, Mecelle-i Umûr-ı Belediyye, gathered the scattered decrees, orders, regulations, and local customs of Ottoman municipal administration into a coherent, five-volume corpus (nine volumes in its Latinized edition). Volume I, though written last (1922), was placed first for its theoretical framing, comparing Islamic and Western municipal models; Volumes II–V (1914–1919) classified municipal laws, concession contracts, health regulations, personnel regimes, and administrative decisions [13], [15]. The work served as both a foundational reference and a reservoir of institutional memory for Republican municipal governance, particularly during the drafting of the 1930 Municipal Law No. 1580. Following Hauriou's institutionalist logic, this codification endowed the municipality with a rational administrative reason that transcended personal initiative, comparable to the codifying function of the French Conseil d'État tradition [5].

B. Rationalization of Urban Space: Street Naming and the City Guide

Ergin's second major contribution concerns the rationalization of urban space. As Director of Numbering under the 1927 Law No. 1003 on building numbering and street naming, Ergin led a five-month survey that renamed or newly named 6,215 of Istanbul's roughly 10,000 streets, subsequently producing the 38-map İstanbul Şehri Rehberi [2], [8]. This operation exemplifies Scott's [6] "legible city": the city's oral and locally particular spatial references were replaced by cartographic mapping and standardized addressing, strengthening the state's capacity for taxation, conscription, policing, and civil registration [14]. In Foucauldian terms, the resulting guide functioned not merely as a wayfinding tool but

as an administrative database enabling the land-registry, census, and police authorities to query and intervene in urban space.

The renaming project also carried a toponymic, nation-building dimension, pursued through three strategies: eliminating duplicate street names, removing names deemed comic or indecent, and — most saliently — nationalizing toponymy by replacing non-Muslim, Greek, or Armenian place names with figures and concepts drawn from Turkish history, for instance “Papas Mihal” Street becoming “Dede Korkut” and “Rus Sokağı” becoming “Türk Beyi” [14].

C. *A Hybrid Technocratic Identity*

Ergin rose from a hybrid self. At Darüşşafaka and the Darülfûnun he became a modern (“Mektepli”) recruit, but on his own went for traditional (“Medreseli”), self-taught in Islamic jurisprudence, Arabic, and the waqf system under Çarşambalı Hacı Ahmed Efendi's direction [12], [1]. This dual formation granted Ergin a symbolic capital read through the lens of Bourdieu's [11] cultural capital; he became a person without whom no technocratic ruling circle in the Hamidian, Constitutional, and Republican periods could proceed. That continuity did not rest on political allegiance to any of the three regimes. It relied on the de facto monopoly he had over administrative knowledge and archival knowledge.

IV. Discussion

These results support the main argument of the study: Ergin's professional life conforms to a paradigm articulated by continuity within change, not the radical break commonly associated with Turkish modernization. Ergin did not see 1855 or the broader Westernization project as a starting-point of zero. Municipal-type services had already followed a centuries-old pathway under the leadership of the “ferdiyet” (individual-initiative) system and the waqf institution; the author claimed that “the task with which he was charged is to carry out this historical path modernly [10]. This reduced the constitutive cost of creating the municipal institution and in stylistic terms, imported a certain quantity of social capital into the new bureaucracy as Pierson [4] puts it, which in turn strengthened its legitimacy.

This reading is in line with Streeck and Thelen's understanding of institutional change as conversion rather than replacement [16], wherein existing structures are redeployed toward new ends rather than demolished entirely. The Islamic İmaret and the Western Cité served as a bundling of urban cores with religious, welfare and administrative functions, and this equation [9], [12] is a clear exemplification of such conversion that however, renders something -the common good- out. His work belongs within the European codification tradition of administrative law at least as it developed in Napoleonic France and, later, Hauriou’s droit administratif[5], but it does not belong only to that tradition. Western codifications usually compiled texts of laws and judicial decisions. Ergin’s Mecelle did that too, but it also built an encyclopedic comparison of Islamic and Western urban institutions, in a method he himself called “local, national, and comparative” [12].

The technocratic nature of Ergin’s durability spanning three different political regimes suggests something broader: that the presence of above-politics expertise can shore up a transitioning regime. He was always a reluctant ally of the Westernizing or the conservative camp, and he had something like a monopoly on the technical know-how needed to run Istanbul. This combination made him (in the terms of this study) an indispensable bureaucrat [1] This status, in turn, is articulated by Bourdieu’s [11] cultural-capital framework as the permanent transformation of hybridised capital into technocratic legitimacy.

Collectively, such findings inform the broader international scholarship of administrative modernization in ex-imperial contexts. The "radical break" story that has characterized the Turkish experience of state formation undervalues the extent to which the structures of the Republic were constructed upon the administrative legacy of the Ottomans, rather than in opposition to it. In this sense, the Ergin case serves as an example that could apply elsewhere: for societies transitioning from empire to nation-state, institutional memory must often be managed.

V. Conclusion

If this study had to be reduced to one sentence, it would be that Osman Nuri Ergin was the architect of modern Turkish municipalism. Positioned at the historical rupture between the Ottoman Empire and the Republic of Türkiye, he functioned not merely as a bureaucrat or historian but as the builder of the

institutional and intellectual infrastructure of modern Turkish local government, pouring the administrative memory of a collapsing empire into the rational forms of the new nation-state and thereby securing a vital continuity between the two. Three lasting legacies of this contribution can be identified. First, Mecelle-i Umûr-ı Belediyye — sometimes called the "constitution" of Turkish municipalism — remains today a primary reference for understanding the institutional origins of Turkish local government. Second, Ergin's street-naming and building-numbering work, and the İstanbul Şehri Rehberi it produced, prefigure Turkey's modern address-registration and urban-information systems. Third, his redefinition of the waqf and imaret institutions as forerunners of municipal social services supplies indigenous, historical reference points for the concept of "social municipalism" still debated in Turkish public administration today.

Beyond its national significance, the Ergin case contributes to the international literature on administrative modernization in societies undergoing an imperial-to-national transition. His biography and body of work reflect an approach grounded in "continuity within change" rather than rupture: rather than abolishing local institutions such as the waqf and imaret, he advocated their functional conversion within modern institutions, while treating the preservation of institutional memory during regime change as a central administrative responsibility. In this sense, Osman Nuri Ergin may properly be described as institutional memory made flesh.

ACKNOWLEDGMENT

The author wishes to thank everyone whose support contributed to his education.

REFERENCES

[1] Sayar, A. G. (1995). Ergin, Osman Nuri. *Türkiye Diyanet Vakfı İslâm Ansiklopedisi* (Vol. 11, pp. 297–298). İstanbul: TDV Yayınları.
[2] Yılmaz, M. (2013). Osman Nuri Ergin: Hayatı, Şahsiyeti ve Eserleri (1883–1961). *Türkiyat Mecmuası*, 23(1), 141–188.
[3] Ünver, A. S. (1962). Osman Ergin Çalışma Hayatı ve Eserleri: 1883–1961. *Belleten*, 26(101), 163–178.
[4] Pierson, P. (2000). Increasing returns, path dependence, and the study of politics. *American Political Science Review*, 94(2), 251–267.
[5] Węgliński, C. (2022). Maurice Hauriou's theory of the institution: Legal institutionalism and the science of the state. *Archiwum Filozofii Prawa i Filozofii Społecznej*, 88–102.
[6] Scott, J. C. (1999). *Seeing Like a State.* New Haven, CT: Yale University Press.
[7] Foucault, M. (1991). Governmentality. In G. Burchell, C. Gordon, & P. Miller (Eds.), *The Foucault Effect: Studies in Governmentality* (pp. 87–104). Chicago, IL: University of Chicago Press.
[8] Ergin, O. N. (1934). *İstanbul şehri rehberi.* İstanbul: İstanbul Belediyesi Matbaası.
[9] Ergin, O. N. (1939). *Türk şehirlerinde imaret sistemi.* İstanbul: Cumhuriyet Matbaası.
[10] Ergin, O. N. (1944). *Türk imar tarihinde vakıflar, belediyeler, patrikhaneler.* İstanbul: Türkiye Basımevi.
[11] Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.), *Handbook of Theory and Research for the Sociology of Education* (pp. 241–258). New York: Greenwood.
[12] İmamoğlu, A. T. (2005). Osman Nuri Ergin: Cumhuriyet döneminin ilk şehir tarihçisi. *Türkiye Araştırmaları Literatür Dergisi*, 3(6), 553–570.
[13] Özcan, A. (2003). Mecelle-i Umûr-ı Belediyye. *Türkiye Diyanet Vakfı İslâm Ansiklopedisi* (Vol. 28, pp. 235–236). Ankara: TDV Yayınları.
[14] Salman, C. (2022). Kent hafızasında bir kayıt ve kayıp belgesi: Osman Nuri Ergin'in İstanbul Şehri Rehberi. *Çağdaş Yerel Yönetimler*, 31(4), 67–84.
[15] Oktay, T. (2005). Osmanlı belediye tarihi araştırmaları ve kaynak malzeme üzerine notlar. *Türkiye Araştırmaları Literatür Dergisi*, 3(6), 257–281.
[16] Streeck, W., & Thelen, K. (Eds.). (2005). *Beyond Continuity: Institutional Change in Advanced Political Economies.* Oxford: Oxford University Press.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Mardin'de Özel Mülkiyete Ait Bir Bina Örneği Üzerinden Konutta Müşterek Alanların Tanımı Ve Hesabı

Rojat AKSOY IŞIK[1*], Cuvan YILDIRIM [2]

[1] *Mimarlık Bölümü, Mühendislik-Mimarlık Fakültesi, Mardin Artuklu Üniversitesi, Türkiye*
[2]*Mimarlık ABD, Lisansüstü Eğitim Enstitüsü, Mardin Artuklu Üniversitesi, Türkiye*

[*](*rojataksoy@artuklu.edu.tr)*

Özet **–** Konut yapılarında müşterek alanlar, yapıların işlevselliğinin sağlanması, teknik gereksinimlerin karşılanması ve kullanıcıların ortak ihtiyaçlarının giderilmesi açısından önemli bir yere sahiptir. Bu çalışma, Mardin ilinde özel mülkiyete ait bir konut yapısı örneği üzerinden müşterek alanların tanımını, hesaplama yöntemlerini ve mimari planlama sürecindeki rollerini incelemeyi amaçlamaktadır. Araştırma kapsamında literatür taraması yapılmış, ilgili mevzuat hükümleri değerlendirilmiş ve belediye tarafından onaylanmış mimari uygulama projesi analiz edilmiştir. Çalışmada, Kızıltepe ilçesi Alimışmış Mahallesi'nde gerçekleştirilecek 33 bağımsız bölümlü bir konut projesinde yer alan müşterek alanlar; sığınak, hidrofor deposu, enerji ve haberleşme odası, jeneratör odası, asansör, kat holleri, merdivenler ve mekanik tesisat şaftları kapsamında ele alınmıştır. Yapılan inceleme ve hesaplamalar sonucunda, sığınak için 77,34 m², hidrofor deposu için 3,73 m², enerji ve haberleşme odası için 10,26 m², jeneratör odası için 7,66 m², asansör alanı için 5,87 m², kat holleri için toplam 85,29 m² ve merdiven alanı için 10,77 m² büyüklüğünde alan gereksinimi belirlenmiştir. Ayrıca mekanik ve havalandırma sistemlerine ait şaftların asgari 60×60 cm boyutlarında olması gerektiği tespit edilmiştir. Araştırma sonucunda incelenen konut projesindeki müşterek alanların ilgili mevzuat hükümlerine uygun olarak tasarlandığı ve asgari alan gereksinimlerini karşıladığı belirlenmiştir. Elde edilen bulgular, müşterek alanların yalnızca teknik zorunlulukları karşılayan mekânlar olmadığını, aynı zamanda yapıların kullanım kalitesini ve sürdürülebilir işleyişini destekleyen temel bileşenler olduğunu göstermektedir. Bu nedenle mimari tasarım ve projelendirme süreçlerinde müşterek alanların ilgili yönetmelikler doğrultusunda doğru tanımlanması, analiz edilmesi ve yeterli büyüklüklerde planlanması büyük önem taşımaktadır.

Anahtar Kelimeler –Konut yapıları, Müşterek alanlar, Kat mülkiyeti, Mardin, Emsal

I. GİRİŞ

İnşaat sektörü, dünya genelinde olduğu gibi Türkiye ekonomisine de önemli katkılar sağlayan başlıca sektörlerden biridir. Sektörün büyük ölçüde yerli sanayiye dayanması, istihdam yaratma kapasitesi ve imalat sektörüyle olan güçlü ilişkileri, inşaat sektörünü Türkiye ekonomisinin temel ve güçlü sektörlerinden biri hâline getirmektedir (Tanrıvermiş, 2019). Yeni inşaat malzemelerinin ve yapım tekniklerinin geliştirilmesi, sektördeki teknolojik ve uygulamaya yönelik gelişmeleri desteklemekte; bu doğrultuda, tekil yapıların yanı sıra çok sayıda bloktan oluşan toplu konut projeleri gibi büyük ölçekli yapılaşmalar giderek yaygınlaşmaktadır (Kılıç & Hardal, 2019). Büyük ölçekli projelerin planlanması ve uygulanması, önemli miktarda finansal, teknik ve fiziksel kaynak kullanımını ve çok sayıda çalışanın farklı görev ve sorumluluklar üstlenmesini gerektirmektedir. Bu nedenle kaynakların etkin ve verimli kullanılması, proje süreçlerinde görev ve sorumlulukların açık biçimde tanımlanması ve belirlenen sorumlulukların eksiksiz şekilde yerine getirilmesi büyük önem taşımaktadır (Açıkalın, 2019; Şengül, 2012).

Büyük ölçekli yapı projelerinin başarılı bir biçimde tamamlanabilmesi için projelerin doğru, eksiksiz ve uygulanabilir bir biçimde tasarlanması gerekmektedir. Tasarım sürecinin önemli bileşenlerinden biri de yapı kullanıcılarının ortak ve sosyal ihtiyaçlarına cevap veren alanların planlanmasıdır (Yüksel Gürez, 2019). Yapıların ortak kullanımına ayrılan bu alanlar zaman içerisinde gelişmiş ve ilgili mevzuat kapsamında belirlenen yönetmeliklerle çeşitli standartlara bağlanmıştır. Yapı içerisinde yer alan ortak kullanım alanları, kat maliklerinin ortak kullanımına sunulan ve bireysel mülkiyete konu olmayan müşterek alanlar olarak değerlendirilmektedir. Bu kapsamda sığınak, kazan dairesi, asansör, çatı, merdiven boşlukları, avlular, kapıcı dairesi, kömürlük, bacalar, taşıyıcı sistem elemanları, su depoları, elektrik ve haberleşme odaları gibi çeşitli birimler müşterek alanlar içerisinde değerlendirilmektedir (Cemiloğlu, 2021; Hilmioğlu & İşanç, 2022).

Özellikle konut projelerinde kullanıcıların ortak ihtiyaçlarını karşılayan müşterek alanlar; yapının işleyişini destekleyen teknik birimler, sosyal etkileşimi sağlayan mekânlar ve düşey-yatay sirkülasyonu sağlayan alanlar gibi farklı işlevleri bünyesinde barındırmaktadır. Söz konusu alanların doğru biçimde planlanması, kullanıcı konforunun sağlanmasının yanı sıra yapının işlevselliği ve sürdürülebilirliği açısından da önemli bir rol üstlenmektedir (Alsibaai, 2022; Yenice, 2019).

Bu çalışma, Kızıltepe ilçesi Alimışmış Mahallesi'nde gerçekleştirilecek bir konut projesi kapsamında yapı içerisindeki müşterek alanların planlanması, hesaplanması ve mekânsal ilişkilerinin değerlendirilmesine odaklanmaktadır. Çalışma alanının çevresinde yer alan toplu konut yapıları belirli bir kullanıcı nüfusuna hizmet etmekte olup, bölgedeki konut yapıları genel olarak zemin kat üzerine dört veya beş katlı yapılardan oluşmaktadır. Alimışmış Mahallesi'nde ele alınan konut projesinin uygulama sürecine geçilmeden önce, arazinin mevcut durumunun belirlenmesi amacıyla harita ve kadastro birimlerinden aplikasyon krokisinin temin edilmesi ve emsal hesabının yapılabilmesi için ilgili belediyeden imar durum belgesinin alınması gerekmektedir. Mimari projenin emsal hesabı belirlendikten sonra aplikasyon krokisi üzerinden parsel sınırları tanımlanmış ve uygulama projesinin hazırlanmasına yönelik gerekli veriler elde edilmiştir.

Çalışma kapsamında ele alınan konut projesinde toplam 33 bağımsız bölümden oluşan ve 1+1, 2+1 ve 3+1 daire tiplerini içeren karma bir planlama yaklaşımı benimsenmiştir. Yapının işlevsel ve teknik açıdan gereksinimlerini karşılayabilmesi amacıyla gerekli müşterek alanlar belirlenmiştir. Bu kapsamda sığınak, su deposu, enerji ve haberleşme odası, mekanik ve havalandırma şaftları, asansör, kat holü ve ortak kat merdiveni gibi birimler projede yer almaktadır. Söz konusu birimlerin büyüklükleri ve mekânsal gereklilikleri, ilgili mevzuat ve yönetmeliklerde belirtilen hesaplama yöntemleri doğrultusunda belirlenmiştir.

Sığınak alanları, ilgili yönetmeliklerde bağımsız bölümlerin oda sayılarına bağlı olarak belirlenen minimum alan koşulları doğrultusunda tasarlanmıştır. Buna göre bir odalı bağımsız bölümler için 2 m², iki odalı bağımsız bölümler için 3 m², üç ve daha fazla odalı bağımsız bölümler için ise her bağımsız bölüm başına 4 m² sığınak alanı ayrılması gerekmektedir. Bununla birlikte sığınakların tasarımında sığınak odasının yanı sıra havalandırma odası, mutfak nişi, kadın ve erkek WC ile duş birimlerinin de çözümleme içerisinde değerlendirilmesi gerekmektedir. Ayrıca sığınak girişinin birbirine dik

konumlandırılmış iki kapıdan oluşması öngörülmektedir (Sığınak Yönetmeliği, 1988). Su deposu (hidrofor deposu), yapının çeşitli nedenlerle şebeke suyuna erişiminin kesintiye uğraması durumunda kullanıcıların su ihtiyacının karşılanmasına yönelik olarak tasarlanan teknik müşterek alanlardan biridir. Bu nedenle su deposunun, yapının ihtiyacını karşılayabilecek yeterli kapasitedeki su depolama birimlerini barındırabilecek büyüklükte planlanması gerekmektedir. Alan hesabı, ilgili mevzuatta belirtilen gereklilikler doğrultusunda gerçekleştirilmiştir (Su Yalıtım Yönetmeliği, 2017). Enerji ve haberleşme odası; binaya ait elektrik tesisatı, anten ve internet kablolarının düzenli bir biçimde bir araya getirilmesini, gerekli durumlarda bu sistemlere müdahale edilmesini ve kontrollerinin gerçekleştirilmesini sağlayan teknik birimdir. Yapı içerisindeki elektrik sistemlerinin ilgili standart ve yönetmelikler doğrultusunda düzenlenmesi ve gerekli panolarda toplanması gerekmektedir (Elektrik İç Tesisleri Proje Hazırlama Yönetmeliği, 2003).

Mekanik ve havalandırma şaftları ise yapı içerisinde tesisat sistemlerinin dağıtımını ve hava sirkülasyonunun sağlanmasını mümkün kılan teknik birimlerdir. Bu kapsamda havalandırma şaftı, mekanik şaft, elektrik şaftı ve sığınak havalandırma şaftı gibi farklı birimler yapı içerisindeki teknik gereksinimlere göre düzenlenmektedir (Mekanik Tesisat Fenni Mesuliyet Hizmetleri Yönetmeliği, 2012). Jeneratör odası ise elektrik enerjisinin kesintiye uğradığı durumlarda yapının belirli bir süre boyunca ihtiyaç duyduğu enerjinin sağlanmasına yönelik olarak tasarlanan teknik birimdir. Bu kapsamda yapının gereksinimlerini karşılayabilecek kapasitede bir jeneratörün yerleştirilebilmesi için yeterli alanın ayrılması gerekmektedir (Elektrik İç Tesisleri Proje Hazırlama Yönetmeliği, 2003). Yapı içerisindeki düşey sirkülasyonu sağlayan temel müşterek alanlardan biri asansördür. Asansör, farklı katlar arasında kullanıcıların erişimini kolaylaştıran ve yapının erişilebilirliğini artıran önemli bir ulaşım birimidir. Asansör gereksinimi ve sayısına ilişkin hesaplamalar, yapının bağımsız bölüm sayısı ve bina yüksekliği gibi çeşitli parametreler dikkate alınarak ilgili mevzuat doğrultusunda yapılmaktadır (Asansör Yönetmeliği, 2014). Kat holü, bağımsız bölümlere erişimi sağlayan ve yapı kullanıcılarının ortak kullanımına ayrılmış önemli bir sirkülasyon alanıdır. Yapının girişinden bağımsız bölümlere geçişi sağlayan kat hollerinin, ilgili mevzuatta belirlenen minimum genişlik koşullarını karşılaması gerekmektedir. Planlı Alanlar İmar Yönetmeliği'ne göre kat holünün dar kenarının 150 cm'den az olmaması gerekmektedir (Planlı Alanlar İmar Yönetmeliği, 2017). Ortak kat merdiveni de yapı içerisindeki düşey sirkülasyonu sağlayan temel müşterek alanlardan biridir. Merdiven tasarımında basamak sayısı, kat yüksekliğinin rıht yüksekliğine bölünmesiyle belirlenmekte; basamak ölçüleri ve merdiven genişliği ise ilgili mevzuatta tanımlanan kriterlere göre düzenlenmektedir. Planlı Alanlar İmar Yönetmeliği'nde ortak merdivenlerin dar kenarının 120 cm'den az olmaması gerektiği belirtilmektedir (Planlı Alanlar İmar Yönetmeliği, 2017). Belirtilen hesaplama ve tasarım kriterleri doğrultusunda mimari uygulama projesine ilişkin çizimler hazırlanmıştır. İlgili mevzuat, müşterek alanların gerekli minimum büyüklüklerinin belirlenmesinde ve mekânsal organizasyonlarının oluşturulmasında yönlendirici bir rol üstlenmektedir. Bu doğrultuda hazırlanan mimari projelerin, belediye onay sürecinde ilgili mimari uygulama ve teknik çizim standartlarına uygun olarak hazırlanması gerekmektedir. Standartlaştırılmış çizim dili, projenin farklı meslek disiplinleri tarafından doğru ve anlaşılır biçimde okunabilmesini sağlamakta ve tasarımın uygulama aşamasında ortaya çıkabilecek koordinasyon sorunlarının azaltılmasına katkıda bulunmaktadır.

II. MATERYAL VE YÖNTEM

Bu çalışmada, konut yapılarında müşterek alanların yapı içerisindeki mekânsal organizasyonu, birbirleriyle olan ilişkileri, sirkülasyon sistemi içerisindeki konumları ve kullanıcı hareketleri açısından işlevsel yeterlilikleri incelenmiştir. Ayrıca müşterek alanların gerekli minimum büyüklüklerinin ilgili mevzuat hükümleri doğrultusunda belirlenmesi ve bu alanların mimari çizimler aracılığıyla çözümlenmesi amaçlanmıştır. Çalışmanın yöntemsel süreci üç temel aşamadan oluşmaktadır. İlk aşamada müşterek alanlar konusunda kapsamlı bir literatür taraması gerçekleştirilmiştir. Bu kapsamda konut projelerinde ortak kullanım alanları, mekânsal organizasyon, kullanıcı gereksinimleri ve mekânsal ölçeklendirme konularında yayımlanmış akademik çalışmalar incelenmiştir. Bunun yanı sıra, belediyeler

tarafından onaylanmış çeşitli konut projeleri değerlendirilerek müşterek alanların projelendirilme biçimleri ve mekânsal organizasyonları karşılaştırmalı olarak ele alınmıştır. Çalışmanın ikinci aşamasında, müşterek alanların hukuki, işlevsel ve boyutsal gerekliliklerinin belirlenmesi amacıyla ilgili mevzuat ve yönetmelikler incelenmiştir. Bu kapsamda kat mülkiyeti ve kat irtifakı ile konut yapılarının tasarımını doğrudan veya dolaylı olarak ilgilendiren yönetmelikler değerlendirilmiştir. İlgili mevzuat kapsamında müşterek alanların kullanım esasları, minimum alan gereksinimleri, sirkülasyon ölçütleri ve teknik birimlere ilişkin hesaplama yöntemleri belirlenmiştir.

Çalışmanın üçüncü aşamasında ise uygulamaya yönelik plan çözümleme analizleri gerçekleştirilmiştir. Bu doğrultuda seçilen konut uygulama projesi üzerinden müşterek alanların mekânsal organizasyonu, büyüklükleri, dağılımları, birbirleriyle olan ilişkileri ve işlevsel yeterlilikleri değerlendirilmiştir. Projeye ait mimari çizimler, bilgisayar destekli tasarım programı AutoCAD kullanılarak analiz edilmiş; müşterek alanlara ilişkin gerekli alan ve boyut hesaplamaları gerçekleştirilmiştir. Elde edilen veriler, ilgili yönetmeliklerde belirtilen minimum ölçütler ile karşılaştırılarak değerlendirilmiştir.

Analizler sonucunda, konut projelerinde müşterek alanların planlanması ve hesaplanmasına ilişkin mevcut uygulamalar değerlendirilmiş; mevzuat tarafından belirlenen minimum koşullar ile mimari tasarım gereklilikleri arasındaki ilişki ortaya konulmuştur. Bu kapsamda, müşterek alanların yalnızca mevzuatta belirtilen minimum ölçütleri karşılamasının yeterli olup olmadığı; aynı zamanda kullanıcı hareketleri, mekânsal ilişkiler, işlevsellik ve yapı içerisindeki sirkülasyon açısından yeterli bir çözüm sunup sunmadığı da incelenmiştir. Elde edilen bulgular doğrultusunda müşterek alanların planlanması ve hesaplanmasına ilişkin mevcut sorunlar belirlenmiş ve bu sorunların giderilmesine yönelik öneriler geliştirilmiştir.

III. BULGULAR

Bu bölümde, Kızıltepe ilçesi Alimışmış Mahallesi'nde incelenen konut projesine ilişkin arazi verileri, emsal ve taban alanı hesaplamaları ile yapı içerisinde yer alan müşterek alanlara yönelik bulgular sunulmuştur. Çalışma kapsamında müşterek alanların yalnızca mevzuatta belirtilen minimum koşulları sağlayıp sağlamadığı değil, aynı zamanda yapı içerisindeki işlevleri, kullanıcı sirkülasyonu, erişilebilirlik ve teknik kullanım gereksinimleri de değerlendirilmiştir. Çalışma alanına ilişkin temel veriler, aplikasyon krokisi ve imar durum belgesi üzerinden belirlenmiştir. Aplikasyon krokisi incelenerek parsel sınırları, parsel boyutları ve yapı yaklaşma mesafeleri tespit edilmiştir. Bu veriler, mimari uygulama projesinin parsel üzerinde hangi sınırlar içerisinde geliştirileceğinin belirlenmesinde temel oluşturmuştur.

Parselin toplam alanı 1.484,20 m² olarak belirlenmiştir. İmar durum belgesinde parsel için TAKS değeri 0,40, KAKS/Emsal değeri ise 2,00 olarak belirtilmiştir. Bu değerler kullanılarak parselin yapılaşma kapasitesi hesaplanmıştır. Taban alanı, parsel alanının TAKS değeri ile çarpılması sonucunda 1.484,20 m² × 0,40 = 593,68 m² olarak bulunmuştur. Buna göre parsel üzerinde oluşturulabilecek maksimum taban alanı 593,68 m²'dir. Toplam emsal alanı ise 1.484,20 m² × 2,00 = 2.968,40 m² olarak hesaplanmıştır. Çalışmada ayrıca Kızıltepe Belediyesi tarafından belirtilen %30 emsal dışı alan oranı esas alınmış ve 2.968,40 m² × 0,30 = 890,52 m² hesaplaması sonucunda emsal dışında değerlendirilebilecek alan 890,52 m² olarak belirlenmiştir. Elde edilen veriler doğrultusunda iki bloktan oluşan konut projesinin yapılaşma sınırları oluşturulmuş ve proje kararlarının belirlenen emsal sınırları içerisinde kalması sağlanmıştır.

Arazi ve Emsal Verisi	Değer
Parsel alanı	1.484,20 m²
TAKS	0,40
Maksimum taban alanı	593,68 m²
KAKS / Emsal	2,00
Toplam emsal alanı	2.968,40 m²
Emsal dışı alan oranı	%30
Emsal dışı alan	890,52 m²

Mimari projenin bağımsız bölüm organizasyonu incelendiğinde, yapıda farklı büyüklüklerde toplam 33 bağımsız bölüm bulunduğu belirlenmiştir. Projede 1+1, 2+1 ve 3+1 konut tipleri bir arada kullanılmıştır. Bağımsız bölümlerin oda sayıları ve alanları bağımsız bölüm tablosuna aktarılmış ve bu veriler, toplam yapı alanı ile emsal alanı arasındaki ilişkinin değerlendirilmesinde kullanılmıştır. Bağımsız bölüm sayısı ve oda dağılımı, özellikle kullanıcı sayısına bağlı olarak belirlenen sığınak alanının hesaplanmasında temel veri olarak değerlendirilmiştir. Projenin mimari planları incelendiğinde başlıca müşterek alanların sığınak, su deposu, enerji ve haberleşme odası, mekanik ve havalandırma şaftları, jeneratör odası, asansör, kat holü ve ortak kat merdiveni olduğu belirlenmiştir. Bu alanlar, yapı kullanıcılarının ortak ihtiyaçlarının karşılanması, teknik sistemlerin çalıştırılması ve yapı içerisindeki yatay ve düşey sirkülasyonun sağlanması açısından önemli işlevlere sahiptir.Sığınak alanının belirlenmesinde bağımsız bölümlerin oda sayıları esas alınmıştır. Projede sığınak hesabına esas alınan bağımsız bölüm dağılımında bir adet bir odalı ve 17 adet iki odalı bağımsız bölüm bulunduğu belirlenmiştir. Yapılan hesaplama sonucunda sığınak için gerekli minimum alan 71,00 m² olarak bulunmuştur. Mimari uygulama projesinde sığınak için ayrılan net alanın ise 77,34 m² olduğu tespit edilmiştir. Böylece projede ayrılan sığınak alanının hesaplanan minimum gereksinimden 6,34 m² daha fazla olduğu belirlenmiştir. Sığınak içerisindeki yardımcı işlevler ve kullanıcı hareketleri de dikkate alındığında, mevcut alanın minimum gereksinimi karşılamanın yanında işlevsel kullanım için de yeterli olduğu görülmüştür.Yapının teknik ihtiyaçları kapsamında su deposu ve hidrofor alanı da incelenmiştir. Projede su deposu için 3,73 m² alan ayrıldığı belirlenmiştir. Planlı Alanlar İmar Yönetmeliği'nde su deposu için doğrudan bir minimum metrekare koşulu bulunmaması nedeniyle çalışma kapsamında Kızıltepe Belediyesi tarafından belirtilen yaklaşık 3,00 m²'lik değer esas alınmıştır. Projede ayrılan 3,73 m² alanın bu değerin üzerinde olduğu görülmektedir. Su deposunun büyüklüğü belirlenirken teknik ekipmanların yerleşimi, bakım işlemleri ve personelin erişimi de dikkate alınmıştır. Bu nedenle mevcut çözümün yapının teknik ihtiyaçlarını karşılayacak şekilde düzenlendiği değerlendirilmiştir. Enerji ve haberleşme sistemleri projede aynı mekân içerisinde çözülmüş, ancak iki sistem için ayrı panolar kullanılmıştır. Enerji tesisatına ait kablolar ile haberleşme sistemlerine ait kabloların ayrı panolarda toplanması sağlanmıştır. Bu amaçla toplam 10,26 m² alan ayrılmıştır. Pano yerleşimleri ve bakım sırasında gerekli olacak personel erişim mesafeleri dikkate alınarak yapılan düzenlemenin teknik kullanım açısından yeterli olduğu görülmüştür. Böylece iki farklı teknik işlevin aynı mekân içerisinde, birbirinden bağımsız pano sistemleri kullanılarak çözümlendiği belirlenmiştir.Yapının mekanik, elektrik ve havalandırma tesisatlarının düzenlenmesi amacıyla teknik şaftlar oluşturulmuştur. Özellikle banyo ve WC gibi ıslak hacimlerin havalandırılması için gerekli şaftların plan içerisinde uygun şekilde düzenlendiği görülmüştür. Projede ilgili teknik şaftlar için 60 × 60 cm minimum kesit ölçüsü esas alınmıştır. Şaftların bakım ve teknik müdahaleye uygun olması amacıyla erişim kapakları kat holleriyle ilişkilendirilmiştir. Bu düzenleme sayesinde tesisatların yapı içerisinde düzenli biçimde dağıtılması ve gerektiğinde bakım veya müdahale yapılabilmesi için gerekli koşulların sağlandığı belirlenmiştir.Elektrik enerjisinin kesilmesi durumunda yapının temel ihtiyaçlarının karşılanabilmesi amacıyla projede jeneratör odasına yer verilmiştir. Jeneratör odasının büyüklüğü, elektrik projesinde belirlenen jeneratörün boyutları ve çevresinde bırakılması gereken teknik erişim alanları dikkate alınarak belirlenmiştir. Projede jeneratör odası için 7,66 m² alan ayrılmıştır. Ayrılan alanın jeneratörün yerleşimi, bakım işlemleri ve teknik müdahale için gerekli koşulları karşılayacak şekilde düzenlendiği görülmüştür.İncelenen yapı zemin kat ve beş normal kat olmak üzere

toplam altı seviyeden oluşmaktadır. Bina yüksekliği ve bağımsız bölüm sayısı dikkate alınarak yapıda asansör çözümü oluşturulmuştur. Mimari ve asansör projelerinde asansör için ayrılan alan 5,87 m² olarak belirlenmiştir. Asansörün kat holü ile doğrudan bağlantılı olması, kullanıcıların ortak dolaşım alanlarından bağımsız bölümlere ulaşmasını kolaylaştırmaktadır. Bu nedenle asansörün yapının ortak düşey sirkülasyon sisteminin bir parçası olarak çözümlendiği görülmüştür.Kat holleri, bağımsız bölümlere ulaşımı sağlayan ortak dolaşım alanları olarak değerlendirilmiştir. Projede kat holleri, bağımsız bölüm girişleri ile asansör ve ortak merdiven arasında bağlantı sağlayacak şekilde düzenlenmiştir. Planlı Alanlar İmar Yönetmeliği'nde belirtilen geçiş genişlikleri dikkate alınarak yapılan incelemede, kat hollerinin gerekli geçiş koşullarını sağladığı görülmüştür. Projede kat holleri için toplam 85,29 m² alan ayrılmıştır. Kat hollerinin asansör ve merdivenle doğrudan ilişkilendirilmesi, yapı içerisindeki yatay ve düşey sirkülasyonun bütünleşmesini sağlamıştır.Ortak kat merdiveni de yapının düşey sirkülasyonunu sağlayan temel müşterek alanlardan biridir. Projede kat yüksekliği 310 cm ve basamak sayısı 18 olarak belirlenmiştir. Buna göre rıht yüksekliği 310 cm / 18 = 17,22 cm olarak hesaplanmış ve yaklaşık 17,2 cm olarak belirlenmiştir. Ayrıca merdivenin dar kenarının 120 cm minimum genişlik koşulunu karşılayacak şekilde düzenlendiği görülmüştür. Elde edilen ölçüler, merdivenin yapı içerisindeki düşey sirkülasyon ihtiyacını karşılayacak şekilde tasarlandığını göstermektedir (Tablo 1)

Tablo 1. Projede Yer Alan Müşterek Alanların Özellikleri

Müşterek Alan	Projedeki Alan / Ölçü	Esas Alınan Değer	Değerlendirme
Sığınak	77,34 m²	71,00 m² minimum	Minimum gereksinimin 6,34 m² üzerindedir.
Su deposu	3,73 m²	Yaklaşık 3,00 m²	Belirtilen minimum değerin üzerindedir.
Enerji ve haberleşme odası	10,26 m²	Teknik ekipman ve erişim	İki teknik işlev aynı mekânda ayrı panolarla çözülmüştür.
Mekanik ve havalandırma şaftları	60 × 60 cm	Minimum kesit	Tesisat ve bakım ihtiyaçlarına göre düzenlenmiştir.
Jeneratör odası	7,66 m²	Teknik ekipman ve erişim	Jeneratör yerleşimi ve teknik müdahale için alan sağlanmıştır.
Asansör	5,87 m²	Bina ve sirkülasyon gereksinimi	Kat holü ile doğrudan ilişkilendirilmiştir.
Kat holü	85,29 m²	Gerekli geçiş genişlikleri	Bağımsız bölümlere, asansöre ve merdivene erişim sağlanmıştır.
Ortak kat merdiveni	18 basamak / 17,22 cm rıht	310 cm kat yüksekliği	120 cm minimum dar kenar genişliği sağlanmıştır.

Elde edilen bulgular birlikte değerlendirildiğinde, projede müşterek alanların arazi ve yapılaşma koşulları, bağımsız bölüm sayısı, teknik gereksinimler ve kullanıcı ihtiyaçları ile birlikte ele alındığı görülmektedir. Parsel alanı, TAKS ve KAKS değerleri kullanılarak yapılaşma sınırları belirlenmiş, daha sonra bağımsız bölümlerin ve müşterek alanların organizasyonu gerçekleştirilmiştir. Müşterek alanların belirlenmesinde mevzuatta yer alan ölçütlerin yanı sıra kullanıcı hareketleri, erişilebilirlik, teknik ekipmanların yerleşimi ve bakım gereksinimleri de dikkate alınmıştır. Özellikle sığınak ve su deposu alanlarında belirlenen minimum değerlerin üzerinde alan ayrılması, diğer teknik alanlarda ise kullanım ve erişim koşullarının dikkate alınması, müşterek alanların yalnızca minimum ölçüleri karşılamak amacıyla değil, yapının genel işleyişine katkı sağlayacak şekilde tasarlandığını göstermektedir. Bu doğrultuda, incelenen konut projesinde müşterek alanların mevzuat koşulları ile mimari ve teknik gereksinimler arasında ilişki kurularak çözümlendiği sonucuna ulaşılmıştır.

IV. TARTIŞMA

Bu çalışmada, konut yapılarında yer alan müşterek alanların yapı içerisindeki mekânsal organizasyonu, birbirleriyle olan ilişkileri, kullanıcı sirkülasyonu ve işlevsel gereklilikleri incelenmiştir. Ayrıca söz konusu alanların planlama ve boyutlandırma süreçleri; mimari çizimler, kullanıcı hareketleri, tefriş elemanlarının yerleşimi ve ilgili mevzuat hükümlerinde belirtilen minimum ölçütler doğrultusunda değerlendirilmiştir. Bu kapsamda, müşterek alanların yalnızca yapı içerisinde ayrılması gereken fiziksel alanlar olmadığı; yapının işlevselliği, kullanıcı güvenliği, erişilebilirliği ve teknik sürekliliği açısından önemli tasarım bileşenleri olduğu ortaya konulmuştur.

İncelenen 33 bağımsız bölümlü konut uygulama projesinde sığınak, su (hidrofor) deposu, enerji ve haberleşme odası, mekanik ve havalandırma şaftları, jeneratör odası, asansör, kat holü ve ortak kat merdiveni gibi temel müşterek alanlar değerlendirilmiştir. Yapılan incelemelerde, söz konusu birimlerin ilgili yönetmeliklerde belirtilen teknik ve boyutsal gereklilikler doğrultusunda planlandığı ve mimari uygulama projesine aktarıldığı görülmüştür. Bununla birlikte, müşterek alanların yeterliliği yalnızca metrekare değerleri üzerinden değil, alan içerisindeki tefrişlerin yerleşimi, kullanıcı hareketleri, erişim koşulları ve diğer yapı birimleriyle olan ilişkileri üzerinden de değerlendirilmiştir.

Sığınak alanına ilişkin yapılan hesaplamalarda, bağımsız bölümlerin oda sayıları dikkate alınarak gerekli minimum alan 71,00 m² olarak belirlenmiştir. Projede ise sığınak için 77,34 m² net alan ayrılmıştır. Böylece mevzuat doğrultusunda hesaplanan minimum alan koşulunun sağlandığı ve gerekli kullanım alanının üzerinde bir çözüm oluşturulduğu belirlenmiştir. Sığınak planlamasında birbirine dik konumlandırılmış iki giriş kapısı, mutfak nişi, kadın ve erkek WC-duş birimleri ile engelli WC-duş alanı gibi yardımcı işlevlerin de çözümlenmesi, sığınağın yalnızca minimum alan koşulunu karşılayan bir mekân olarak değil, kullanıcı gereksinimlerine cevap verebilecek işlevsel bir birim olarak ele alındığını göstermektedir.

Su (hidrofor) deposu açısından yapılan değerlendirmede, su deposunun teknik ekipmanlarının yerleştirilmesi ve personel sirkülasyonunun sağlanması için gerekli alan dikkate alınmıştır. Projede su deposu için 3,73 m² alan ayrılmıştır. Bu alanın, yapının su ihtiyacının karşılanması ve gerekli teknik müdahalelerin gerçekleştirilebilmesi açısından yeterli bir çözüm sunduğu değerlendirilmiştir. Böylece su deposunun yalnızca teknik bir gereklilik olarak değil, bakım ve kullanım koşulları da dikkate alınarak planlandığı görülmüştür.

Enerji ve haberleşme odasının değerlendirilmesinde, iki farklı teknik işlevin tek bir mekân içerisinde çözümlendiği görülmüştür. Elektrik ve haberleşme sistemlerine ait panoların yerleşimi, bakım ve müdahale için gerekli kullanım alanları ve personel sirkülasyonu dikkate alınarak yapılan çözümlemede toplam 10,26 m² alan ayrılmıştır. Bu durum, farklı teknik sistemlerin aynı mekân içerisinde bir arada çözümlenmesinin, gerekli erişim ve kullanım koşulları sağlandığında işlevsel bir alternatif oluşturabileceğini göstermektedir.

Mekanik ve havalandırma şaftlarının değerlendirilmesinde, yapı içerisindeki hava sirkülasyonu ve mekanik tesisatın sürekliliğinin sağlanması temel kriterler olarak ele alınmıştır. Projede mekanik, havalandırma ve elektrik şaftları için gerekli minimum kesitlerin sağlandığı görülmüştür. Şaftların 60 × 60 cm minimum ölçü üzerinden çözümlenmesi ve ilgili teknik birimlerle ilişkilendirilmesi, yapı içerisindeki tesisat sistemlerinin düzenli bir biçimde organize edilmesine olanak sağlamıştır. Bu nedenle mekanik ve havalandırma şaftlarının, mimari planlama sürecinin erken aşamalarında dikkate alınması gereken teknik müşterek alanlar olduğu değerlendirilmiştir.

Jeneratör odası için yapılan değerlendirmede, jeneratörün yerleşimi, tefriş elemanlarının konumlandırılması ve bakım amacıyla gerekli personel hareket alanları dikkate alınmıştır. Projede jeneratör odası için 7,66 m² alan ayrılmıştır. Elde edilen bulgular, teknik ekipmanların yalnızca fiziksel olarak yerleştirilmesinin yeterli olmadığını; ekipman çevresinde bakım, kontrol ve müdahaleye olanak sağlayacak yeterli sirkülasyon alanlarının da oluşturulması gerektiğini göstermektedir.

Asansör alanının değerlendirilmesinde, yapı içerisindeki düşey sirkülasyonun etkin biçimde sağlanması temel kriter olarak ele alınmıştır. Projede asansör için 5,87 m² alan ayrılmıştır. Asansör kabininin iç ölçülerinin ilgili mevzuatta belirtilen koşulları sağlaması ve asansörün kat holüyle doğrudan ilişkilendirilmesi, yapı içerisindeki erişilebilirlik ve düşey sirkülasyon açısından önem taşımaktadır. Bu doğrultuda asansörün, yalnızca katlar arasındaki ulaşımı sağlayan teknik bir birim olmadığı; aynı zamanda ortak sirkülasyon sisteminin temel bileşenlerinden biri olduğu sonucuna ulaşılmıştır. Kat hollerine ilişkin değerlendirmede, bu alanların yalnızca belirli bir katta değil, yapının tüm katlarında bağımsız bölümlere erişimi sağlayacak biçimde düzenlenmesi gerektiği belirlenmiştir. Projede kat hollerinin en dar noktalarında 150 cm minimum geçiş koşulunun sağlandığı ve toplam 85,29 m² alan ayrıldığı görülmüştür. Kat holünün asansör, merdiven ve bağımsız bölüm girişleriyle doğrudan ilişkilendirilmesi, yapı içerisindeki kullanıcı sirkülasyonunun kesintisiz bir biçimde gerçekleştirilmesini sağlamaktadır.

Ortak kat merdivenine ilişkin değerlendirmelerde ise kat yüksekliği, basamak sayısı, rıht yüksekliği ve merdiven genişliği temel tasarım kriterleri olarak ele alınmıştır. İncelenen projede kat yüksekliği 310 cm, basamak sayısı 18 ve rıht yüksekliği yaklaşık 17,2 cm olarak belirlenmiştir. Merdivenin toplamda 10,77 m² alan kapladığı tespit edilmiştir. Ayrıca merdivenin dar kenarının 120 cm'den az olmaması yönündeki mevzuat koşulunun sağlandığı görülmüştür. Bu veriler, merdiven tasarımının hem mevzuat hükümleri hem de kullanıcı güvenliği ve ergonomisi açısından değerlendirilmesi gerektiğini göstermektedir. Genel olarak değerlendirildiğinde, incelenen projede müşterek alanların planlanmasında iki temel yaklaşımın birlikte ele alındığı görülmektedir. Birinci yaklaşım, ilgili mevzuat tarafından belirlenen minimum alan ve boyut koşullarının karşılanmasıdır. İkinci yaklaşım ise bu alanların gerçek kullanım koşulları içerisinde işlevsel, erişilebilir ve kullanıcı gereksinimlerine cevap verebilir nitelikte tasarlanmasıdır. Bu iki yaklaşımın birlikte değerlendirilmesi, müşterek alanların yalnızca mevzuata uyum açısından değil, yapı performansı ve kullanıcı memnuniyeti açısından da yeterli olmasını sağlamaktadır. Müşterek alanların eksik, yetersiz veya hatalı planlanması durumunda yapı kullanım sürecinde çeşitli sorunların ortaya çıkması mümkündür. Teknik ekipmanlara erişimin kısıtlanması, bakım ve onarım faaliyetlerinin güçleşmesi, kullanıcı sirkülasyonunun olumsuz etkilenmesi, erişilebilirlik sorunları ve mevzuata aykırılıkların oluşması bu sorunların başlıcaları arasında değerlendirilebilir. Bu nedenle müşterek alan gereksinimlerinin mimari tasarımın son aşamasında çözümlenmesi yerine, projenin başlangıç aşamalarından itibaren tasarım sürecine dahil edilmesi gerekmektedir.

Sonuç olarak, müşterek alanların konut yapılarının işlevselliğini artıran, kullanıcıların ortak gereksinimlerini karşılayan, teknik sistemlerin sürekliliğini sağlayan ve yapı içerisindeki sirkülasyonu düzenleyen temel bileşenler olduğu belirlenmiştir. İncelenen 33 bağımsız bölümlü konut projesi üzerinden yapılan hesaplama ve plan çözümlemeleri, müşterek alanların her birinin kendine özgü kullanım, boyutlandırma ve mekânsal ilişki gerekliliklerine sahip olduğunu ortaya koymuştur. Bu kapsamda, müşterek alanların tasarımında yalnızca yönetmeliklerde belirtilen minimum metrekare değerlerinin esas alınmasının yeterli olmadığı; aynı zamanda kullanıcı hareketleri, tefriş yerleşimleri, teknik ekipmanların konumu, bakım ve müdahale olanakları ile farklı mekânlar arasındaki sirkülasyon ilişkilerinin de birlikte değerlendirilmesi gerektiği sonucuna ulaşılmıştır. Bu doğrultuda, mimari uygulama projelerinde müşterek alanların erken tasarım aşamasında belirlenmesi ve mevzuat, kullanıcı gereksinimleri ve teknik gereklilikler birlikte değerlendirilerek boyutlandırılması; hem proje onay süreçlerinin sağlıklı yürütülmesi hem de yapının kullanım sürecinde güvenli, işlevsel ve sürdürülebilir bir çevre oluşturulması açısından önem taşımaktadır.

KAYNAKLAR

Açıkalın, D. (2019). *Kurumsal kaynak planlaması sistemi entegrasyonu sürecinde proje yönetimi yaklaşımı ve bir firma incelemesi* [Yüksek lisans tezi, Dokuz Eylül Üniversitesi].

Alsibaai, L. (2022). *Sürdürülebilir mimarlık bağlamında yapılı çevrede değişim ve uyarlanabilirlik kapasitelerinin işlevsellik açısından incelenmesi*.

Asansör Yönetmeliği. (2014). *Resmî Gazete* (Sayı 29396).

Cemiloğlu, N. (2021). Kat mülkiyetinde ortak yer kavramı ile ortak yerlerdeki hak ve yükümlülükler, 82–95.

Elektrik İç Tesisleri Proje Hazırlama Yönetmeliği. (2003). *Resmî Gazete* (Sayı 25305).

Hilmioğlu, S., & İşanç, D. (2022). Konut yapılarına ait ortak alanların evrensel tasarım ilkeleri kapsamında incelenmesi. *Sürdürülebilir Mühendislik Uygulamaları ve Teknolojik Gelişmeler Dergisi, 5*(1), 66–80.

Kılıç, T., & Hardal, S. (2019). İstanbul'daki kentsel dönüşüm projelerinin genel bir eleştirisi. *Journal of International Social Research, 12*(62), 348–355.

Mekanik Tesisat Fenni Mesuliyet Hizmetleri Yönetmeliği. (2012). *Resmî Gazete* (Sayı 28177).

Planlı Alanlar İmar Yönetmeliği. (2017). *Resmî Gazete* (Sayı 30113).

Sığınak Yönetmeliği. (1988). *Resmî Gazete* (Sayı 19910).

Su Yalıtım Yönetmeliği. (2017). *Resmî Gazete* (Sayı 30223).

Şengül, M. (2012). Kat mülkiyetinde ortak yer ve tesislerin niteliğinin değiştirilmesi. *Marmara Üniversitesi Hukuk Fakültesi Hukuk Araştırmaları Dergisi, 18*(3), 201–232.

Tanrıvermiş, H. (2019). Türkiye ekonomisinde inşaat ve gayrimenkul sektörlerinin rolleri ve gelişme eğilimlerinin değerlendirilmesi. *Kent, İnşaat ve Ekonomi Kongresi*, 36–57.

Yenice, T. K. (2019). Hasan Kalyoncu Üniversitesi öğrenci yurdu mimari proje tasarım süreci. *Journal of Architectural Sciences and Applications, 4*(2), 183–19.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Uluslararası Karayolu Taşımacılığında Sınır Geçişlerinden Kaynaklanan Riskler ve Operasyonel Etkileri

Muhammed Turgut*

[1]*Uluslararası Ticaret ve Lojistik / Uygulamalı Bilimler Fakültesi, Tarsus Üniversitesi, Türkiye*

**(muhammedturgut@tarsus.edu.tr)*

Özet – Gerçekleştirilen araştırma, uluslararası karayolu taşımacılığında sınır geçişlerinden kaynaklanan riskleri belirlemeyi ve bu risklerin taşıma operasyonları üzerindeki etkilerini kavramsal olarak değerlendirmeyi amaçlamaktadır. Araştırmada nitel araştırma yaklaşımı kapsamında doküman incelemesi ve literatür taramasına dayalı kavramsal değerlendirme yöntemi benimsenmiştir. Bu doğrultuda uluslararası kuruluşların ve kamu kurumlarının yayımladığı rapor, düzenleme ve uygulama belgeleri ile sınır geçişleri, karayolu yük taşımacılığı, gümrük işlemleri, taşıma maliyetleri, transit süreçleri ve ticaretin kolaylaştırılması konularındaki akademik çalışmalar birlikte incelenmiştir. İncelenen kaynaklarda öne çıkan risk unsurları tematik olarak sınıflandırılmış ve operasyonel etkileri bakımından değerlendirilmiştir. Yapılan değerlendirme sonucunda sınır geçişlerinden kaynaklanan riskler; gümrük ve belge kaynaklı riskler, bekleme ve sınır kapısı yoğunluğu riskleri, geçiş belgesi ve transit izin riskleri, mevzuat ve idari uygulama farklılıkları, politik, jeopolitik ve güvenlik riskleri ile dijital sistem ve koordinasyon riskleri olmak üzere altı temel grupta ele alınmıştır. Bulgular, söz konusu risklerin çoğu zaman birbirinden bağımsız gelişmediğini, bir risk unsurunun diğerini tetikleyerek sınır geçiş süresinin uzaması, teslimat planlarının bozulması, güzergâh değişiklikleri, araç ve sürücü kullanımında verimsizlik ve taşıma maliyetlerinde artış gibi sonuçlar doğurabildiğini göstermektedir. Çalışma ayrıca sınır geçiş risklerinin azaltılmasında yalnızca fiziksel kapasite yatırımlarının yeterli olmadığını dijitalleşme, prosedürlerin uyumlaştırılması, kurumlar arası koordinasyon, veri paylaşımı ve risk temelli sınır yönetiminin birlikte ele alınması gerektiğini ortaya koymaktadır.

Anahtar Kelimeler – Uluslararası karayolu taşımacılığı, sınır geçişleri, risk yönetimi, taşıma riski, uluslararası ticarette risk yönetimi

I. GİRİŞ

Uluslararası ticaretin fiziksel olarak sürdürülebilmesi, üretim ve tüketim noktaları arasındaki yük hareketlerinin zamanında, güvenli ve öngörülebilir biçimde gerçekleştirilebilmesine bağlıdır. Karayolu taşımacılığı, kapıdan kapıya taşıma imkânı sunması, güzergâh ve teslimat planlamasında esnek olması ve diğer taşıma modlarını tamamlayıcı bir işlev üstlenmesi nedeniyle uluslararası yük taşımacılığının temel unsurlarından biridir. Türkiye açısından da karayolu taşımacılığı dış ticarette önemli bir paya sahiptir. T.C. Ticaret Bakanlığı verilerine göre 2025 yılında parasal değer itibarıyla Türkiye ihracatının %31,9'u ve ithalatının %19'u karayoluyla gerçekleştirilmiştir. Aynı yıl karayolunun taşınan miktar açısından payı ihracatta %18,42, ithalatta ise %7,64 olarak gerçekleşmiştir (T.C. Ticaret Bakanlığı, 2026).

Uluslararası karayolu taşımacılığını yurtiçi karayolu taşımacılığından ayıran temel unsurlardan biri, taşıma sürecinin bir veya birden fazla ülke sınırının geçilmesini gerektirmesidir. Sınır geçişlerinde gümrük işlemleri, transit prosedürleri, belge kontrolleri, araç ve yük denetimleri, güvenlik kontrolleri ve farklı kamu kurumlarının yürüttüğü işlemler taşımanın kesintisiz ilerleyişini doğrudan etkileyebilmektedir. Bu nedenle sınır kapıları yalnızca ülkelerin giriş ve çıkış noktaları değil, aynı zamanda uluslararası taşıma zincirinde zaman ve maliyet performansının belirlendiği kritik operasyon noktalarıdır. Dünya Ticaret Örgütü Ticaretin Kolaylaştırılması Anlaşması da eşyanın serbest bırakılmasını hızlandırmak amacıyla varış öncesi işlemlerin

geliştirilmesini, risk yönetiminin kullanılmasını ve sınır kontrollerinden sorumlu kurumların koordinasyon içinde çalışmasını öngörmektedir (WTO, 2017).

Sınır geçişlerinde ortaya çıkan bekleme ve işlem süreleri, karayolu taşımacılığında yalnızca operasyonel bir gecikme olarak değerlendirilmemelidir. Bekleyen bir taşıtın sürücü, yakıt, araç kullanım süresi ve filo kapasitesiyle ilişkili maliyetleri devam ederken taşımanın planlanan teslim süresinden sapması tedarik zincirinin diğer aşamalarında da ek maliyet ve belirsizlik yaratabilmektedir. Doğu Avrupa ve Orta Asya ülkelerindeki karayolu yük taşımalarını inceleyen Iimi (2023), sınır geçişinin ortalama karayolu yük taşıma maliyetlerine yaklaşık %3–4 oranında ilave maliyet oluşturduğunu ortaya koymuştur. Benzer şekilde sınır geçişlerinde bekleme sürelerinin azaltılmasını inceleyen çalışmalar, bekleme süresindeki değişimin taşıma maliyetlerine yansıdığını ve sınır işlemlerinin etkinliğinin ticari rekabet gücü açısından ekonomik sonuçlar doğurduğunu göstermektedir (Roberts vd., 2014; Avetisyan vd., 2015).

Teslim süresinin uzaması kadar teslim süresinin öngörülemez hale gelmesi de uluslararası ticaret açısından önem taşımaktadır. Taşımanın ne zaman tamamlanacağının güvenilir biçimde tahmin edilememesi, firmaların stok, üretim, dağıtım ve müşteri teslimat planlarını doğrudan etkileyebilmektedir. Hummels ve Schaur (2013), taşıma süresinin uluslararası ticarette ekonomik değeri bulunan önemli bir unsur olduğunu ve transit süresindeki her ilave günün özellikle zaman hassas ürünlerde ticari maliyet yarattığını ortaya koymaktadır. Bu yaklaşım, sınır kapılarında meydana gelen gecikmelerin yalnızca taşıyıcı işletmelerin operasyonel performansıyla sınırlı olmadığını, dış ticaret işlemlerinin zaman ve maliyet yapısına yayılan daha geniş bir risk alanı oluşturduğunu göstermektedir.

Sınır geçiş performansı, gümrük işlemlerinin yanı sıra altyapı kapasitesi, belge ve izin süreçleri, ülkeler arasındaki idari uygulama farklılıkları, güvenlik tedbirleri ve sınır kurumları arasındaki koordinasyon gibi çok sayıda faktörden etkilenmektedir. Nitekim uluslararası uygulamalarda sınır geçişlerinin kolaylaştırılmasına yönelik programların bekleme sürelerini azaltabildiği, ancak elde edilen faydanın altyapı kapasitesi ve uygulama koşullarına göre farklılaşabildiği görülmektedir. Bradbury (2010), Kanada-ABD sınırındaki Free and Secure Trade uygulamasını incelediği çalışmasında programın incelenen beş büyük sınır geçiş noktasının dördünde ortalama bekleme süresini azalttığını, ancak uygulamanın faydalarının sınır kapıları ve işletmeler arasında eşit dağılmadığını belirlemiştir. Bu durum sınır geçiş risklerinin tek bir nedene indirgenemeyeceğini ve idari, operasyonel, teknik ve dış çevre kaynaklı faktörlerin birlikte değerlendirilmesi gerektiğini göstermektedir.

Bu çalışmanın temel problemi, uluslararası karayolu taşımacılığında sınır geçişlerinde yaşanan sorunların çoğu zaman yalnızca bekleme süresi veya gümrük gecikmesi üzerinden değerlendirilmesi, buna karşılık bu sorunların farklı risk kaynakları ile teslim süresi ve taşıma maliyetleri arasındaki bağlantısının bütüncül biçimde ele alınmamasıdır. Bu doğrultuda çalışmanın amacı, uluslararası karayolu yük taşımacılığında sınır geçişlerinden kaynaklanan temel riskleri sistematik biçimde sınıflandırmak ve söz konusu risklerin başta teslim süresi ve taşıma maliyetleri olmak üzere taşıma operasyonları üzerindeki etkilerini kavramsal olarak değerlendirmektir. Çalışmada ayrıca bu risklerin azaltılmasına yönelik uluslararası ticaretin kolaylaştırılması ve sınır yönetimi uygulamaları incelenmektedir. Bu amaç doğrultusunda çalışma iki temel araştırma sorusuna odaklanmaktadır:

AS1: Uluslararası karayolu yük taşımacılığında sınır geçişlerinden kaynaklanan temel risk alanları nelerdir?

AS2: "Sınır geçişlerinden kaynaklanan riskler uluslararası karayolu taşıma operasyonlarını, özellikle teslim süresi ve taşıma maliyetlerini hangi mekanizmalar üzerinden etkilemektedir?"

II. ARAŞTIRMA YÖNTEMİ

Bu çalışma, uluslararası karayolu yük taşımacılığında sınır geçişlerinden kaynaklanan risklerin mevcut akademik ve kurumsal bilgi birikimi üzerinden değerlendirilmesini amaçladığından nitel araştırma yaklaşımı kapsamında tasarlanmıştır. Araştırmada doküman incelemesi ile yapılandırılmış literatür taraması birlikte kullanılmıştır. Doküman incelemesi, araştırma problemiyle ilişkili yazılı kaynakların belirli bir araştırma amacı doğrultusunda sistematik biçimde seçilmesi, incelenmesi ve yorumlanmasına dayanan nitel bir araştırma yöntemidir (Bowen, 2009). Literatür taramasının araştırma yöntemi olarak kullanılmasında ise kaynakların gelişigüzel aktarılması yerine araştırma sorusuna dayalı açık bir seçim ve sentez sürecinin izlenmesi araştırmanın güvenilirliği açısından önem taşımaktadır (Snyder, 2019). Bu

nedenle çalışma klasik anlatısal bir literatür özeti şeklinde değil, belirlenen araştırma soruları çerçevesinde kaynakların sınıflandırılması ve karşılaştırılmasına dayalı kavramsal bir inceleme olarak yürütülmüştür.

Araştırmada kullanılan dokümanlar iki temel kaynak grubundan oluşturulmuştur. Birinci grupta sınır geçişleri, uluslararası karayolu taşımacılığı, transit işlemleri ve ticaretin kolaylaştırılması konusunda doğrudan düzenleyici veya uygulamaya yönelik bilgi sağlayan Dünya Ticaret Örgütü, Birleşmiş Milletler Avrupa Ekonomik Komisyonu, Dünya Bankası, Asya Kalkınma Bankası, T.C. Ticaret Bakanlığı ile T.C. Ulaştırma ve Altyapı Bakanlığı kaynakları yer almaktadır. İkinci grup ise uluslararası karayolu yük taşımacılığı, sınır geçiş süreleri, gümrük gecikmeleri, taşıma maliyetleri ve sınır yönetimi konularında yayımlanmış hakemli akademik çalışmalardan oluşmaktadır. Akademik literatürde özellikle sınırda bekleme süreleri ile yük taşıma maliyetleri arasındaki ilişkiyi inceleyen çalışmalar, çalışmanın teslim süresi ve maliyet boyutunun oluşturulmasında temel referanslar olarak değerlendirilmiştir (Roberts vd., 2014; Avetisyan vd., 2015; Iimi, 2023).

Kaynak seçiminde araştırma sorularıyla doğrudan ilişki, kaynağın kurumsal veya akademik güvenilirliği, bulguların karayolu yük taşımacılığına uygulanabilirliği ve kaynağın doğrulanabilir olması temel ölçütler olarak kullanılmıştır. Akademik kaynaklarda hakemli dergilerde yayımlanmış çalışmalara, kurumsal kaynaklarda ise ilgili kuruluşların resmî rapor, sözleşme, veri seti ve mevzuat dokümanlarına öncelik verilmiştir. Konu literatüründe 2010–2026 dönemindeki çalışmalara ağırlık verilmekle birlikte araştırmanın kavramsal veya yöntemsel çerçevesi açısından temel nitelik taşıyan daha eski çalışmalar da kapsamda tutulmuştur. Yalnızca yolcu geçişlerini ele alan ve yük taşımacılığı açısından aktarılabilir sonuç içermeyen çalışmalar, sınır geçişiyle doğrudan bağlantısı bulunmayan genel lojistik çalışmaları, kaynağı doğrulanamayan internet içerikleri ve aynı bulguyu tekrar eden ikincil kaynaklar değerlendirme dışında bırakılmıştır.

Dokümanların analizinde tematik içerik analizi yaklaşımı kullanılmıştır. Analiz çerçevesi üç üst boyut üzerinden oluşturulmuştur: risk kaynağı, operasyonel etki ve risk azaltma yaklaşımı. Operasyonel etkiler başta teslim süresi ve taşıma maliyetleri olmak üzere güzergâh ve taşıma planlaması ile araç ve sürücü kullanımına ilişkin sonuçlar üzerinden değerlendirilmiştir. İlk aşamada kaynaklarda sınır geçişinde gecikme, bekleme, belge veya izin sorunu, kontrol, altyapı yetersizliği, mevzuat farklılığı, güvenlik, dijital sistem ve kurumlar arası koordinasyon gibi risk ifadeleri belirlenmiştir. İkinci aşamada benzer risk ifadeleri ortak temalar altında birleştirilmiştir. Üçüncü aşamada her risk grubunun teslim süresi ve taşıma maliyetleriyle ilişkisi kaynaklarda ortaya konulan nedensel mekanizmalar dikkate alınarak değerlendirilmiştir. Son aşamada ise risklerin azaltılmasına yönelik uygulamalar uluslararası kuruluşların ticaretin kolaylaştırılması ve sınır yönetimi çerçeveleriyle karşılaştırılmıştır. Bu süreç sonunda sınır geçiş riskleri gümrük ve belge kaynaklı riskler, bekleme ve yoğunluk riskleri, geçiş belgesi ve transit izin riskleri, mevzuat ve idari uygulama farklılıkları, politik ve güvenlik riskleri ile dijital sistem ve koordinasyon riskleri olmak üzere altı ana başlık altında sınıflandırılmıştır.

III. ULUSLARARASI KARAYOLU TAŞIMACILIĞI VE SINIR GEÇİŞ SÜRECİ

Uluslararası karayolu taşımacılığı, yükün çıkış noktasından varış noktasına kadar doğrudan taşınabilmesine imkân vermesi ve farklı taşıma türleriyle bütünleşebilmesi nedeniyle dış ticaret lojistiğinin önemli unsurlarından biridir. Ancak taşımanın birden fazla ülkenin egemenlik alanında gerçekleştirilmesi, yurtiçi taşımacılıktan farklı olarak sınır geçişleri, gümrük işlemleri ve transit prosedürleri gibi ilave süreçlerin ortaya çıkmasına neden olmaktadır. Bu nedenle uluslararası karayolu taşımacılığının etkinliği yalnızca taşıyıcının operasyonel performansına değil, güzergâh üzerindeki sınır kapılarının ve ilgili kamu kurumlarının işlem kapasitesine de bağlıdır. OSCE ve UNECE tarafından hazırlanan sınır geçişlerine ilişkin iyi uygulamalar el kitabında da sınır yönetimi; gümrükleme ve transit işlemleri, kurumlar arası iş birliği, risk yönetimi, sınır kapısı tasarımı, bilgi teknolojileri ve sınır performansının ölçülmesi gibi birbirini tamamlayan boyutlarla ele alınmaktadır (OSCE & UNECE, 2012).

Sınır geçiş süreci, yükün ve taşıtın bir ülkeden çıkışı ile diğer ülkeye girişinin gerçekleştirilmesi sırasında farklı idari ve operasyonel işlemlerin tamamlanmasını gerektirmektedir. Gümrük beyanlarının ve taşıma belgelerinin kontrolü, transit rejimine ilişkin işlemler, araç ve yük kontrolleri, gerekli izinlerin doğrulanması ve güvenlik amacıyla gerçekleştirilen denetimler bu sürecin temel unsurları arasında yer

almaktadır. Bununla birlikte gerçekleştirilen işlemlerin türü ve yoğunluğu taşınan eşyanın niteliğine, kullanılan transit rejimine ve ilgili ülkelerin mevzuatına göre farklılaşabilmektedir. Sınır geçişlerinin kolaylaştırılmasına yönelik uluslararası çerçevenin temel belgelerinden biri olan 1982 tarihli Sınır Kontrollerinin Uyumlaştırılması Uluslararası Sözleşmesi, sınır kontrollerinin uyumlaştırılması ve koordinasyonunun geliştirilmesini amaçlamaktadır. Sözleşmeye 2008 yılında yürürlüğe giren Karayolu Taşımacılığına ilişkin Ek 8 ile uluslararası karayolu taşımalarında sınır geçiş prosedürlerinin hızlandırılması, kontrol süreçlerinin daha etkin yürütülmesi ve sınır kapılarındaki operasyonların kolaylaştırılmasına yönelik özel hükümler getirilmiştir (UNECE, 2008).

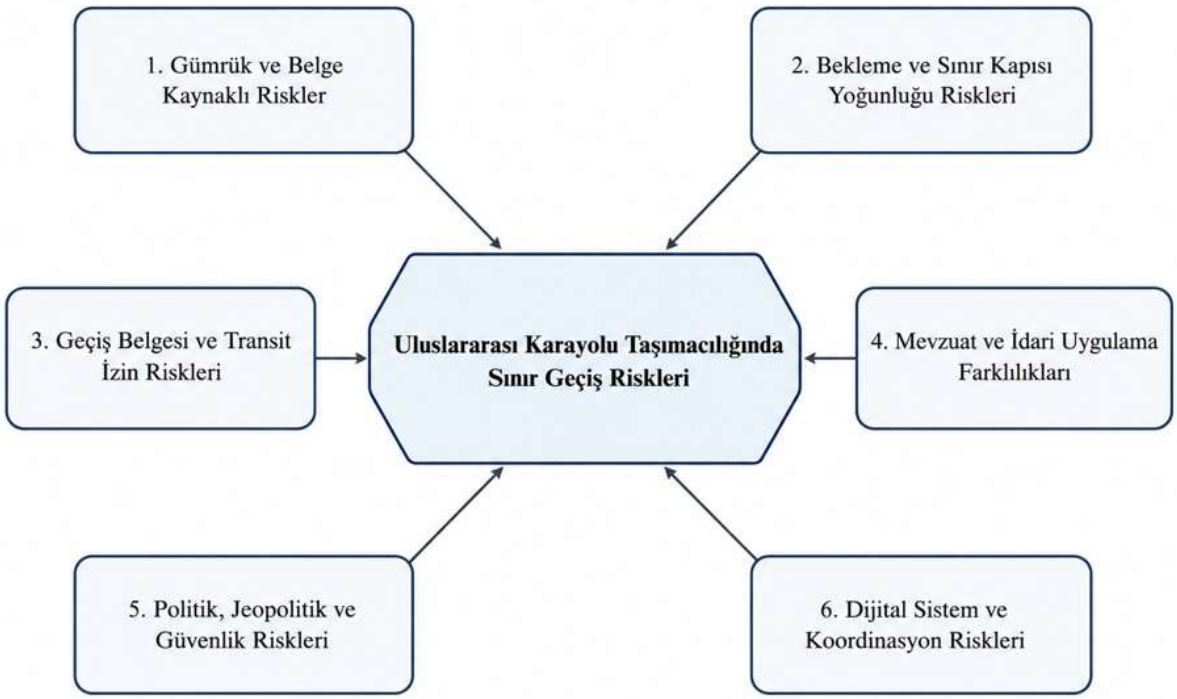
Uluslararası karayolu taşımacılığında transit işlemlerinin kolaylaştırılmasına yönelik önemli mekanizmalardan biri de TIR sistemidir. TIR sistemi, eşyanın hareket noktası ile varış noktası arasındaki uluslararası taşımada gümrük transit işlemlerinin ortak bir sistem üzerinden yürütülmesine imkân vermektedir. TIR kapsamında taşınan yüklerin mühürlü taşıtlarda veya yükleme birimlerinde hareket etmesi ve uluslararası bir garanti sisteminin bulunması, transit ülkelerde tekrarlanan gümrük kontrollerinin azaltılmasına katkı sağlamaktadır. UNECE, TIR sisteminin uluslararası eşya taşımalarını kolaylaştırmayı amaçladığını ve sistem kapsamında mühürlü taşıtlardaki eşyanın kural olarak sınır geçişlerinde yeniden fiziksel kontrole tabi tutulmadığını belirtmektedir (UNECE, 2024).

Sınır geçişlerinin etkinliği açısından yalnızca gümrük idaresinin performansı yeterli değildir. Aynı sınır kapısında gümrük, güvenlik, ulaştırma, sağlık, tarım veya ürün denetiminden sorumlu farklı kamu kurumlarının işlem yapabilmesi, kurumlar arasındaki koordinasyonu sınır performansının temel unsurlarından biri haline getirmektedir. WTO Ticaretin Kolaylaştırılması Anlaşması'nın 8. maddesi, ithalat, ihracat ve transit işlemlerinden sorumlu sınır kurumlarının birbirleriyle iş birliği ve koordinasyon içinde çalışmasını öngörmekte; ortak sınıra sahip ülkeler arasında çalışma gün ve saatlerinin uyumlaştırılması, prosedür ve formalitelerin eşgüdümü, ortak tesislerin kullanılması, ortak kontroller ve tek duraklı sınır kapısı uygulamaları gibi yöntemlere işaret etmektedir. Aynı Anlaşma, varış öncesi belge işleme, risk yönetimi ve ithalat, ihracat ve transit formalitelerinin sadeleştirilmesini de sınır işlemlerinin hızlandırılmasına yönelik araçlar arasında düzenlemektedir (WTO, 2017).

Bu yapı, sınır geçiş sürecinin uluslararası karayolu taşımacılığında yalnızca idari bir zorunluluk olmadığını, taşımanın zaman ve maliyet performansını doğrudan etkileyebilen bir operasyon alanı olduğunu göstermektedir. Belgelerin eksik veya uyumsuz olması, sınır kapılarında yoğunluk oluşması, transit ve geçiş izinlerinde sorun yaşanması, farklı idari uygulamalar, güvenlik tedbirleri veya dijital sistemlerde meydana gelen aksaklıklar taşıma sürecinin planlanan şekilde devam etmesini engelleyebilmektedir. Dolayısıyla sınır geçişinde yer alan işlem ve aktör sayısının artması, taşıyıcılar açısından farklı nitelikte risk kaynaklarının ortaya çıkmasına zemin hazırlamaktadır. İzleyen bölümde bu riskler, uluslararası karayolu taşımacılığındaki etkileri dikkate alınarak sistematik biçimde sınıflandırılmaktadır.

IV. ULUSLARARASI KARAYOLU TAŞIMACILIĞINDA SINIR GEÇİŞLERİNDEN KAYNAKLANAN RİSKLER

Sınır geçişleri, uluslararası karayolu taşımacılığında taşıma süresinin ve maliyet yapısının farklı idari, operasyonel ve dış çevre faktörlerinden etkilenebildiği kritik aşamalardır. Sınır kapılarında tamamlanması gereken gümrük işlemleri, belge kontrolleri, transit prosedürleri, geçiş izinleri, güvenlik kontrolleri ve farklı kurumların denetimleri taşımanın planlanan akışından sapmasına neden olabilecek çeşitli riskler ortaya çıkarmaktadır. Bu riskler birbirinden tamamen bağımsız değildir. Örneğin belge kaynaklı bir eksiklik ilave gümrük kontrolüne, kontrol süresinin uzaması sınır kapısında beklemeye, bekleme ise teslim süresinin uzamasına ve taşıma maliyetinin artmasına neden olabilmektedir. Bu nedenle sınır geçiş risklerinin yalnızca ortaya çıktıkları işlem aşamasına göre değil, taşıma süresi ve maliyetleri üzerinde oluşturdukları zincirleme etkiler dikkate alınarak değerlendirilmesi gerekmektedir.



Şekil 1. Uluslararası Karayolu Taşımacılığında Sınır Geçişlerinden Kaynaklanan Riskler

A. Gümrük ve Belge Kaynaklı Riskler

Uluslararası karayolu taşımacılığında gümrük ve belge kaynaklı riskler, sınır geçiş sürecinin temel risk alanlarından birini oluşturmaktadır. Bir uluslararası taşımanın gerçekleştirilebilmesi, yükün niteliğine, taşıma güzergâhına ve uygulanan gümrük rejimine bağlı olarak çeşitli ticari, taşıma ve gümrük belgelerinin hazırlanmasını ve ilgili idarelere sunulmasını gerektirmektedir. Belgelerde yer alan bilgilerin eksik, hatalı veya birbirleriyle uyumsuz olması, gümrük idaresinin ek bilgi ya da belge talep etmesine, işlemin yeniden kontrol edilmesine veya yükün serbest bırakılmasının gecikmesine neden olabilmektedir. WTO Ticaretin Kolaylaştırılması Anlaşması'nın ithalat, ihracat ve transit formalitelerine ilişkin hükümlerinde belge ve formalite gerekliliklerinin azaltılması ve basitleştirilmesi, işlemlerin hızlı biçimde tamamlanması ve uyum maliyetlerinin düşürülmesi temel amaçlar arasında yer almaktadır. Anlaşmada ayrıca destekleyici belgelerin kâğıt veya elektronik kopyalarının kabul edilmesi ve belgelerin tek pencere üzerinden sunulabilmesine yönelik düzenlemeler bulunması, belge süreçlerinin sınır geçiş performansı açısından taşıdığı önemi göstermektedir (WTO, 2017).

Gümrük kaynaklı riskler yalnızca belgelerin eksik veya hatalı hazırlanmasından doğmamaktadır. Sunulan bilgilerin riskli görülmesi, yükün ilave kontrole yönlendirilmesi, farklı kurumların aynı sevkiyat üzerinde ayrı işlemler gerçekleştirmesi veya gümrük prosedürlerinin öngörülenden uzun sürmesi de taşıma süresinde belirsizlik yaratabilmektedir. Volpe Martincus vd. (2015), Uruguay'ın ihracat işlemlerine ilişkin ayrıntılı gümrük verilerini kullanarak gümrük işlem sürelerindeki artışın firmaların ihracat performansını olumsuz etkilediğini ortaya koymuştur. Çalışmada gümrük gecikmelerindeki %10'luk artışın ihracatta yaklaşık %3,8'lik azalmayla ilişkili olduğu belirlenmiştir. Bu bulgu doğrudan karayolu sınır kapılarına özgü olmamakla birlikte, gümrük işlem sürelerinin uluslararası ticarette yalnızca idari bir bekleme değil ekonomik sonuçlar doğuran bir zaman maliyeti olduğunu göstermesi bakımından önemlidir (Volpe Martincus vd., 2015).

Belge akışının ve gümrük işlemlerinin dijitalleştirilmesi, bu risklerin azaltılmasına yönelik önemli uygulamalardan biridir. Caballini ve Benzi (2023), İtalya'daki Fast Corridor uygulamasını IKEA örneği üzerinden inceleyerek gümrük süreçlerinin teknoloji ve veri paylaşımıyla yeniden tasarlanmasının yük akışının hızlandırılması ve tedarik zinciri rekabetçiliğinin geliştirilmesi açısından sağlayabileceği katkıları ortaya koymuştur. Bu tür uygulamalar, gümrük işlemlerinin fiziksel sınır veya terminal noktasında yoğunlaşmasının azaltılması, belgelerin önceden işlenmesi ve yük hareketlerinin dijital olarak izlenmesi

yoluyla işlem sürelerinin daha öngörülebilir hale getirilmesine katkı sağlayabilmektedir (Caballini & Benzi, 2023). WTO'nun varış öncesi işlemeye ve elektronik belge sunumuna ilişkin hükümleri de yük sınır noktasına ulaşmadan önce gerekli bilgilerin işlenebilmesini ticaretin kolaylaştırılmasının temel araçları arasında değerlendirmektedir (WTO, 2017).

Bu çerçevede uluslararası karayolu taşımacılığında gümrük ve belge kaynaklı risklerin temel sonucu, sınır geçiş süresinin hem uzaması hem de öngörülebilirliğinin azalmasıdır. Eksik veya uyumsuz belge, ilave kontrol ya da yeniden işlem gerektiren bir durum taşıtın sınır kapısında planlanandan daha uzun süre kalmasına neden olabilir. Taşıyıcı açısından bu durum yalnızca teslimat gecikmesi değil, araç ve sürücü zamanının verimsiz kullanılması ve sonraki taşıma operasyonlarının aksaması gibi sonuçlar da doğurabilir. Bu nedenle doğru ve tutarlı belge yönetimi, bilgilerin sınır geçişinden önce kontrol edilmesi ve elektronik veri paylaşımının etkin kullanılması gümrük ve belge kaynaklı risklerin azaltılmasında temel unsurlar olarak değerlendirilebilir.

B. Bekleme ve Sınır Kapısı Yoğunluğu Riskleri

Sınır kapılarında meydana gelen bekleme ve yoğunluk, uluslararası karayolu yük taşımacılığının en görünür operasyonel risklerinden biridir. Taşıtların sınır kapısına geliş hızının kontrol ve işlem kapasitesini aşması, gümrük ve diğer sınır kurumlarında işlem sürelerinin uzaması, personel veya fiziksel kapasite yetersizlikleri ve belirli dönemlerde artan taşıma talebi kuyrukların oluşmasına neden olabilmektedir. Bu durum, sınır geçiş süresinin taşıyıcı tarafından tam olarak kontrol edilemeyen bir değişkene dönüşmesine yol açmaktadır. Dolayısıyla sınır kapısındaki yoğunluk yalnızca belirli bir süre bekleme anlamına gelmemekte, taşımanın toplam transit süresinin öngörülebilirliğini de azaltmaktadır.

Türkiye'nin Avrupa ile karayolu taşımacılığındaki en önemli geçiş noktalarından biri olan Kapıkule üzerine yapılan çalışmalar, sınır kapısı kapasitesi ile taşıt akışı arasındaki ilişkinin önemini açık biçimde ortaya koymaktadır. Ülengin vd. (2021), Kapıkule Sınır Kapısı'ndaki mevcut süreçleri yetkililerle yapılan derinlemesine görüşmeler ve ayrık olay simülasyonu yardımıyla incelemiş, uzun kamyon kuyrukları ve uzun hizmet sürelerinin sınır geçiş performansını olumsuz etkilediğini belirlemiştir. Çalışmada farklı iyileştirme senaryoları test edilerek kuyrukların ve gecikmelerin azaltılmasına yönelik alternatifler değerlendirilmiştir. Bu bulgular, sınır kapılarındaki yoğunluğun yalnızca taşıma talebindeki artıştan değil, işlem süreçlerinin ve mevcut kapasitenin taşıt akışına cevap verme düzeyinden de kaynaklanabildiğini göstermektedir (Ülengin vd., 2021).

Sınırda bekleme süresinin uzaması taşıma işletmesi açısından doğrudan zaman ve maliyet kaybına dönüşebilmektedir. Roberts vd. (2014), ABD'deki giriş noktalarında bekleme sürelerindeki değişimin ekonomik etkilerini inceleyerek sınırdaki bekleme süresinin taşıma maliyetleriyle bağlantılı olduğunu göstermiştir. Çalışmada bekleme sürelerindeki değişimler, yolcu ve yük hareketlerinde ortaya çıkan zaman ve lojistik maliyetleri üzerinden değerlendirilmiştir. Avetisyan vd. (2015) ise ABD'deki on iki büyük kara yük sınır geçişinde gümrük personelinin artırılması sonucunda bekleme sürelerinde meydana gelebilecek azalmanın yük taşıma maliyetlerine nasıl yansıdığını incelemiştir. Araştırmada bekleme süresindeki değişim önce lojistik model aracılığıyla taşıma maliyetine dönüştürülmüş, daha sonra bu maliyet değişiminin ticaret ve makroekonomik etkileri analiz edilmiştir. Bu yaklaşım, sınır kapısında geçirilen ek zamanın yalnızca operasyonel bir gecikme değil, ölçülebilir ekonomik maliyet oluşturan bir unsur olduğunu göstermektedir (Roberts vd., 2014; Avetisyan vd., 2015).

Bekleme ve sınır geçiş maliyetleri arasındaki ilişki farklı coğrafyalarda da gözlenmektedir. Iimi (2023), Doğu Avrupa ve Orta Asya'daki uzun mesafeli karayolu yük taşımalarına ilişkin geniş ölçekli sevkiyat verilerini kullanarak sınır geçişinin ortalama yük taşıma maliyetlerine yaklaşık %3–4 oranında ilave maliyet oluşturduğunu belirlemiştir. Çalışmada taşıma maliyetlerinin sürücü ve yakıt gibi araç işletme maliyetlerinin yanı sıra sınır geçişleri ve kontrol noktalarındaki idari ve ticari harcamaları da içerdiği belirtilmektedir. Bu bulgu, sınır geçişlerinin karayolu yük taşımacılığının toplam maliyet yapısı içinde ayrı bir maliyet unsuru oluşturabildiğini göstermektedir (Iimi, 2023).

Sınır kapısı yoğunluğunun oluşturduğu riskin önemli bir diğer boyutu bekleme süresindeki değişkenliktir. Taşıyıcı açısından sınır geçişinin ortalama olarak ne kadar sürdüğü kadar, bu sürenin ne ölçüde öngörülebilir olduğu da operasyon planlamasını etkilemektedir. Bekleme süresinin beklenenden uzun gerçekleşmesi sürücünün çalışma ve dinlenme planını, aracın sonraki taşıma görevine tahsis

edilmesini ve müşteriye verilen teslim zamanını etkileyebilir. Bu nedenle sınır geçiş performansının iyileştirilmesinde yalnızca ortalama bekleme süresinin azaltılması değil, işlem kapasitesinin talep dalgalanmalarına karşı daha dayanıklı hale getirilmesi ve geçiş sürelerinin öngörülebilirliğinin artırılması da önem taşımaktadır. Bu çerçevede sınır kapısı altyapısı, personel kapasitesi, işlem süreçlerinin organizasyonu ve yoğunluk yönetimi bekleme ve sınır kapısı yoğunluğu riskinin azaltılmasında birlikte değerlendirilmesi gereken unsurlardır.

C. Geçiş Belgesi ve Transit İzinlerinden Kaynaklanan Riskler

Uluslararası karayolu yük taşımacılığında geçiş belgeleri ve taşıma izinleri, taşıyıcının belirli bir ülkeye giriş yapabilmesi, o ülke üzerinden transit geçebilmesi veya üçüncü ülke taşımacılığı gerçekleştirebilmesi açısından önemli düzenleyici araçlardır. İkili ve çok taraflı anlaşmalar çerçevesinde kullanılan bu belgelerin türü, miktarı ve kullanım koşulları ülkelere ve taşımanın niteliğine göre farklılık gösterebilmektedir. Türkiye'de geçiş belgelerinin yıllık ve dönemsel kullanımına ilişkin uygulamalar da belge arzının sınırsız olmadığını ve bazı belgelerin belirlenen kotalar çerçevesinde kullanıldığını göstermektedir. Ulaştırma ve Altyapı Bakanlığının 2026 yılı geçiş belgelerine ilişkin düzenlemesinde, bazı ülke kotalarının önceki yıl tükenmiş olması, belgelerin belirli dönemlerde kullanıma açılması ve transit veya üçüncü ülke belgelerinin belirli taşıma türlerine tahsis edilmesi bu yapının güncel örnekleri arasında yer almaktadır (T.C. Ulaştırma ve Altyapı Bakanlığı, 2025).

Geçiş belgesi kotalarının taşıma talebini karşılamaması, uluslararası karayolu taşımacılığında güzergâh ve taşıma planlaması açısından önemli bir risk oluşturabilmektedir. Ülengin vd. (2015), 18 Avrupa Birliği ülkesi için 2005–2012 dönemini kapsayan panel veri kullanarak karayolu taşıma kotalarının Türkiye'nin dış ticareti üzerindeki etkisini incelemiştir. Araştırma sonuçları, kotaların Türkiye'nin karayoluyla gerçekleştirdiği ihracat üzerinde anlamlı olumsuz etkiye sahip olduğunu göstermiştir. Çalışmada incelenen dönem ve ülkeler için kota sınırlamalarının bulunmaması durumunda Türkiye'nin karayoluyla ihracatının yaklaşık 10,6 milyar ABD doları daha yüksek olabileceği tahmin edilmiştir. Araştırmada ayrıca firmaların kota engelleriyle karşılaştıklarında diğer taşıma modlarına yönelebildiklerine ilişkin bulgular elde edilmiştir (Ülengin vd., 2015).

Geçiş belgelerinin etkisi, belgenin türüne göre de farklılaşabilmektedir. Çekyay vd. (2017), Avrupa ülkelerinin Türkiye'ye uyguladığı ikili ve transit taşıma kotalarını 2005–2014 dönemindeki karayolu ihracatı üzerinden incelemiş ve maksimum akış modeli ile çekim modelini bütünleştiren bir yaklaşım kullanmıştır. Çalışmanın sonuçları, uygulanan kotaların Türkiye'nin karayoluyla gerçekleştirdiği ihracatta kayba neden olduğunu, ancak bu etkinin kota türleri arasında aynı düzeyde olmadığını göstermektedir. Araştırmada olumsuz etkinin büyük bölümünün ikili taşıma kotalarından kaynaklandığı, transit kotaların etkisinin ise daha sınırlı kaldığı belirlenmiştir (Çekyay vd., 2017).

Bu bulgular, geçiş belgesi ve kota riskinin yalnızca bir belgenin mevcut olup olmamasıyla sınırlı değerlendirilemeyeceğini göstermektedir. Belgenin hangi ülke için geçerli olduğu, ikili veya transit taşımaya izin verip vermediği, üçüncü ülke taşımalarında kullanılabilmesi ve yıllık kotanın kullanım düzeyi taşıyıcının güzergâh seçimini ve operasyon planını etkileyebilmektedir. Özellikle planlanan güzergâha ilişkin belgenin temin edilememesi veya kotanın tükenmesi durumunda taşımanın alternatif bir güzergâha yönlendirilmesi, farklı bir taşıma modunun tercih edilmesi ya da operasyonun belgenin temin edilmesine kadar ertelenmesi söz konusu olabilir. Bu durum taşıma mesafesinin ve teslim süresinin uzamasına, araç ve sürücü kullanım planlarının değişmesine ve ilave operasyonel maliyetlerin ortaya çıkmasına neden olabilecek bir risk alanı oluşturmaktadır (Ülengin vd., 2015; Çekyay vd., 2017).

Geçiş belgesi ve transit izinlerinden kaynaklanan risklerin azaltılmasında belge kotalarının kullanım düzeyinin düzenli izlenmesi, taşıma işletmelerinin alternatif güzergâhları önceden planlaması ve taşıma yapılacak ülkelerde geçerli belge türlerinin operasyon öncesinde kontrol edilmesi önem taşımaktadır. Bunun yanında ülkeler arasında gerçekleştirilen ikili karayolu taşımacılığı görüşmeleri yoluyla kotaların taşıma talebiyle daha uyumlu hale getirilmesi ve çok taraflı izin sistemlerinin etkin kullanılması, belge kaynaklı operasyonel belirsizliklerin azaltılmasına katkı sağlayabilir. Türkiye'de geçiş belgelerinin kullanım oranları ve kota durumları dikkate alınarak belgelerin dönemsel biçimde kullanıma açılması da mevcut kapasitenin yönetilmesine yönelik uygulamalardan biridir (T.C. Ulaştırma ve Altyapı Bakanlığı, 2025).

D. Mevzuat ve İdari Uygulama Farklılıklarından Kaynaklanan Riskler

Uluslararası karayolu taşımacılığının birden fazla ülkenin hukuki ve idari sistemi içerisinde gerçekleştirilmesi, taşıyıcıların güzergâh boyunca farklı sınır ve gümrük uygulamalarıyla karşılaşmasına neden olabilmektedir. Sınır kapılarının çalışma saatleri, uygulanacak kontrol prosedürleri, belge ve formalite gereklilikleri ile farklı kamu kurumlarının işlem yöntemleri ülkeler arasında değişiklik gösterebilir. Bu farklılıklar taşıma operasyonunun önceden planlanmasını güçleştirebildiği gibi, taşıyıcının bir ülkede sorunsuz biçimde yürüttüğü bir işlemin başka bir sınır geçişinde ek işlem veya kontrol gerektirmesi sonucunu doğurabilir. WTO Ticaretin Kolaylaştırılması Anlaşması'nın ortak sınıra sahip ülkeler arasında çalışma gün ve saatlerinin, prosedürlerin ve formalitelerin uyumlaştırılmasını, ortak kontrollerin geliştirilmesini ve tek duraklı sınır kapılarının kurulmasını teşvik etmesi, sınır yönetimindeki idari farklılıkların azaltılmasının ticaretin kolaylaştırılması açısından önemini göstermektedir (WTO, 2017).

Karayolu taşımacılığındaki sınır prosedürlerinin uyumlaştırılması gerekliliği UNECE düzenlemelerinde de açık biçimde görülmektedir. Sınır Kontrollerinin Uyumlaştırılması Uluslararası Sözleşmesi'nin Karayolu Taşımacılığına ilişkin Ek 8'i, profesyonel sürücülere yönelik işlemlerin kolaylaştırılması, özellikle acil ve bozulabilir yüklerde sınır geçiş süreçlerinin hızlandırılması ve karayolu taşıtlarına ilişkin kontrollerde uyumlaştırılmış teknik hükümlerin geliştirilmesi gibi düzenlemeler içermektedir. Bu çerçeve, uluslararası karayolu taşımacılığında yalnızca sınır kapısının fiziksel kapasitesinin değil, sınırın iki tarafındaki işlem ve kontrol uygulamalarının birbirleriyle uyumunun da taşıma performansı açısından önemli olduğunu göstermektedir (UNECE, 1982).

İdari işlem ve prosedürlerin taşıma ve ticaret maliyetleri üzerindeki etkisi akademik ve kurumsal çalışmalarda da ortaya konulmaktadır. Moïsé vd. (2011), OECD ülkelerine yönelik ticaretin kolaylaştırılması göstergelerini inceleyerek özellikle bilgiye erişim, formaliteler ve prosedürler, ön karar uygulamaları ve kurumlar arası iş birliğinin ticaret hacmi ve ticaret maliyetleri açısından önemli alanlar olduğunu belirlemiştir. Çalışmada ticaretin kolaylaştırılmasına ilişkin göstergelerin birlikte iyileştirilmesinin ticaret maliyetlerinde yaklaşık %10'a ulaşabilecek bir azaltma potansiyeli taşıdığı tahmin edilmiştir. Bu sonuç doğrudan karayolu sınır kapılarına özgü olmamakla birlikte, idari süreçlerin sadeleştirilmesi ve kurumlar arası uygulamaların uyumlaştırılmasının uluslararası ticaret maliyetleri üzerindeki önemini göstermektedir (Moïsé vd., 2011).

Mevzuat ve idari uygulama farklılıklarının taşıyıcı açısından temel sonucu, sınır geçişinde ihtiyaç duyulacak süre ve işlemlerin öngörülebilirliğinin azalmasıdır. Liu ve Yue (2013), 96 ülkenin gümrükleme sürelerine ilişkin verileri kullanarak etkin olmayan gümrük ve idari prosedürlerden kaynaklanan zaman kayıplarının uluslararası ticaret üzerinde tarife dışı bir engel niteliği taşıyabildiğini göstermiştir. Özellikle zaman hassas ve bozulabilir ürünlerde sınırdaki gecikmelerin ürün kalitesi, fiyatı ve ticaret akışları üzerinde olumsuz sonuçlar oluşturabileceği belirlenmiştir. Dolayısıyla uluslararası karayolu taşımacılığında mevzuatın şeffaf ve erişilebilir olması, sınır prosedürlerinin mümkün olduğunca uyumlaştırılması ve uygulama değişikliklerinin taşıyıcılara zamanında aktarılması, bu risklerin azaltılması ve teslim sürelerinin daha öngörülebilir hale getirilmesi açısından önem taşımaktadır (Liu & Yue, 2013; WTO, 2017).

E. Politik, Jeopolitik ve Güvenlik Riskleri

Uluslararası karayolu taşımacılığı, taşıma güzergâhının birden fazla ülkenin siyasi, hukuki ve güvenlik ortamından etkilenmesi nedeniyle dış çevre kaynaklı risklere açıktır. Ülkeler arasında ortaya çıkan siyasi gerilimler, bölgesel çatışmalar, yaptırımlar, sınırların geçici olarak kapatılması veya sınır geçişlerine yönelik ilave güvenlik tedbirleri, önceden planlanan taşıma operasyonlarının aksamasına neden olabilmektedir. Bu tür riskler taşıyıcı işletmenin doğrudan kontrol alanı dışında ortaya çıkmakla birlikte güzergâh seçimi, transit süre, sınır geçişi ve teslimat planlaması üzerinde doğrudan sonuçlar doğurabilmektedir. Politik veya güvenlik kaynaklı bir gelişmenin sınır kapısının kullanımını kısıtlaması durumunda taşımanın alternatif bir sınır kapısına veya güzergâha yönlendirilmesi gerekebilmekte, bu durum taşıma mesafesi ve süresinin artmasına neden olabilmektedir.

Jeopolitik belirsizliklerin uluslararası yük taşımacılığı üzerindeki etkisi ampirik çalışmalarda da ortaya konulmaktadır. Liu ve Fu (2024), jeopolitik riskin hava, deniz ve karayolu uluslararası yük taşımacılığı

üzerindeki etkisini ayrı ayrı incelemiş ve jeopolitik riskteki artışın üç taşıma türünde de yük hacmi üzerinde negatif ve istatistiksel olarak anlamlı etkiye sahip olduğunu belirlemiştir. Araştırmada karayolu taşımacılığının jeopolitik riske karşı denizyolu ve havayoluna kıyasla daha düşük esneklik gösterdiği belirlenmekle birlikte, etkinin karayolu yük taşımacılığında da negatif olduğu görülmüştür. Bu sonuç, uluslararası karayolu taşımacılığının coğrafi esneklik ve alternatif güzergâh kullanımı gibi bazı avantajlara sahip olmasına rağmen jeopolitik gelişmelerden tamamen bağımsız olmadığını göstermektedir (Liu & Fu, 2024).

Sınır kapılarında güvenliğin artırılmasına yönelik uygulamalar ise uluslararası taşımacılık açısından çift yönlü bir yapı taşımaktadır. Güvenlik kontrolleri yasa dışı ticaretin, kaçakçılığın ve diğer güvenlik tehditlerinin önlenmesi açısından gerekli olmakla birlikte kontrol yoğunluğunun artması sınır geçiş sürelerini ve kuyrukları etkileyebilmektedir. Hwang vd. (2015), ABD-Kanada arasındaki karayolu yük ağını sınır geçişleri üzerinden inceleyerek sınır kapılarının yük akışındaki kritik rolünü ortaya koymuştur. Çalışmada bir sınır geçiş noktasının kapanmasının veya kapasitesinin azalmasının yük hareketlerinin diğer sınır kapılarına yönelmesine ve ağın farklı bölümlerinde ikincil etkilerin ortaya çıkmasına neden olabileceği belirtilmektedir. Araştırmada ayrıca sınır güvenliği önlemleri ve denetimlerden kaynaklanan gecikmelerin sınır kapılarındaki yoğunluk açısından önemli olduğu vurgulanmaktadır (Hwang vd., 2015).

Bu nedenle sınır güvenliği ile ticaretin ve taşımacılığın kolaylaştırılması arasında dengeli bir yaklaşımın oluşturulması önem taşımaktadır. OSCE ve UNECE tarafından hazırlanan sınır geçişleri iyi uygulamalar el kitabında da sınır güvenliği ile ticaretin kolaylaştırılması birbirini dışlayan amaçlar olarak değil, etkin sınır yönetiminin birlikte yürütülmesi gereken unsurları olarak ele alınmaktadır (OSCE & UNECE, 2012). Güvenilir taşıyıcılara yönelik kolaylaştırılmış uygulamalar bu yaklaşımın örneklerinden biridir. Bradbury (2010), ABD-Kanada sınırında güvenlik ve ticaretin kolaylaştırılması amacıyla uygulanan Free and Secure Trade programının incelenen beş büyük sınır geçiş noktasının dördünde ortalama bekleme süresini azalttığını belirlemiştir. Bu sonuç, risk temelli ve önceden doğrulanmış taşıyıcı uygulamalarının güvenlik gerekliliklerinden vazgeçmeden sınır geçişlerinin daha etkin yürütülmesine katkı sağlayabileceğini göstermektedir (Bradbury, 2010).

Politik, jeopolitik ve güvenlik risklerinin taşıyıcı işletmeler açısından temel özelliği ortaya çıkma zamanlarının ve etkilerinin tam olarak öngörülememesidir. Bu nedenle söz konusu risklerin tamamen ortadan kaldırılmasından çok, etkilerinin azaltılmasına yönelik operasyonel hazırlık yapılması önem taşımaktadır. Alternatif sınır kapılarının ve güzergâhların önceden belirlenmesi, güzergâh üzerindeki siyasi ve güvenlik gelişmelerinin izlenmesi, sınır kapılarının çalışma durumuna ilişkin güncel bilginin operasyon planlamasına dâhil edilmesi ve yükün niteliğine uygun güvenlik prosedürlerinin taşıma öncesinde tamamlanması bu kapsamda değerlendirilebilir. Böylece beklenmeyen bir sınır kısıtlamasının teslim süresi ve taşıma maliyeti üzerinde oluşturabileceği etkilerin daha kontrollü yönetilmesi mümkün olabilir.

F. Dijital Sistem ve Koordinasyon Riskleri

Gümrük ve sınır işlemlerinin dijitalleşmesi, belge akışının hızlandırılması, bilgilerin taşıt sınır kapısına ulaşmadan önce işlenebilmesi ve farklı kamu kurumlarının aynı veriye elektronik ortamda erişebilmesi açısından önemli kolaylıklar sağlamaktadır. WTO Ticaretin Kolaylaştırılması Anlaşması'nda varış öncesi işlem, elektronik belge kullanımı ve tek pencere uygulamalarına yer verilmesi, dijital sistemlerin modern sınır yönetiminin temel araçlarından biri haline geldiğini göstermektedir (WTO, 2017). Bununla birlikte sınır işlemlerinin elektronik sistemlere giderek daha fazla bağımlı hale gelmesi, sistemlerin kesintisiz çalışması, veri doğruluğu ve farklı kurumların bilgi sistemleri arasındaki uyumun sınır geçiş performansı açısından kritik hale gelmesine neden olmaktadır.

Dijital sistemlerden beklenen etkinliğin sağlanabilmesi, yalnızca her kurumun kendi işlemlerini elektronik ortama aktarmasına bağlı değildir. Gümrük idaresi, diğer sınır kurumları, taşıyıcılar ve farklı ülkelerin idareleri tarafından kullanılan sistemlerin birbirleriyle veri alışverişi yapabilmesi gerekmektedir. Holloway (2009), elektronik sınır yönetiminin ulusal ve uluslararası düzeyde gelişebilmesi için ilgili

sistemlerin birlikte çalışabilirliğini sağlayacak ortak politika, hukuki çerçeve ve veri yapılarının gerekli olduğunu vurgulamaktadır. Çalışmada farklı veri içerikleri ve formatlarının sistemler arasında uyumsuzluk yaratabileceği ve hatalı elektronik mesajların sistem işleyişini aksatabilecek teknik sorunlara dönüşebileceği belirtilmektedir. Dolayısıyla sınır geçişlerinde dijitalleşme, veri standardizasyonu ve sistemler arası birlikte çalışabilirlik sağlanmadığında yeni operasyonel riskler oluşturabilmektedir (Holloway, 2009).

Benzer şekilde WCO ve WTO tarafından hazırlanan dijital teknolojilere ilişkin çalışma raporu, gümrük idarelerinin küresel tedarik zincirindeki tek aktör olmadığını ve dijital dönüşümün diğer kamu kurumları, özel sektör ve lojistik sistemleriyle bağlantılı biçimde ele alınması gerektiğini belirtmektedir. Raporda veri kalitesinin geliştirilmesi, sistemler arası bağlantının güçlendirilmesi ve farklı paydaşlar arasında bilgi paylaşımının artırılması gümrük süreçlerinin etkinleştirilmesi açısından önemli unsurlar olarak değerlendirilmektedir. Ayrıca dijital teknolojilerin sınır yönetiminde kullanılmasının yalnızca teknolojik yatırım meselesi olmadığı, iş süreçlerinin yeniden düzenlenmesi, paydaşlar arasında iş birliği kurulması ve uygun veri yönetimi yapılarının oluşturulmasını gerektirdiği vurgulanmaktadır (WCO & WTO, 2022).

Uluslararası karayolu taşımacılığı açısından dijital sistemlerde meydana gelebilecek kesintiler, veri aktarım hataları veya kurumların elektronik sistemleri arasındaki uyumsuzluklar sınırda tamamlanması gereken işlemlerin gecikmesine neden olabilir. Özellikle aynı sevkiyata ilişkin bilgilerin farklı sistemlerde farklı formatlarda tutulması, elektronik belgenin bir kurum tarafından tanınıp diğer kurum tarafından doğrulanamaması veya sınırın iki tarafındaki sistemlerin birbirleriyle veri paylaşamaması, dijitalleşmenin sağladığı zaman avantajını azaltabilir. Bu nedenle dijital sınır yönetiminde temel amaç yalnızca kâğıt belgelerin elektronik ortama taşınması değil, verinin bir kez üretilerek ilgili kurumlar arasında güvenli, doğru ve yeniden kullanılabilir biçimde paylaşılabildiği bütünleşik bir yapı oluşturmaktır. WCO ve WTO'nun dijital sınır yönetimine ilişkin yaklaşımı da teknoloji kullanımının veri paylaşımı, risk yönetimi ve kurumlar arası iş birliğiyle birlikte ele alınması gerektiğini ortaya koymaktadır (WCO & WTO, 2022).

Bu çerçevede dijital sistem ve koordinasyon risklerinin azaltılması için sınır kapılarında kullanılan bilgi sistemlerinin teknik sürekliliğinin sağlanması, ortak veri standartlarının benimsenmesi, elektronik belgelerin karşılıklı tanınması, kurumlar arasında gerçek zamanlı veri paylaşımının geliştirilmesi ve sistem kesintilerine karşı alternatif işlem prosedürlerinin oluşturulması önem taşımaktadır. Dijitalleşmenin sınır geçişlerindeki riskleri azaltıcı etkisi, ancak teknolojik altyapının kurumlar arası ve ülkeler arası koordinasyonla desteklenmesi halinde tam olarak ortaya çıkabilir. Aksi halde dijital sistemler işlemleri hızlandıran araçlar olmalarının yanında sınır geçiş sürecinde yeni teknik ve operasyonel bağımlılıklar da oluşturabilir (Holloway, 2009; WCO & WTO, 2022).

V. SONUÇLAR VE ÖNERİLER

Bu çalışma, uluslararası karayolu taşımacılığında sınır geçişlerinden kaynaklanan riskleri belirlemek ve bu risklerin taşıma operasyonları üzerindeki etkilerini kavramsal olarak değerlendirmek amacıyla gerçekleştirilmiştir. Yapılan inceleme sonucunda sınır geçişlerinin yalnızca gümrük işlemlerinin tamamlandığı idari noktalar olmadığı, aynı zamanda taşıma süresinin, güzergâh planlamasının, araç ve sürücü kullanımının ve taşıma maliyetlerinin doğrudan etkilenebildiği kritik operasyon alanları olduğu görülmektedir. Uluslararası karayolu taşımalarının birden fazla ülkenin mevzuatına, sınır yönetimi uygulamalarına ve ulaştırma politikalarına bağlı olarak gerçekleştirilmesi, taşıyıcı işletmeler açısından sınır geçişlerini önemli bir belirsizlik ve risk kaynağı haline getirmektedir.

Çalışma kapsamında sınır geçişlerinden kaynaklanan riskler gümrük ve belge kaynaklı riskler, bekleme ve sınır kapısı yoğunluğu riskleri, geçiş belgesi ve transit izinlerinden kaynaklanan riskler, mevzuat ve idari uygulama farklılıkları, politik, jeopolitik ve güvenlik riskleri ile dijital sistem ve koordinasyon riskleri olmak üzere altı temel başlık altında değerlendirilmiştir. Bu risk alanlarının her biri farklı nedenlerden ortaya çıkmakla birlikte taşıma operasyonu üzerindeki sonuçları açısından birbirleriyle yakından ilişkilidir. Özellikle sınır kapılarındaki bekleme ve işlem sürelerinin artmasının taşıma maliyetlerine yansıdığı, sınır geçişlerinin karayolu yük taşımacılığında ölçülebilir bir maliyet unsuru oluşturabildiği daha önce yapılan ampirik çalışmalar tarafından da ortaya konulmaktadır (Roberts vd., 2014; Avetisyan vd., 2015; Iimi, 2023).

Çalışmanın temel çıkarımlarından biri, sınır geçiş risklerinin çoğu zaman birbirinden bağımsız biçimde ortaya çıkmamasıdır. Örneğin belgelerde meydana gelen bir eksiklik veya uyumsuzluk ilave kontrol gerektirebilmekte, kontrol süresinin uzaması sınır kapısında beklemeye, bekleme süresindeki artış ise teslimat planının bozulmasına ve ilave operasyonel maliyetlere neden olabilmektedir. Benzer şekilde geçiş belgesi kotalarında yaşanan sorunlar taşıyıcının planlanan güzergâhı kullanmasını güçleştirebilmekte ve taşıma planının değiştirilmesini gerektirebilmektedir. Türkiye'nin uluslararası karayolu taşımacılığı üzerine yapılan çalışmalar da sınır kapısı kapasitesi ve geçiş belgesi kısıtlarının taşıma ve dış ticaret performansı açısından önem taşıdığını göstermektedir (Ülengin vd., 2015; Çekyay vd., 2017; Ülengin vd., 2021). Bu nedenle sınır geçiş risklerinin tek tek işlem sorunları şeklinde değil, birbirini tetikleyebilen operasyonel riskler bütünü olarak değerlendirilmesi daha uygun görünmektedir.

Sınır geçişlerinden kaynaklanan risklerin azaltılmasında yalnızca fiziksel sınır kapısı kapasitesinin artırılması yeterli değildir. Gümrük ve taşıma belgelerinin sınır geçişinden önce kontrol edilmesi, elektronik belge ve veri paylaşımının geliştirilmesi, sınır kurumları arasındaki koordinasyonun güçlendirilmesi, gümrük işlemlerinde risk temelli kontrol mekanizmalarının kullanılması ve ortak sınıra sahip ülkeler arasında prosedürlerin mümkün olduğunca uyumlaştırılması önem taşımaktadır. WTO Ticaretin Kolaylaştırılması Anlaşması'nda yer verilen varış öncesi işlem, tek pencere, risk yönetimi ve sınır kurumları arasında iş birliği gibi uygulamalar bu yaklaşımı desteklemektedir (WTO, 2017). Bununla birlikte dijitalleşmenin tek başına yeterli olmadığı, kullanılan sistemlerin birlikte çalışabilirliği, veri standartlarının uyumu ve kurumlar arası bilgi paylaşımının da sağlanması gerektiği görülmektedir (Holloway, 2009; WCO & WTO, 2022).

Türkiye açısından sınır geçiş risklerinin yönetimi, ülkenin Avrupa ve Asya arasındaki uluslararası karayolu taşımacılığındaki konumu nedeniyle ayrıca önem taşımaktadır. Kapıkule üzerine gerçekleştirilen çalışma, sınır kapısındaki işlem kapasitesi ve süreç organizasyonunun uzun araç kuyruklarının ve hizmet sürelerinin azaltılmasında belirleyici olabileceğini göstermektedir (Ülengin vd., 2021). Bunun yanında geçiş belgesi ve kota uygulamalarının Türkiye'nin karayoluyla dış ticareti üzerinde oluşturabileceği etkiler, sınır geçiş performansının yalnızca sınır kapısının fiziksel işleyişiyle sınırlı olmadığını ortaya koymaktadır (Ülengin vd., 2015; Çekyay vd., 2017). Bu nedenle sınır kapısı altyapısının geliştirilmesi, geçiş belgesi süreçlerinin taşıma talebiyle uyumlu hale getirilmesi, dijital sınır uygulamalarının yaygınlaştırılması ve komşu ülkelerle idari koordinasyonun güçlendirilmesi birlikte ele alınması gereken alanlar olarak değerlendirilebilir.

Bu çalışmanın temel sınırlılığı, sınır geçiş risklerinin akademik çalışmalar ve kurumsal dokümanlar üzerinden kavramsal olarak değerlendirilmiş olması ve taşıyıcı işletmeler, sürücüler veya sınır kapılarında görev yapan aktörlerden doğrudan birincil veri elde edilmemesidir. Bu nedenle çalışmada belirlenen risklerin göreli önem düzeyleri veya taşıma işletmeleri üzerindeki etkilerinin büyüklüğü karşılaştırılmamıştır. Gelecek araştırmalarda uluslararası karayolu taşıma işletmeleri ve sürücülerle gerçekleştirilecek görüşmeler yoluyla sınır geçişlerinde karşılaşılan risklerin uygulamadaki karşılıkları incelenebilir. Farklı sınır kapılarının karşılaştırılması, sınır geçiş sürelerinin gerçek operasyon verileri üzerinden analiz edilmesi veya belirlenen risklerin uzman değerlendirmelerine dayalı olarak önceliklendirilmesi de konunun ampirik olarak geliştirilmesine katkı sağlayabilir. Bu yönüyle çalışma, uluslararası karayolu taşımacılığında sınır geçiş risklerini bütüncül bir çerçevede sınıflandırarak ileride gerçekleştirilecek uygulamalı araştırmalar için kavramsal bir temel sunmaktadır.

KAYNAKLAR

[1] Avetisyan, M., Heatwole, N., Rose, A., & Roberts, B. (2015). Competitiveness and macroeconomic impacts of reduced wait times at U.S. land freight border crossings. *Transportation Research Part A: Policy and Practice, 78*, 84–101. https://doi.org/10.1016/j.tra.2015.04.027

[2] Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal, 9*(2), 27–40.

[3] Bradbury, S. L. (2010). An assessment of the free and secure trade (FAST) program along the Canada–US border. *Transport Policy, 17*(6), 367–380. https://doi.org/10.1016/j.tranpol.2010.04.003

[4] Caballini, C., & Benzi, M. (2023). Fast Corridors: Innovative customs processes and technology to increase supply chain competitiveness. The case of IKEA Italy. *Transportation Research Interdisciplinary Perspectives, 21*, 100909. https://doi.org/10.1016/j.trip.2023.100909

[5] Çekyay, B., Toktaş Palut, P., Kabak, Ö., Ülengin, F., Özaydın, Ö., & Ülengin, B. (2017). Analysis of the impact of bilateral and transit quotas on Turkey's international trade by road transport: An integrated maximum flow and gravity model approach. *Research in Transportation Economics, 66*, 70–77. https://doi.org/10.1016/j.retrec.2017.04.006
[6] Holloway, S. (2009). The transition from eCustoms to eBorder management. *World Customs Journal, 3*(1), 13–25. https://doi.org/10.55596/001c.91351
[7] Hummels, D. L., & Schaur, G. (2013). Time as a trade barrier. *American Economic Review, 103*(7), 2935–2959. https://doi.org/10.1257/aer.103.7.2935
[8] Hwang, H., Park, J. Y., Kwon, C., Friedman, K., Attard, N., Chang, S. H., & Wells, S. (2015). The ties that bind: Bi-national trade implications of the US and Canada using bi-national freight movement network via border crossings. *Research in Transportation Business & Management, 16*, 32–49. https://doi.org/10.1016/j.rtbm.2015.08.002
[9] Iimi, A. (2023). Estimating road freight transport costs in Eastern Europe and Central Asia using large shipping data. World Bank.
[10] Liu, K., & Fu, Q. (2024). How does geopolitical risk affect international freight? *Journal of Air Transport Management, 118*, 102614. https://doi.org/10.1016/j.jairtraman.2024.102614
[11] Liu, L., & Yue, C. (2013). Investigating the impacts of time delays on trade. *Food Policy, 39*, 108–114. https://doi.org/10.1016/j.foodpol.2013.01.001
[12] Moïsé, E., Orliac, T., & Minor, P. (2011). *Trade facilitation indicators: The impact on trade costs* (OECD Trade Policy Papers No. 118). OECD Publishing. https://doi.org/10.1787/5kg6nk654hmr-en
[13] Organization for Security and Co-operation in Europe, & United Nations Economic Commission for Europe. (2012). *Handbook of best practices at border crossings: A trade and transport facilitation perspective*. Organization for Security and Co-operation in Europe.
[14] Roberts, B., Rose, A., Heatwole, N., Wei, D., Avetisyan, M., Chan, O., & Maya, I. (2014). The impact on the US economy of changes in wait times at ports of entry. *Transport Policy, 35*, 162–175. https://doi.org/10.1016/j.tranpol.2014.05.010
[15] Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104*, 333–339. https://doi.org/10.1016/j.jbusres.2019.07.039
[16] T.C. Ticaret Bakanlığı. (2026). *Dış ticaret lojistiği 2026*. Uluslararası Hizmet Ticareti Genel Müdürlüğü.
[17] T.C. Ulaştırma ve Altyapı Bakanlığı, Ulaştırma Hizmetleri Düzenleme Genel Müdürlüğü. (2025, 25 Aralık). *2026 yılı geçiş belgeleri hakkında bilgilendirme*.
[18] Ülengin, F., Çekyay, B., Işık, M., Özaydın, Ö., Kabak, Ö., Toktaş-Palut, P., Önsel Ekici, Ş., & Topcu, Y. I. (2021). A simulation-based approach for improving the largest border crossing between Europe and Turkey. *Transport Policy, 114*, 350–363. https://doi.org/10.1016/j.tranpol.2021.10.011
[19] United Nations Economic Commission for Europe. (2008). New Annex 8 (Annex on Road Transport) to the International Convention on the Harmonization of Frontier Controls of Goods, 1982. United Nations Economic Commission for Europe.
[20] United Nations Economic Commission for Europe. (2024). *TIR handbook: 2024 edition*. United Nations.
[21] Volpe Martincus, C., Carballo, J., & Graziano, A. (2015). Customs. *Journal of International Economics, 96*(1), 119–137. https://doi.org/10.1016/j.jinteco.2015.01.011
[22] World Customs Organization, & World Trade Organization. (2022). *WCO/WTO study report on disruptive technologies*. World Customs Organization.
[23] World Trade Organization. (2017). *Agreement on Trade Facilitation*.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Theoretical investigation and molecular docking analysis of a phenolic Schiff base targeting acetylcholinesterase

Sevgi Kansız [*1]

[1]*Department of Fundamental Sciences, Samsun University, Türkiye*

[*]*(sevgi.kansiz@samsun.edu.tr) Email of the corresponding author*

Abstract **–** Schiff bases containing phenolic moieties have attracted considerable attention because of their diverse chemical properties and potential biological activities, including their relevance to neurodegenerative disorders such as Alzheimer's disease. In this study, the electronic structure, molecular properties, and potential acetylcholinesterase (AChE) inhibitory activity of a phenolic Schiff base, $C_{14}H_{12}BrNO$, were investigated using an integrated computational approach. Density Functional Theory (DFT) calculations were performed at the B3LYP/6-311+G(d,p) level using the Gaussian 16W program to investigate the optimized molecular structure and electronic properties of the compound. Molecular docking simulations against human AChE (PDB ID: 4EY7) using AutoDock Vina revealed a strong predicted binding affinity of −10.14 kcal mol^{-1}. The ligand was stabilized within the active-site region through hydrogen bonding with HIS447, together with π–alkyl and π–π stacked interactions involving TRP286 and a π–π T-shaped interaction with PHE338. In addition, in silico physicochemical, pharmacokinetic, and drug-likeness assessments indicated high gastrointestinal absorption and predicted blood–brain barrier permeability, while the compound satisfied Lipinski's rule of five without violations. These computational findings provide useful insights into the structural, electronic, and molecular recognition properties of the phenolic Schiff base and support its further investigation as an AChE-targeting compound.

Keywords – Schiff base, phenol, DFT, molecular docking, ADME, drug-likeness

I. INTRODUCTION

Phenolic Schiff bases have attracted considerable attention in molecular and medicinal chemistry because of their versatile structural features and diverse biological properties [1-3]. The presence of the azomethine (–CH=N–) linkage together with phenolic functionality provides favorable sites for hydrogen bonding and other non-covalent interactions, which can play an important role in molecular recognition and biological activity [4-6]. In addition, the electronic characteristics of Schiff bases can be effectively investigated through density functional theory (DFT), molecular electrostatic potential (MEP), and frontier molecular orbital analyses, providing valuable information on their reactivity and interaction potential [7-9]. Acetylcholinesterase (AChE) is an important biological target in the investigation of compounds with potential neuroactive properties, making the study of ligand–AChE interactions particularly relevant [10-12]. In the present study, the phenolic Schiff base $C_{14}H_{12}BrNO$ was investigated using DFT calculations, molecular docking, and SwissADME analyses. The molecular structure was optimized at the B3LYP/6-311+G(d,p) level, while MEP analysis was employed to characterize the electrostatic features of the molecule. Molecular docking was subsequently performed against acetylcholinesterase to examine its binding mode and possible inhibitory potential. Furthermore, SwissADME calculations were carried out to assess the physicochemical, pharmacokinetic, and drug-likeness properties of the compound.

II. MATERIALS AND METHOD

Molecular docking calculations for $C_{14}H_{12}BrNO$ were carried out with AutoDock Vina [13]. Prior to the docking procedure, the molecular structure was geometrically optimized using Gaussian 16W [14] in conjunction with the GaussView6 graphical interface [15]. The optimization was performed at the DFT/B3LYP/6-311+G(d,p) level of theory (Fig. 1). The optimized ligand was subsequently converted into the required PDB and PDBQT file formats using Discovery Studio Visualizer 4.0 (DSV 4.0) [16]. For the receptor, the crystal structure of the antifungal target with PDB ID 4EY7 was retrieved from the RCSB Protein Data Bank [17] and employed in the docking calculations. The SwissADME web-based platform [18] was utilized to evaluate the compound in terms of its physicochemical characteristics, pharmacokinetic behavior, drug-likeness, and related molecular properties.

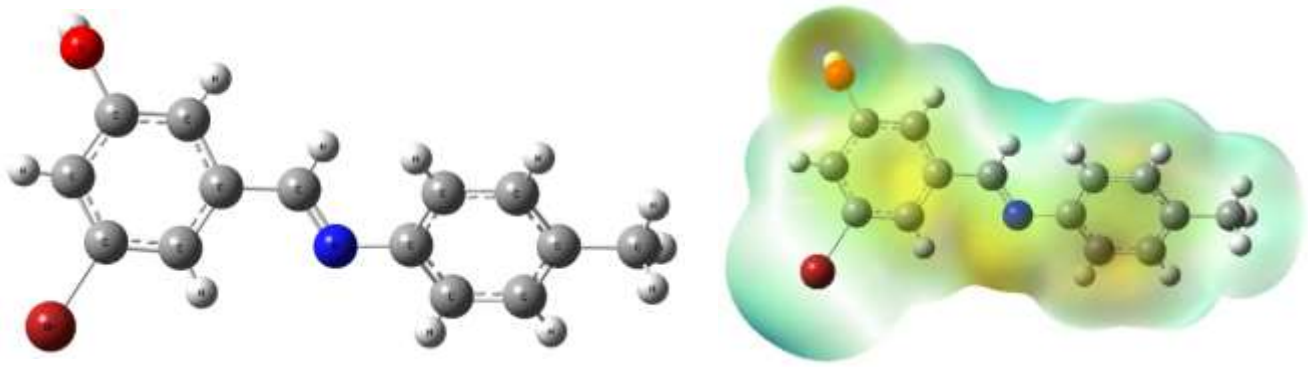


Fig. 1 The optimized geometry and MEP map of $C_{14}H_{12}BrNO$

III. RESULTS AND DISCUSSION

The optimized structure of $C_{14}H_{12}BrNO$ exhibits a C=N bond length of 1.266 Å, while the C–N and C–O distances are calculated as 1.432 and 1.430 Å, respectively. The C–C distances within the aromatic ring range from 1.36 to 1.42 Å, with an average value of approximately 1.39 Å, consistent with the expected bond-length characteristics of an aromatic system [19-24]. The C–Br bond length of 1.910 Å is also in accordance with the structural features expected for an aryl bromide. The MEP surface of $C_{14}H_{12}BrNO$ reveals a predominantly yellowish region over the aromatic ring, indicating a relatively moderate electrostatic potential in this part of the molecule. A pronounced blue region is observed around the hydrogen atom of the phenolic OH group, reflecting its positive electrostatic character and potential hydrogen-bond donor behavior [25-27]. Meanwhile, the imine nitrogen of the C=N group is associated with an orange-yellow region, suggesting a relatively electron-rich site that may participate in intermolecular interactions.

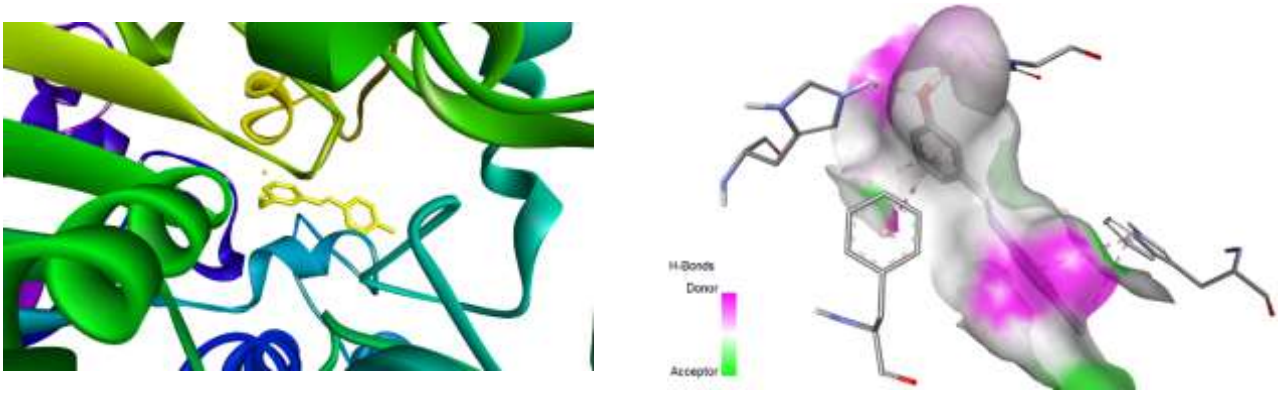


Fig. 2 The placement of $C_{14}H_{12}BrNO$ in PDB: 4EY7

Molecular docking is a widely employed computational approach for exploring ligand–protein recognition and predicting the possible binding modes and affinities of small molecules within specific biological targets [28-30]. In the present study, docking of $C_{14}H_{12}BrNO$ into human acetylcholinesterase (AChE, PDB ID: 4EY7) using AutoDock Vina resulted in a favorable binding affinity of −10.14 kcal

mol^{-1} (Table 1). As illustrated in Fig. 2, the ligand was accommodated within the active-site region of AChE and stabilized by a hydrogen bond involving HIS447, together with π–alkyl and π–π stacked interactions with TRP286 and a π–π T-shaped interaction with PHE338. These interactions suggest a favorable ligand–enzyme recognition pattern and support the potential AChE inhibitory activity of the investigated Schiff base.

Table 1. The results of the binding affinity values.

Mode	Affinity (kcal/mol)	rmsd l.b.	rmsd u.b.
1	-10.14	0.000	0.000
2	-9.813	4.009	7.497
3	-9.649	3.327	7.326
4	-9.501	2.204	2.902
5	-9.332	2.135	2.723
6	-9.312	4.144	7.739
7	-8.473	3.758	5.085
8	-8.453	1.416	2.160
9	-8.431	2.580	3.322
Number of hydrogen bonding: 1			

The SwissADME results indicate a molecular weight of 290.16 g mol^{-1}, a consensus LogP value of 4.04, and a TPSA of 32.59 Å² (Table 2). The molecule contains two hydrogen-bond acceptors and one hydrogen-bond donor and satisfies the Lipinski criteria without any violation, while its predicted gastrointestinal absorption is high. Although the compound is predicted to be BBB permeant and not to act as a P-gp substrate, its relatively high lipophilicity may contribute to its membrane permeability. The calculated bioavailability score of 0.55 further indicates a moderate predicted oral bioavailability. The BOILED-Egg model (Fig. 3) places the molecule within the yellow region, corresponding to a favorable prediction for passive gastrointestinal absorption. This location is consistent with the high GI absorption predicted by SwissADME and can be associated with the compound's relatively low TPSA and lipophilic character. The ADME profile suggests that $C_{14}H_{12}BrNO$ possesses physicochemical characteristics compatible with good membrane permeation and oral absorption, although its relatively high lipophilicity should be taken into consideration.

Table 2. Some SwissADME results for $C_{14}H_{12}BrNO$.

Physicochemical Properties		*Water Solubility*	
Formula	$C_{14}H_{12}BrNO$	Log S (ESOL)	-4.50
Molecular weight	290.16 g/mol	Solubility	9.20e-03 mg/ml ; 3.17e-05 mol/l
Num. heavy atoms	17	Class	Moderately soluble
Num. arom. heavy atoms	12	Log S (Ali)	-4.30
Fraction Csp3	0.07	Solubility	1.44e-02 mg/ml ; 4.97e-05 mol/l
Num. rotatable bonds	2	Class	Moderately soluble
Num. H-bond acceptors	2	Log S (SILICOS-IT)	-5.67
Num. H-bond donors	1	Solubility	6.28e-04 mg/ml ; 2.16e-06 mol/l
Molar Refractivity	74.83	Class	Moderately soluble
TPSA	32.59 Å²	*Medicinal Chemistry*	
Lipophilicity		PAINS	0 alert
Log Po/w (XLOGP3)	3.92	Brenk	1 alert: imine_1
Log Po/w (WLOGP)	4.21	Leadlikeness	No; 1 violation: XLOGP3>3.5
Log Po/w (MLOGP)	3.52	Synthetic accessibility	2.48
Log Po/w (SILICOS-IT)	4.52	*Druglikeness*	
Consensus Log Po/w	4.04	Lipinski	Yes; 0 violation
Pharmacokinetics		Ghose	Yes
GI absorption	High	Veber	Yes
BBB permeant	Yes	Egan	Yes
P-gp substrate	No	Muegge	Yes
Log Kp (skin permeation)	-5.29 cm/s	Bioavailability Score	0.55

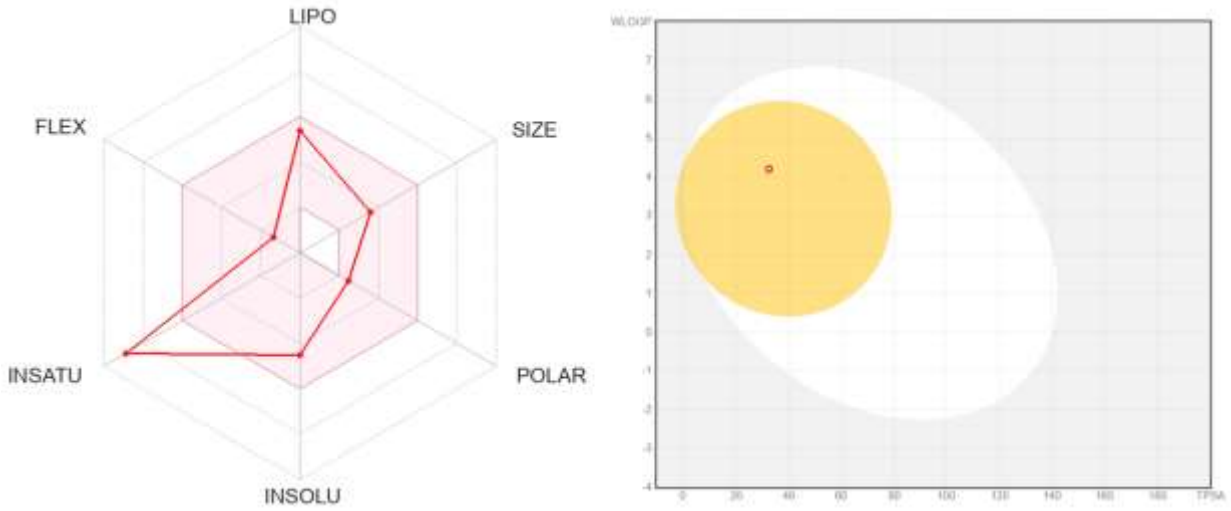


Fig. 3 The radar and boiled egg model of $C_{14}H_{12}BrNO$.

IV. CONCLUSION

In conclusion, the computational investigation of $C_{14}H_{12}BrNO$ provided consistent information on its structural, electronic, binding, and pharmacokinetic characteristics. DFT calculations confirmed a stable optimized molecular geometry, while MEP analysis identified the phenolic OH group and imine nitrogen as prominent electrostatic regions. Molecular docking revealed a favorable interaction with human AChE, with a binding affinity of −10.14 kcal mol^{-1}, supported by hydrogen-bonding and aromatic interactions within the active site. The SwissADME and BOILED-Egg analyses further indicated high gastrointestinal absorption, predicted BBB permeation, and no Lipinski violations. These findings suggest that the compound may represent a promising scaffold for further investigation as a potential AChE inhibitor.

REFERENCES

[1] I. Selatnia, A. Sid, H. Lgaz, A. Bouhraoua, T. Roisnel, O.M.A. Khamaysa, S. Kansız, K. Metwally, A.G. Soliman, H.-S. Lee, and S. Boudiba, "Synthesis, structural characterization, and biological evaluation of novel iodophenyl-anthracene Schiff bases: Insights from X-ray crystallography, DFT, antibacterial, and molecular docking studies," *J. Mol. Struct.*, vol. 1350, pp. 144066, 2025.

[2] S. Kansiz, "Investigating the structural and surface properties of copper and nickel complexes fluorine-based salen type Schiff bases," *J. Struct. Chem*., vol. 64(12), pp. 2295-2310, 2023.

[3] O. Şimşek, Ş. Atalay, M. Taşdoğan, E. Ağar, H. Bülbül, and S. Kansız, "Synthesis, Characterization, and Hirshfeld Surface Analysis of Two Schiff Bases: (*E*)-3-(((4-iodophenyl) imino) methyl)benzene-1,2-diol and (*E*)-3-(((3-bromo-4-methylphenyl)imino)methyl)-benzene-1,2-diol," *J. Struct. Chem.*, vol. *67*(4), pp. 827-840, 2026.

[4] S. Yeşilbağ, S. Kansız, N. Dege, and E. Ağar, "Structural investigation and hirshfeld surface analysis of two [ONO]-Type schiff bases," *J. Struct. Chem*., vol. 65(3), pp.464-477, 2024.

[5] P. Şen, and S. Kansız, "Crystal Structure, Spectroscopic and Computational Studies of N2O2 Type Salen-Based Schiff Base and Its Di-boron Complex," *J. Chem. Crystallogr.*, vol. *52*(1), pp. 1-16, 2022.

[6] M. K. Gumus, F. Sen, S. Kansiz, N. Dege, and E. Saif, "Crystal structure and Hirshfeld surface analysis of 2-{[(E)-(3-cyclobutyl-1H-1, 2, 4-triazol-5-yl) imino] methyl} phenol," *Acta Crystallogr. E: Crystallogr. Commun*., vol. 77(12), pp. 1267-1271, 2021.

[7] H. Kargar, M. Fallah-Mehrjardi, K. S. Munawar, E. Acar, N. Dege, and S. Kansız, "Structural dynamics of a phenoxy-bridged heteronuclear Cu (II)/Hg (II)-Salen complex: synthesis, characterization, DFT insights, and Hirshfeld surface analysis," *J. Iran. Chem. Soc.*, vol. 22(4), pp. 769-780, 2025.

[8] S. Kansız, A. Tolan, H. İçbudak, and N. Dege, "Synthesis, crystallographic structure, theoretical calculations, spectral and thermal properties of trans-diaquabis (trans-4-aminoantipyrine) cobalt (II) acesulfamate," *J. Mol. Struct*., vol. 1190, pp. 102-115, 2019.

[9] N. Arumugam, A. I. Almansour, R. S. Kumar, S. Yeswanthkumar, R. Padmanaban, Y. Arun, S. Kansız, N. Dege, T. S. Manohar, and S. Venketesh, "Design, stereoselective synthesis, computational studies and cholinesterase inhibitory activity of novel spiropyrrolidinoquinoxaline tethered indole hybrid heterocycle," *J. Mol. Struct*., 2021, vol. 1225, pp. 129165.

[10] N. Dege, A. S. Aydın, E. Ağar, S. Kansız, S. JoseKavitha, K. BalaSubramani, M. Hemamalini, and V. Rajakannan, "Synthesis, crystal structure, Hirshfeld surface analysis, In-Silico assessment of druggability and molecular docking studies of Schiff base compound," *Chem. Data Collect.*, vol. *25*, 100320, 2020.

[11] O. Simsek, M. Ashfaq, M. N. Tahir, S. Ozturk, and E. Agar, "Synthesis and charaterizations of the Schiff base derived from 2-hydroxy-5-nitrobenzaldehyde alongwith Hirshfeld surface analysis and computational study," *J. Struct. Chem.*, vol. 64(5), pp. 942-953, 2023.
[12] H. Ozsanlı, O. Sımsek, O. R. Yılmaz, C. Cıcek, E. Agar, U. Çoruh, and E. M. Vazquez-Lopez, "Synthesis, crystal structure, Hirshfeld surface analysis, supramolecular and computational investigation of the (E)-4-methoxy-2-(((2-methoxy-5-(trifluoromethyl) phenyl) imino) methyl) phenol compound," J. Struct. Chem., vol. 65(5), pp. 946-957, 2024.
[13] O. Trott, and A. J. Olson, "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading," *J. Comput. Chem*., vol. 31(2), pp. 455-461, 2010.
[14] M. J. Frisch, G. Trucks, H. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, and G. Petersson, 2009. Gaussian 09, Revision D. 01, Gaussian, Inc.: Wallingford, CT.
[15] R. Dennington, T. Keith, and J. Millam, 2009. GaussView, version 5, Semichem Inc.: Shawnee Mission, KS.
[16] https://www.3dsbiovia.com/.
[17] https://www.rcsb.org/.
[18] A. Daina, O. Michielin, and V. Zoete, "SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules," *Sci. Rep*., vol. 7(1), pp. 42717, 2017.
[19] R. Ilmi, S. Kansız, N. K. Al Rasbi, J. Husband, N. Dege, and M. S. Khan, "Synthesis, X-ray crystal structure and determination of non-covalent interactions through Hirshfeld surface analysis of a pure red emitting asymmetrical octacoordinated Sm (III) complex," *Polyhedron*, vol. 246, pp. 116673, 2023.
[20] S. Kansiz, S. Malinkin, and N. Dege, "Synthesis, crystal structure and Hirshfeld surface analysis of tetraaquabis (isonicotinamide-κN1) cobalt (II) succinate," *Acta Crystallogr. E: Crystallogr. Commun*., vol. 74(7), pp. 1026-1029, 2018.
[21] K. Chkirate, S. Kansiz, K. Karrouchi, J. T. Mague, N. Dege, and E. M. Essassi, "Crystal structure and Hirshfeld surface analysis of N-{2-[(E)-(4-methylbenzylidene) amino] phenyl}-2-(5-methyl-1-H-pyrazol-3-yl) acetamide hemihydrate," *Acta Crystallogr. E: Crystallogr. Commun*., vol. 75(2), pp. 154-158, 2019.
[22] S. Cakmak, S. Kansiz, M. Azam, C. C. Ersanlı, O. Idil, A. Veyisoglu, H. Yakan, H. Kütük, and A. Chutia, "Synthesis, structural investigation, Hirshfeld surface analysis, and biological evaluation of N-(3-Cyanothiophen-2-yl)-2-(thiophen-2-yl) acetamide," *ACS omega*, vol. 7(13), pp. 11320-11329, 2022.
[23] S. Kansız, G. Kaştaş, M. K. Gümüş, N. Dege, and R. Aydın, "Synthesis, structural characterization, and supramolecular assembly of a one-dimensional cobalt (II) coordination polymer with succinato and isonicotinamide ligands," *J. Struct. Chem*., vol. 66(4), pp. 722-732, 2025.
[24] M. K. Gümüş, S. Kansız, G. B. Tulemisova, and N. Dege, "Synthesis, Crystal Structure, and Hirshfeld Surface Analysis of 4,4A,5,6,7,8-Hexahydro-4A-Methyl-2,5,7-Triphenyl-2H-Pyrido[4,3-d][1,3] Oxazine," *J. Struct. Chem*., vol. 66(1), pp. 201-213, 2025.
[25] M. K. Alenzi, M. Azam, S. I. Al-Resayes, S. Kansız, A. S. Çamaş, D. C. Agurokpon, H. Louis, V. Kumar, N. Dege, and M. Alam, "Experimental and theoretical investigation into the design of nickel (II), copper (II), and zinc (II) complexes substituted with pyrazole ligand," *Polyhedron*, vol. 259, pp. 117073, 2024.
[26] Y. El Bakri, M. Kurbanova, A. Ahsin, S. Karthikeyan, A. Maharramov, R. Al-Salahi, S. Kansız, S. N. Mali, and M. Ashfaq, "Synthesis, crystal structure investigation, and theoretical approaches to discover potential 6-bromo-3-cyanocoumarin as a potent inhibitor MetAP (methionine aminopeptidase)2," *Chem. Phys. Impact*, vol. 8, 100477, 2024.
[27] S. Kansız, M. Azam, N. Dege, N. Ermiş, S.I. Al-Resayes, and M. Alam, "Supramolecular assembly in designing co-crystals of fumaric acid and pyrimidine/picolinate derivatives," *Green Chem. Lett. Rev*., vol. 15 (3), pp. 825-836, 2022.
[28] A.-A. B. OmarAli, A. J. M. Al-Karawi, A. A. Awad, N. Dege, S. Kansız, E. Agar, Z. A. Hussein and I. R. Mohammed, "Two new zinc(II) and mercury(II) complexes based on N,V-(cyclohexane-1,2-diylidene)bis(4-fluorobenzohydrazide): synthesis, crystal structures and antibacterial activities," *Acta Crystallogr. C: Struct. Chem*., vol. 76, pp. 476-482, 2020.
[29] O. Şimşek, S. Kansız, İ. Yılmaz, A. Çukurovalı, and M. Dinçer, "Structural characterization, DFT-based electronic analysis, Hirshfeld surface study, molecular docking, and in silico drug-likeness evaluation of N-(4-(3-methyl-3mesitylcyclobutyl)-1H-imidazol-2-yl) acetamide," *J. Mol. Struct.*, vol. 1358, pp. 145399, 2026.
[30] S. Kansız, Y. Sert, S. Meral, A. Alaman Ağar, and N. Dege, "Synthesis, Experimental and Theoretical Characterization of (E)-1-((5-nitrothiophen-2-yl) methylene)-2-Phenylhydrazine," *J. Struct. Chem.*, vol. *67*(2), pp. 291-308, 2026.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

DFT, molecular docking and ADME studies of a Schiff base compound

Sevgi Kansız [*1]

[1]*Department of Fundamental Sciences, Samsun University, Türkiye*

[*]*(sevgi.kansiz@samsun.edu.tr) Email of the corresponding author*

Abstract – Phenolic Schiff bases have received sustained interest owing to their structural versatility and diverse biological properties, particularly in the development of molecular scaffolds relevant to neurodegenerative diseases. Their ability to establish hydrogen-bonding and aromatic interactions with biologically important targets makes them attractive candidates for investigating acetylcholinesterase (AChE) recognition. In the present study, the electronic and molecular properties of the phenolic Schiff base $C_{16}H_{16}ClNO$ were investigated through a combination of Density Functional Theory (DFT), molecular electrostatic potential (MEP) mapping, molecular docking, and in silico pharmacokinetic analyses. The molecular geometry was optimized at the B3LYP/6-311+G(d,p) level using Gaussian 16W. The MEP surface revealed a pronounced positive electrostatic potential around the phenolic O–H hydrogen, whereas comparatively electron-rich regions were observed around the imine nitrogen, providing insights into the possible reactive and interaction sites of the molecule. Molecular docking against human AChE (PDB ID: 4EY7) using AutoDock Vina yielded a predicted binding affinity of −11.14 kcal mol^{-1}, corresponding to an estimated inhibition constant (K_i) of approximately 6.8 nM at 298 K. Within the active site, the ligand was involved in hydrogen bonding with SER293, π–alkyl interactions with VAL294, and π–π T-shaped interactions with TYR337 and TYR341. In silico ADME and drug-likeness analyses further indicated high gastrointestinal absorption, predicted blood–brain barrier permeability, and full compliance with Lipinski's rule of five. The combined electronic-structure and molecular docking results provide insight into the AChE recognition behavior of this chlorinated phenolic Schiff base and warrant further investigation of its biological potential.

Keywords – Schiff base, MEP, AChE, molecular docking, ADME, drug-likeness

I. INTRODUCTION

Schiff bases containing phenolic groups represent an important class of compounds owing to their readily tunable electronic structures and potential biological activities [1-3]. Their characteristic azomethine functionality can facilitate intermolecular interactions with biological macromolecules, while the phenolic group may contribute to hydrogen-bonding interactions and molecular recognition processes [4-6]. Computational approaches such as DFT and molecular electrostatic potential (MEP) analysis provide useful insights into the electronic structure, reactive regions, and interaction characteristics of these molecules [7-9]. From a biological perspective, acetylcholinesterase (AChE) is a well-established enzyme target, and computational assessment of ligand binding to its active site can provide preliminary information regarding potential inhibitory activity [10-12]. Herein, the phenolic Schiff base $C_{16}H_{16}ClNO$ was subjected to a combined computational investigation involving DFT calculations, molecular docking, and ADME/drug-likeness prediction. The electronic structure of the compound was examined using the B3LYP/6-311+G(d,p) method, and its electrostatic properties were evaluated through MEP analysis. Docking simulations with AChE were conducted to characterize the ligand–enzyme interactions and to estimate the binding potential of the compound. In parallel, SwissADME was employed to predict key physicochemical and pharmacokinetic descriptors and to evaluate the drug-like profile of the investigated molecule.

II. MATERIALS AND METHOD

The computational investigation was initiated by optimizing the $C_{16}H_{16}ClNO$ ligand at the DFT/B3LYP/6-311+G(d,p) level using Gaussian 16W [13] and GaussView6 [14]. The optimized molecular geometry is presented in Fig. 1. For subsequent docking calculations, the ligand structure was prepared in PDB and PDBQT formats with Discovery Studio Visualizer 4.0 [15]. AutoDock Vina [16] was then employed to investigate the binding behavior of the compound toward the selected antifungal protein. The three-dimensional structure of the receptor (PDB ID: 4EY7) was downloaded from the RCSB Protein Data Bank [17] and used as the docking target. Complementary in silico characterization was performed through SwissADME [18], providing information on the compound's physicochemical parameters, pharmacokinetic profile, drug-likeness, and other relevant molecular descriptors.

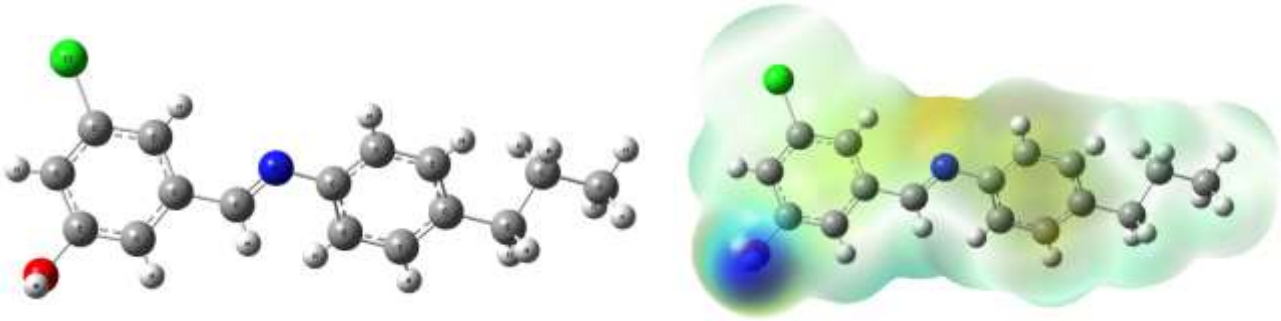


Fig. 1 The optimized geometry and MEP map of $C_{16}H_{16}ClNO$

III. RESULTS AND DISCUSSION

For $C_{16}H_{16}ClNO$, the optimized molecular geometry gives a C=N bond distance of 1.266 Å, together with C–N and C–O bond lengths of 1.432 and 1.430 Å, respectively. The aromatic C–C bonds fall within the 1.36–1.42 Å interval and display an average length of about 1.39 Å, supporting the preservation of the expected aromatic bonding pattern [19-24]. The calculated C–Cl distance is 1.760 Å, which is characteristic of the carbon–chlorine linkage in the molecule. The MEP mapping of the compound shows a yellowish distribution on the aromatic ring, corresponding to regions of intermediate electrostatic potential [28-27]. In contrast, the hydrogen atom of the phenolic OH group exhibits a distinct blue coloration, indicating positive potential and its possible role as a hydrogen-bond donor. The imine nitrogen presents an orange-to-yellow region on the MEP surface, highlighting this atom as an electrostatically differentiated site that can contribute to interactions with the biological target.

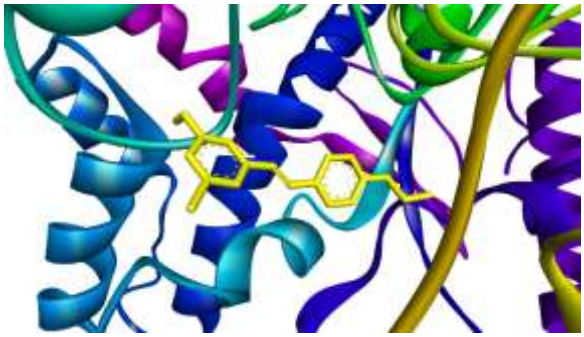


Fig. 2 The placement of $C_{16}H_{16}ClNO$ in PDB: 4EY7

Computational molecular docking is commonly applied to characterize the recognition of small molecules by biological macromolecules and to estimate their binding preferences within specific protein binding sites [28-30]. For $C_{16}H_{16}ClNO$, AutoDock Vina calculations performed against human acetylcholinesterase (AChE, PDB ID: 4EY7) produced a binding affinity of −11.14 kcal mol^{-1} (Table 1). This value corresponds to an estimated inhibition constant (Ki) of approximately 6.8 nM at 298 K. As shown in Figs. 2 and 3, the ligand is positioned favorably within the AChE active-site region and participates in hydrogen bonding with SER293, π–alkyl interactions with VAL294, and π–π T-shaped interactions with TYR337 and TYR341. The combination of these interactions contributes to the predicted stability of the ligand–AChE complex and indicates a notable binding potential toward the enzyme.

Table 1. The results of the binding affinity values of $C_{16}H_{16}ClNO$.

Mode	Affinity (kcal/mol)	rmsd l.b.	rmsd u.b.
1	-11.14	0.000	0.000
2	-10.75	1.541	2.023
3	-10.60	3.975	8.609
4	-10.30	3.840	8.301
5	-9.74	3.952	8.434
6	-9.67	2.763	4.462
7	-9.56	2.251	2.822
8	-9.02	6.235	8.591
9	-8.85	3.986	6.516
Number of hydrogen bonding: 1			

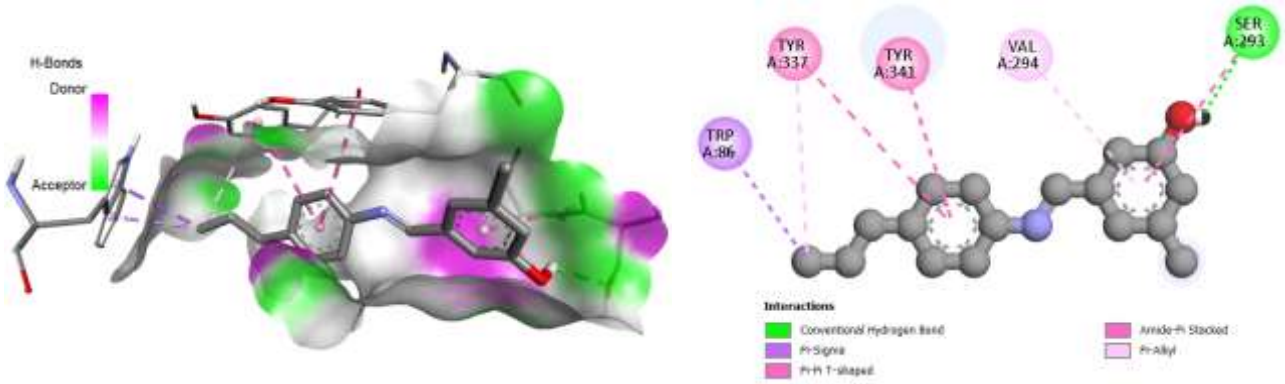


Fig. 3 3D and 2D molecular docking results of the ligand+PDB:4EY7

Table 2. Some SwissADME results for $C_{16}H_{16}ClNO$.

Physicochemical Properties		*Water Solubility*	
Formula	$C_{16}H_{16}ClNO$	Log S (ESOL)	-4.78
Molecular weight	273.76 g/mol	Solubility	4.51e-03 mg/ml ; 1.65e-05 mol/l
Num. heavy atoms	19	Class	Moderately soluble
Num. arom. heavy atoms	12	Log S (Ali)	-5.25
Fraction Csp3	0.19	Solubility	1.55e-03 mg/ml ; 5.65e-06 mol/l
Num. rotatable bonds	4	Class	Moderately soluble
Num. H-bond acceptors	2	Log S (SILICOS-IT)	-6.25
Num. H-bond donors	1	Solubility	1.56e-04 mg/ml ; 5.69e-07 mol/l
Molar Refractivity	81.75	Class	Poorly soluble
TPSA	32.59 Å²	*Medicinal Chemistry*	
Lipophilicity		PAINS	0 alert
Log Po/w (XLOGP3)	4.83	Brenk	1 alert: imine_1
Log Po/w (WLOGP)	4.75	Leadlikeness	No; 1 violation: XLOGP3>3.5
Log Po/w (MLOGP)	3.89	Synthetic accessibility	2.43
Log Po/w (SILICOS-IT)	5.23	*Druglikeness*	
Consensus Log Po/w	4.67	Lipinski	Yes; 0 violation
Pharmacokinetics		Ghose	Yes
GI absorption	High	Veber	Yes
BBB permeant	Yes	Egan	Yes
P-gp substrate	No	Muegge	Yes
Log Kp (skin permeation)	-4.54 cm/s	Bioavailability Score	0.55

The SwissADME analysis gives a molecular weight of 273.76 g mol^{-1} and a consensus LogP value of 4.67, indicating a relatively lipophilic molecular character. The compound has a TPSA of 32.59 Å², together with two hydrogen-bond acceptors and one hydrogen-bond donor, and fulfills the Lipinski criteria without violations (Table 2). Its predicted GI absorption is high, while BBB permeation is also predicted, and the molecule is not identified as a P-gp substrate. The bioavailability score of 0.55 suggests a moderate predicted oral bioavailability. The calculated solubility values, particularly the SILICOS-IT prediction (Log S = −6.25), indicate that limited aqueous solubility may be an important consideration despite the favorable predicted absorption profile.

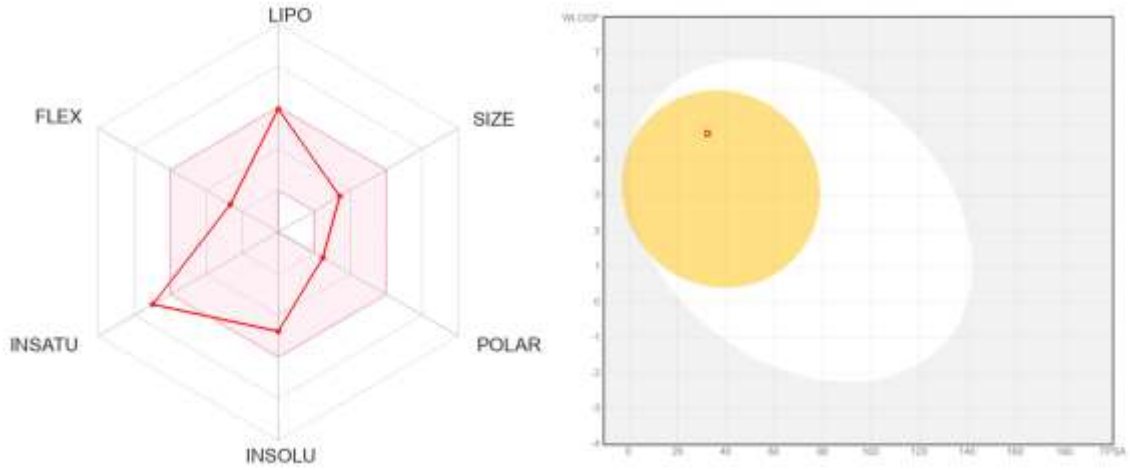


Fig. 4 The radar and boiled egg model of $C_{16}H_{16}ClNO$.

The BOILED-Egg model in Fig. 4 places the ligand within the yellow region, supporting the prediction of favorable passive gastrointestinal absorption. The position of the molecule in this region is consistent with its high GI absorption predicted by SwissADME and its relatively low polar surface area. At the same time, the absence of a position within the white central region should be interpreted together with the independent SwissADME prediction of BBB permeation rather than as a direct experimental indication of brain penetration. $C_{16}H_{16}ClNO$ exhibits a favorable drug-likeness profile according to the major rule-based filters, although its high lipophilicity and relatively low predicted aqueous solubility warrant consideration in further biological evaluation.

IV. CONCLUSION

The present computational study demonstrates that $C_{16}H_{16}ClNO$ possesses favorable structural and electronic features together with notable AChE binding potential. Its optimized geometry and MEP distribution provide insight into the molecular regions involved in intermolecular recognition, while docking calculations yielded a strong binding affinity of -11.14 kcal mol^{-1} and an estimated K_i of approximately 6.8 nM against AChE. SwissADME predictions indicated high GI absorption, BBB permeation, and compliance with Lipinski's rule, although the relatively high lipophilicity and limited predicted aqueous solubility should be considered. The obtained results support the compound as a promising candidate for further experimental evaluation of its AChE inhibitory activity.

REFERENCES

[1] O. Şimşek, Ş. Atalay, M. Taşdoğan, E. Ağar, H. Bülbül, and S. Kansız, "Synthesis, Characterization, and Hirshfeld Surface Analysis of Two Schiff Bases: (*E*)-3-(((4-iodophenyl) imino) methyl)benzene-1,2-diol and (*E*)-3-(((3-bromo-4-methylphenyl)imino)methyl)-benzene-1,2-diol," *J. Struct. Chem.*, vol. *67*(4), pp. 827-840, 2026.

[2] S. Kansiz, "Investigating the structural and surface properties of copper and nickel complexes fluorine-based salen type Schiff bases," *J. Struct. Chem*., vol. 64(12), pp. 2295-2310, 2023.

[3] I. Selatnia, A. Sid, H. Lgaz, A. Bouhraoua, T. Roisnel, O.M.A. Khamaysa, S. Kansız, K. Metwally, A.G. Soliman, H.-S. Lee, and S. Boudiba, "Synthesis, structural characterization, and biological evaluation of novel iodophenyl-anthracene Schiff bases: Insights from X-ray crystallography, DFT, antibacterial, and molecular docking studies," *J. Mol. Struct.*, vol. 1350, pp. 144066, 2025.

[4] M. K. Gumus, F. Sen, S. Kansiz, N. Dege, and E. Saif, "Crystal structure and Hirshfeld surface analysis of 2-{[(E)-(3-cyclobutyl-1H-1, 2, 4-triazol-5-yl) imino] methyl} phenol," *Acta Crystallogr. E: Crystallogr. Commun*., vol. 77(12), pp. 1267-1271, 2021.

[5] S. Yeşilbağ, S. Kansız, N. Dege, and E. Ağar, "Structural investigation and hirshfeld surface analysis of two [ONO]-Type schiff bases," *J. Struct. Chem*., vol. 65(3), pp.464-477, 2024.

[6] P. Şen, and S. Kansız, "Crystal Structure, Spectroscopic and Computational Studies of N2O2 Type Salen-Based Schiff Base and Its Di-boron Complex," *J. Chem. Crystallogr.*, vol. *52*(1), pp. 1-16, 2022.

[7] N. Arumugam, A. I. Almansour, R. S. Kumar, S. Yeswanthkumar, R. Padmanaban, Y. Arun, S. Kansız, N. Dege, T. S. Manohar, and S. Venketesh, "Design, stereoselective synthesis, computational studies and cholinesterase inhibitory activity of novel spiropyrrolidinoquinoxaline tethered indole hybrid heterocycle," *J. Mol. Struct*., 2021, vol. 1225, pp. 129165.

[8] S. Kansız, A. Tolan, H. İçbudak, and N. Dege, "Synthesis, crystallographic structure, theoretical calculations, spectral and thermal properties of trans-diaquabis (trans-4-aminoantipyrine) cobalt (II) acesulfamate," *J. Mol. Struct*., vol. 1190, pp. 102-115, 2019.
[9] H. Kargar, M. Fallah-Mehrjardi, K. S. Munawar, E. Acar, N. Dege, and S. Kansız, "Structural dynamics of a phenoxy-bridged heteronuclear Cu (II)/Hg (II)-Salen complex: synthesis, characterization, DFT insights, and Hirshfeld surface analysis," *J. Iran. Chem. Soc.*, vol. 22(4), pp. 769-780, 2025.
[10] H. Ozsanlı, O. Sımsek, O. R. Yılmaz, C. Cıcek, E. Agar, U. Çoruh, and E. M. Vazquez-Lopez, "Synthesis, crystal structure, Hirshfeld surface analysis, supramolecular and computational investigation of the (E)-4-methoxy-2-(((2-methoxy-5-(trifluoromethyl) phenyl) imino) methyl) phenol compound," J. Struct. Chem., vol. 65(5), pp. 946-957, 2024.
[11] O. Simsek, M. Ashfaq, M. N. Tahir, S. Ozturk, and E. Agar, "Synthesis and charaterizations of the Schiff base derived from 2-hydroxy-5-nitrobenzaldehyde alongwith Hirshfeld surface analysis and computational study," *J. Struct. Chem.*, vol. 64(5), pp. 942-953, 2023.
[12] N. Dege, A. S. Aydın, E. Ağar, S. Kansız, S. JoseKavitha, K. BalaSubramani, M. Hemamalini, and V. Rajakannan, "Synthesis, crystal structure, Hirshfeld surface analysis, In-Silico assessment of druggability and molecular docking studies of Schiff base compound," *Chem. Data Collect.*, vol. *25*, 100320, 2020.
[13] M. J. Frisch, G. Trucks, H. Schlegel, G. Scuseria, M. Robb, J. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, and G. Petersson, 2009. Gaussian 09, Revision D. 01, Gaussian, Inc.: Wallingford, CT.
[14] R. Dennington, T. Keith, and J. Millam, 2009. GaussView, version 5, Semichem Inc.: Shawnee Mission, KS.
[15] O. Trott, and A. J. Olson, "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading," *J. Comput. Chem*., vol. 31(2), pp. 455-461, 2010.
[16] https://www.3dsbiovia.com/.
[17] https://www.rcsb.org/.
[18] A. Daina, O. Michielin, and V. Zoete, "SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules," *Sci. Rep*., vol. 7(1), pp. 42717, 2017.
[19] K. Chkirate, S. Kansiz, K. Karrouchi, J. T. Mague, N. Dege, and E. M. Essassi, "Crystal structure and Hirshfeld surface analysis of N-{2-[(E)-(4-methylbenzylidene) amino] phenyl}-2-(5-methyl-1-H-pyrazol-3-yl) acetamide hemihydrate," *Acta Crystallogr. E: Crystallogr. Commun*., vol. 75(2), pp. 154-158, 2019.
[20] S. Kansız, G. Kaştaş, M. K. Gümüş, N. Dege, and R. Aydın, "Synthesis, structural characterization, and supramolecular assembly of a one-dimensional cobalt (II) coordination polymer with succinato and isonicotinamide ligands," *J. Struct. Chem*., vol. 66(4), pp. 722-732, 2025.
[21] R. Ilmi, S. Kansız, N. K. Al Rasbi, J. Husband, N. Dege, and M. S. Khan, "Synthesis, X-ray crystal structure and determination of non-covalent interactions through Hirshfeld surface analysis of a pure red emitting asymmetrical octacoordinated Sm (III) complex," *Polyhedron*, vol. 246, pp. 116673, 2023.
[22] S. Kansiz, S. Malinkin, and N. Dege, "Synthesis, crystal structure and Hirshfeld surface analysis of tetraaquabis (isonicotinamide-κN1) cobalt (II) succinate," *Acta Crystallogr. E: Crystallogr. Commun*., vol. 74(7), pp. 1026-1029, 2018.
[23] M. K. Gümüş, S. Kansız, G. B. Tulemisova, and N. Dege, "Synthesis, Crystal Structure, and Hirshfeld Surface Analysis of 4,4A,5,6,7,8-Hexahydro-4A-Methyl-2,5,7-Triphenyl-2H-Pyrido[4,3-d][1,3] Oxazine," *J. Struct. Chem*., vol. 66(1), pp. 201-213, 2025.
[24] S. Cakmak, S. Kansiz, M. Azam, C. C. Ersanlı, O. Idil, A. Veyisoglu, H. Yakan, H. Kütük, and A. Chutia, "Synthesis, structural investigation, Hirshfeld surface analysis, and biological evaluation of N-(3-Cyanothiophen-2-yl)-2-(thiophen-2-yl) acetamide," *ACS omega*, vol. 7(13), pp. 11320-11329, 2022.
[25] S. Kansız, M. Azam, N. Dege, N. Ermiş, S.I. Al-Resayes, and M. Alam, "Supramolecular assembly in designing co-crystals of fumaric acid and pyrimidine/picolinate derivatives," *Green Chem. Lett. Rev*., vol. 15 (3), pp. 825-836, 2022.
[26] Y. El Bakri, M. Kurbanova, A. Ahsin, S. Karthikeyan, A. Maharramov, R. Al-Salahi, S. Kansız, S. N. Mali, and M. Ashfaq, "Synthesis, crystal structure investigation, and theoretical approaches to discover potential 6-bromo-3-cyanocoumarin as a potent inhibitor MetAP (methionine aminopeptidase)2," *Chem. Phys. Impact*, vol. 8, 100477, 2024.
[27] M. K. Alenzi, M. Azam, S. I. Al-Resayes, S. Kansız, A. S. Çamaş, D. C. Agurokpon, H. Louis, V. Kumar, N. Dege, and M. Alam, "Experimental and theoretical investigation into the design of nickel (II), copper (II), and zinc (II) complexes substituted with pyrazole ligand," *Polyhedron*, vol. 259, pp. 117073, 2024.
[28] S. Kansız, Y. Sert, S. Meral, A. Alaman Ağar, and N. Dege, "Synthesis, Experimental and Theoretical Characterization of (E)-1-((5-nitrothiophen-2-yl) methylene)-2-Phenylhydrazine," *J. Struct. Chem.*, vol. *67*(2), pp. 291-308, 2026.
[29] O. Şimşek, S. Kansız, İ. Yılmaz, A. Çukurovalı, and M. Dinçer, "Structural characterization, DFT-based electronic analysis, Hirshfeld surface study, molecular docking, and in silico drug-likeness evaluation of N-(4-(3-methyl-3mesitylcyclobutyl)-1H-imidazol-2-yl) acetamide," *J. Mol. Struct.*, vol. 1358, pp. 145399, 2026.
[30] A.-A. B. OmarAli, A. J. M. Al-Karawi, A. A. Awad, N. Dege, S. Kansız, E. Agar, Z. A. Hussein and I. R. Mohammed, "Two new zinc(II) and mercury(II) complexes based on N,V-(cyclohexane-1,2-diylidene)bis(4-fluorobenzohydrazide): synthesis, crystal structures and antibacterial activities," *Acta Crystallogr. C: Struct. Chem*., vol. 76, pp. 476-482, 2020.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Biopolymer Additives for Biological Soil Stabilization in Geotechnical Engineering: A Critical Assessment of MICP Mechanisms, Performance, and Applications

Bayram Ateş [1*]

[1]*Civil Engineering Department, Karadeniz Technical University, Trabzon, Türkiye*
[*]*(bayramates61@hotmail.com)*

Abstract **–** Biological soil stabilization has become an increasingly important research field because it offers potentially lower-carbon and more environmentally compatible alternatives to cement- and lime-based ground improvement. Within this field, bio-mediated approaches such as microbially induced carbonate precipitation (MICP) and enzyme-induced carbonate precipitation (EICP), together with microbial processes such as bioclogging and biofilm-mediated stabilization, are frequently discussed alongside biopolymer-based treatment. Particular attention is given to xanthan gum, guar gum, sodium alginate, chitosan, gellan gum, agar, and lignin, which are among the most widely investigated biopolymers in geotechnical engineering. The principal bio-mediated soil-stabilization routes and the urease-catalyzed urea-hydrolysis pathway underlying MICP and EICP are first outlined. The role of biopolymers as direct binders and complementary additives is then examined in relation to particle bonding, pore filling, hydrogel formation, calcium-mediated crosslinking, and bio-mediated mineralization. Representative studies on sands, residual soils, expansive soils, dredged sediments, and slope materials are comparatively assessed in terms of strength gain, hydraulic behavior, durability, and microstructural change. The available evidence generally indicates that biopolymers can improve unconfined compressive strength, shear strength, cohesion, erosion resistance, and, in many cases, reduce hydraulic conductivity. However, performance is strongly governed by soil type, polymer type, dosage, curing regime, and environmental exposure. Biopolymer-assisted biological stabilization is technically promising, although field-scale validation, long-term durability assessment, and standardized design procedures remain necessary for broader geotechnical implementation.

Keywords – Biological soil stabilization; biopolymer additives; MICP; xanthan gum; sodium alginate; guar gum; sustainable geotechnics.

I. INTRODUCTION

Conventional soil stabilization by cement, lime, and other chemical binders remains effective for many geotechnical applications, yet its environmental burden has stimulated the search for more sustainable alternatives. In response, biological and bio-based improvement techniques have gained substantial attention. These techniques seek to improve engineering performance through microbial activity, bio-mediated mineralization, or natural polymeric binders that interact with the soil skeleton in an environmentally compatible manner [1], [2].

Biological and bio-mediated soil improvement encompasses MICP, enzyme-induced carbonate precipitation (EICP), bioclogging, and biofilm-mediated strengthening. Among these approaches, MICP has emerged as one of the most intensively studied methods because urease-producing bacteria can hydrolyze urea, raise pore-fluid pH, generate carbonate ions, and precipitate calcium carbonate at particle contacts and within pores [1], [3]. The resulting calcite bridges improve interparticle bonding and reduce pore connectivity. However, the treatment efficiency of MICP can be affected by non-uniform precipitation, limited bacterial retention, reagent transport, and durability concerns under aggressive

environmental conditions. In parallel, biopolymer-based stabilization has developed into a major branch of sustainable geotechnics. Biopolymers are naturally derived or biodegradable macromolecules obtained from plants, microorganisms, algae, or animals. In soils, they may function as viscous binders, hydrogels, film-forming agents, crosslinkable matrices, or physicochemical bonding media. The most frequently studied materials include xanthan gum (XG), guar gum (GG), sodium alginate (SA), chitosan, gellan gum, agar, and lignin [2], [5]–[18]. Their main geotechnical benefits include higher cohesion, improved unconfined compressive strength (UCS), reduced permeability, mitigation of erosion, and enhanced integrity under wetting and drying cycles.

Biopolymers are particularly relevant to biological soil stabilization because they can be used in two complementary ways. First, they can act as direct bio-based stabilizers by coating particles, filling pores, and forming interparticle bridges after hydration and curing. Second, they can be incorporated into microbial or enzyme-based systems, where they influence fluid viscosity, bacterial immobilization, nucleation environments, and precipitation distribution. For example, sodium alginate exhibits strong affinity for calcium ions and can form calcium-crosslinked polymer networks [14], [20]. Its relevance to MICP-related systems is further supported by studies involving alginate-mediated bacterial immobilization and regulation of $CaCO_3$ precipitation [4]. Likewise, xanthan gum has been investigated as an additive that enhances the reinforcement efficiency of biocementation in sandy soils [19]. In this study, biological soil stabilization is considered within a broad bio-based and bio-mediated framework; however, direct biopolymer treatment does not necessarily involve active microorganisms and should therefore be distinguished from microbial stabilization processes such as MICP. Against this background, the present paper reviews biopolymer additives in biological soil stabilization, with particular emphasis on MICP-related mechanisms, polymer classification, microstructural interactions, geotechnical performance, and current application trends. The paper aims to connect the microbial and biopolymer perspectives rather than discussing them as isolated topics.

II. Materials and Method

Core experimental studies were selected when they directly addressed biopolymer-based or biopolymer-assisted soil stabilization and reported geotechnically relevant outcomes such as strength, stiffness, permeability, erosion resistance, durability, or microstructural evolution. Foundational studies on MICP/EICP chemistry and polymer classification were additionally included to establish the mechanistic and material-classification framework. The main soil categories represented in the reviewed studies were aeolian sand, sand–bentonite mixtures, expansive soils, granite residual soil, dredged sediment, silty sand, and slope-forming soils. In addition, studies on biopolymer-assisted MICP or calcium-mediated crosslinking were assessed to establish the link between direct biopolymer treatment and bio-mediated stabilization. The findings were synthesized comparatively rather than statistically, because the available studies differ substantially in soil type, specimen preparation, curing method, dosage range, crosslinking condition, and test procedure. The review therefore emphasizes mechanisms, trends, and contrasts instead of formal meta-analysis.

A. Biological and Bio-Mediated Methods for Soil Stabilization and the MICP Formation Mechanism

Biological and bio-mediated soil improvement can be broadly discussed through MICP, EICP, bioclogging, and biofilm-mediated stabilization. Among them, MICP is the best established for geotechnical strengthening because the hydrolysis of urea by urease-producing microorganisms leads to carbonate generation and subsequent calcium carbonate precipitation [1], [3]. Ureolysis-induced carbonate precipitation can be achieved through two closely related bio-mediated pathways, namely microbially induced carbonate precipitation (MICP) and enzyme-induced carbonate precipitation (EICP). As schematically illustrated in Fig. 1, the principal distinction between these approaches lies in the source and mode of action of urease. In MICP, urease-producing bacteria are introduced into the soil and retained within the pore network, where microbial urease catalyzes urea hydrolysis and promotes carbonate generation. In EICP, by contrast, free urease is supplied directly without the need for viable bacterial cells. In both pathways, the carbonate species generated by ureolysis react with available Ca^{2+} ions, leading to $CaCO_3$ precipitation within pore spaces and, more importantly, at particle-contact regions.

The resulting mineral bridges increase interparticle bonding and can decrease pore connectivity, thereby contributing to higher soil strength and lower hydraulic conductivity. However, the efficiency and spatial uniformity of precipitation depend on the transport and distribution of bacteria or free enzyme, as well as on the availability of urea and calcium ions.

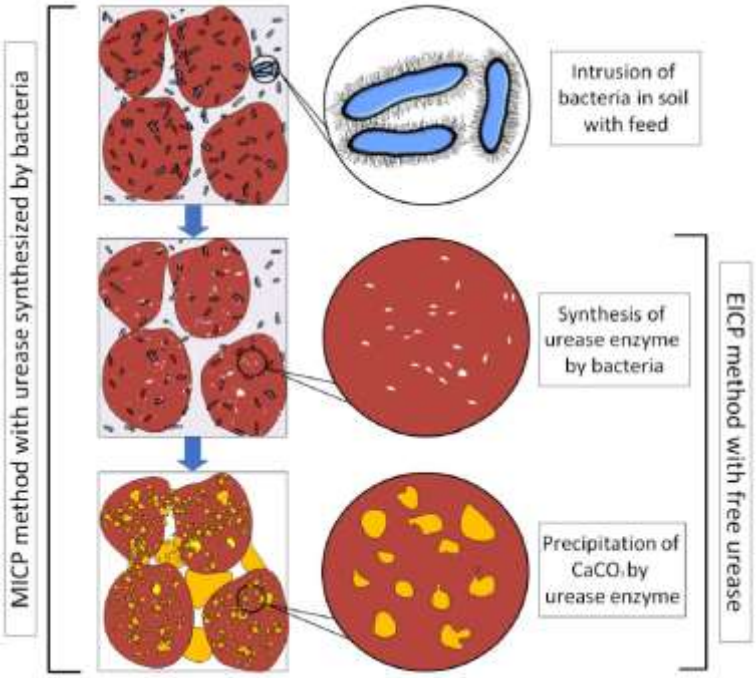


Fig. 1 Mechanisms of microbial-induced calcite precipitation (MICP) and enzyme-induced calcite Precipitation (EICP) in the CaCO3 precipitation [21]

Because ureolysis constitutes the central biochemical step governing carbonate generation, the reaction pathway should be considered in greater detail. Ureolysis constitutes the central biochemical step governing carbonate generation in both MICP and EICP. In these processes, urease strongly accelerates urea hydrolysis, enabling carbonate production at rates suitable for engineering applications. Urea may also hydrolyze very slowly in the absence of urease; however, this uncatalyzed reaction is not the mechanism exploited in practical MICP or EICP treatment. The principal urease-mediated reaction sequence can be expressed as follows [1], [3]:

$$CO(NH_2)_2 + H_2O \rightarrow NH_2COOH + NH_3 \quad (1)$$

$$NH_2COOH + H_2O \rightarrow NH_3 + H_2CO_3 \quad (2)$$

When urease-producing bacteria are present, these reactions proceed rapidly and are followed by ammonium generation, alkalinity increase, and carbonate formation:

$$2NH_3 + 2H_2O \leftrightarrow 2NH_4^+ + 2OH^- \quad (3)$$

$$H_2CO_3 \leftrightarrow HCO_3^- + H^+ \quad (4)$$

$$HCO_3^- + OH^- \leftrightarrow CO_3^{2-} + H_2O \quad (5)$$

$$Ca^{2+} + CO_3^{2-} \rightarrow CaCO_3\downarrow \quad (6)$$

The overall urease-catalyzed hydrolysis route is frequently expressed in the simplified form:

$$CO(NH_2)_2 + 2H_2O \rightarrow 2NH_4^+ + CO_3^{2-} \quad (7)$$

From a geotechnical standpoint, the most important consequence of Eqs. (3)–(7) is the precipitation of $CaCO_3$ at particle contacts and within the pore network. The precipitated calcite progressively transforms initially uncemented or weakly bonded soil grains into a more integrated soil–mineral skeleton by forming cementation bridges at interparticle contacts and partially filling pore spaces. This modification of the soil fabric increases interparticle bonding and contact stiffness, which can lead to higher shear strength, apparent cohesion, unconfined compressive strength, and small-strain stiffness. At the same time, partial pore filling and narrowing of preferential flow channels can reduce hydraulic conductivity and improve resistance to internal erosion and particle migration. In granular soils, sufficiently uniform carbonate precipitation can also restrict particle rearrangement and reduce contractive deformation under

cyclic loading, thereby increasing resistance to liquefaction and cyclic strain accumulation. However, the resulting improvement is strongly dependent on both the amount and spatial distribution of $CaCO_3$; localized or non-uniform precipitation may produce heterogeneous cementation, preferential weak zones, and spatially variable stiffness. Therefore, from an engineering design perspective, MICP performance should be evaluated not only in terms of total carbonate content, but also in terms of precipitation uniformity, bonding at critical particle contacts, permeability evolution, and the resulting changes in strength, stiffness, and cyclic response [1], [3].

B. Biopolymer classification and relevance to biological soil stabilization

Biopolymers used in geotechnical engineering can broadly be grouped according to their hydrogel-forming and environmentally responsive characteristics. Hydrocolloid-type materials such as XG, GG, and SA develop viscous aqueous solutions or hydrogels that can penetrate or occupy soil pores and establish bonding between adjacent particles. Thermally or chemically responsive materials such as agar and gellan gum can undergo sol–gel transitions under specific temperature or ionic conditions, providing additional control over their placement and subsequent solidification [2]. From a broader materials perspective, biodegradable and environmentally degradable polymers can be classified according to their origin and production route, as illustrated in Fig. 2. Naturally or biologically derived polymers include plant- and algae-derived polysaccharides and proteins, animal-derived polymers, and microorganism-derived biopolymers, whereas synthetic biodegradable polymers include materials synthesized from bio-derived monomers as well as biodegradable polymers produced from petrochemical or mixed feedstocks. This distinction is particularly relevant to geotechnical engineering because polymer origin and molecular structure strongly influence hydration, gelation, ionic interaction, biodegradability, and ultimately the soil–polymer bonding mechanism. Plant- and algae-derived materials such as guar gum and alginate, microorganism-derived polymers such as xanthan gum and gellan gum, and chitin-derived chitosan therefore exhibit different interaction pathways within the soil matrix. Accordingly, Fig. 2 provides a conceptual framework for relating biopolymer origin to stabilization mechanism and subsequent geotechnical performance rather than treating all biopolymers as a single homogeneous class [22], [23].

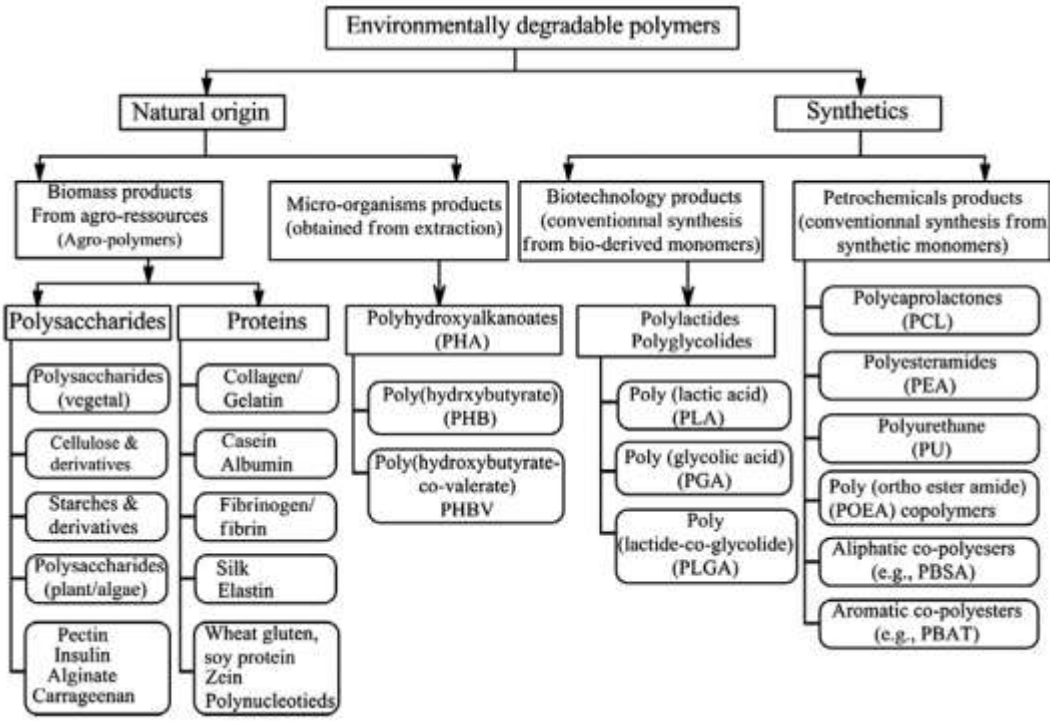


Fig. 2 Classifications of the different biodegradable polymers [23]

Xanthan gum is a microbial polysaccharide widely investigated because of its high viscosity and strong affinity for soil surfaces. During drying and curing, hydrated XG can form coatings around individual grains and polymer bridges between neighbouring particles. In fine-grained soils, hydroxyl and carboxyl functional groups additionally facilitate hydrogen bonding and electrostatic interaction with

mineral surfaces. These processes increase apparent cohesion and frequently result in substantial increases in UCS and shear strength [2], [6], [9], [10].

Guar gum is a plant-derived galactomannan capable of rapid hydration and strong gel formation. Its stabilization effect is generally associated with pore filling, particle encapsulation, and interparticle bonding. However, its long-term behavior may be sensitive to biological degradation. This distinction is important because short-term strength alone may overestimate the suitability of a biopolymer for permanent geotechnical applications [2], [5]–[8].

Sodium alginate is an anionic polysaccharide with pronounced affinity for divalent cations, especially Ca^{2+}. Calcium-induced crosslinking can transform sodium alginate into a three-dimensional calcium-alginate network. Within soil pores, this network can restrict fluid flow, improve particle bonding, and contribute to resilient response [14], [20]. Its calcium-binding characteristics are also particularly relevant to biological stabilization involving carbonate precipitation. Alginate can interact with calcium-rich environments while influencing nucleation, mineral morphology, and microorganism immobilization [4].

Chitosan, which is derived from chitin, differs from many other commonly investigated biopolymers because its amino groups become protonated under acidic to mildly acidic conditions, giving the polymer a pronounced cationic character. Electrostatic attraction between protonated amino groups and negatively charged mineral surfaces can promote flocculation and interparticle bonding. Experimental evidence has demonstrated its potential for improving sandy and clayey soils [15].

Agar and gellan gum exhibit thermally or ionically controlled gelation. Their ability to enter porous materials in a relatively fluid state and subsequently form a gel provides advantages for injection-based stabilization. Agar treatment has also been investigated for liquefaction mitigation [16], while gellan gum has been employed to improve the behavior of bentonite–sand systems and low-permeability barriers [17].

Lignin represents a plant-derived aromatic biopolymer with potential for both soil strengthening and ecological stabilization. Its relevance is particularly apparent in slope bioengineering, where the stabilizer should improve soil properties without suppressing vegetation development [12].

At the microscale, the performance of these materials can therefore be summarized by five interacting mechanisms: hydrogel formation, particle coating, interparticle bridging, pore filling, and physicochemical bonding. Their relative importance varies with polymer chemistry and soil mineralogy. In biological mineralization systems, a sixth mechanism emerges through interaction between the biopolymer network and mineral precipitation [2], [4], [8], [10].

C. Geotechnical Performance and Dosage-Dependent Behavior

Mechanical and hydraulic performance of biopolymer-treated soils: The reviewed literature consistently demonstrates that biopolymers can improve both the mechanical and hydraulic behavior of soils; however, the degree of improvement depends strongly on polymer type, dosage, soil characteristics, and curing conditions. Ai et al. [5] reported that guar gum significantly improved the mechanical behavior of aeolian sand and identified 1.5% GG with 21 days of curing as a favorable treatment condition. Under this condition, cohesion increased markedly, the internal friction angle increased by 63.72%, and hydraulic conductivity decreased by approximately 97.5%. In sand–bentonite mixtures, Al-Jabobi and Bilsel [6] compared XG and GG at concentrations from 0.5% to 3%. Approximately 2% was found to be an optimum concentration for both polymers, and the resulting UCS values were nearly three times those of untreated specimens. Their results also suggested that XG tends to provide slightly stronger and more stable long-term mechanical enhancement than GG under the tested conditions. For expansive soils, Raj et al. [9] showed that XG treatment increased UCS by roughly 1.8–3 times depending on curing duration and dosage. The observed strength gains were supported by microstructural densification and enhanced particle bonding. Similarly, Kalita et al. [11] demonstrated that the combined use of XG and eggshell powder can produce substantial CBR improvements in silty sand, confirming the potential of hybrid low-carbon stabilization systems for pavement subgrades. Liu and Wang [10] reported particularly informative results for granite residual soil because they presented the evolution of both cohesion and internal friction angle with XG and GG content. Their data show that XG exhibits an optimum around

1.5%, after which excessive polymer content causes a decline in strength parameters, whereas GG produces a more gradual and stable enhancement trend up to 2.0%. Figure 3 clearly demonstrates that the shear-strength response of granite residual soil is strongly dependent on both biopolymer type and dosage. For XG-treated specimens, cohesion and internal friction angle increase progressively up to approximately 1.5% XG, reaching 155.29 kPa and 35.54°, respectively. However, a further increase to 2.0% results in a pronounced reduction in both parameters. In contrast, GG produces a more stable enhancement pattern: cohesion remains high between 1.5% and 2.0%, while the internal friction angle continues to increase and reaches 39.01° at 2.0% GG [10]. These contrasting trends confirm that increasing biopolymer content does not necessarily produce a proportional improvement in soil strength and that the optimum dosage is polymer-specific. The different responses observed in Fig. 3 can be related to the distinct soil–polymer interaction mechanisms of XG and GG. At excessive XG contents, localized and overly thick hydrogel accumulations may develop at particle contacts, reducing interparticle friction and producing a lubricating effect within the soil matrix. Drying-induced heterogeneity may further contribute to local microfissuring and loss of mechanical efficiency. GG, by comparison, appears to form a more flexible and conformable hydrogel network that maintains particle bonding without producing the same degree of adverse interfacial lubrication at higher dosages [10]. From an engineering perspective, these results highlight the need to optimize biopolymer dosage on the basis of polymer chemistry, soil mineralogy, and the targeted balance between cohesion and frictional resistance rather than simply maximizing polymer content.

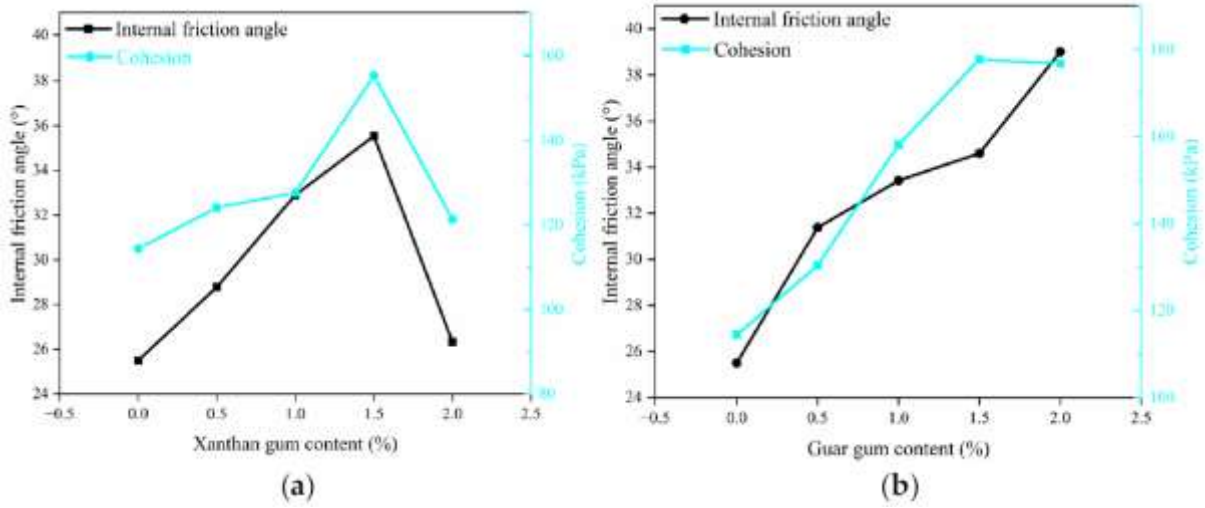


Fig. 3 Cohesion and internal friction angle of biopolymer-treated GRS: (a) XG; (b) GG [10]

D. *Microstructural Evidence, Durability, and Application Trends*

Microstructural observations provide direct evidence for the mechanisms through which biopolymers modify the soil skeleton and ultimately influence macroscopic geotechnical behavior. In dredged sediment, Ghafoori et al. [8] comparatively investigated calcium alginate (AL), xanthan gum (XA in the original study), chitosan (CH), and guar gum (GG) using SEM/BSE imaging coupled with EDS analysis. As shown in Fig. 4, the untreated sediment exhibits a relatively loose and weakly bonded fabric, with irregularly shaped particles separated by visible pore spaces and only limited interparticle contacts. The corresponding EDS spectra are dominated by O, Si, Al, and Fe, consistent with the mineralogical composition of the untreated sediment. In contrast, the treated specimens exhibit clear polymer-induced modifications of the particle contacts and pore structure, demonstrating that biopolymer stabilization does not operate through a single universal bonding mechanism [8].

The morphology of these polymer–soil associations differs substantially among the investigated biopolymers. Calcium alginate partially coats the sediment particles, fills local pore spaces, and forms bridge-like connections between adjacent aggregates, with some polymer bridges having dimensions below approximately 1 μm. Xanthan gum develops elongated fibrous structures, locally reaching widths of approximately 8.6 μm, which wrap around and connect neighboring particles to form a more continuous load-transfer network. Chitosan produces bridge-shaped and cluster-like formations, generally

below about 2 μm in diameter, which promote a denser fabric and reduce open pore space. Guar gum, in contrast, develops comparatively broad and continuous gel-like films, with polymeric bridges extending over distances of up to approximately 60 μm across particle surfaces. The EDS spectra obtained from the treated regions show increased carbon and oxygen signals, providing complementary evidence that the observed coatings, bridges, and clusters correspond to organic biopolymer phases rather than solely to the original mineral skeleton [8].

These microstructural differences are directly reflected in the measured mechanical response. The formation of hydrogel networks and polymer bridges primarily enhances apparent cohesion by increasing adhesion and interparticle connectivity; however, the effect on frictional resistance depends on the morphology and continuity of the polymer phase. In the investigated dredged sediment, xanthan gum produced the largest increase in cohesion, but the associated polymer films were accompanied by a reduction in friction angle, indicating that strong bonding can coexist with a partial lubricating effect at particle interfaces. Chitosan generated rougher and more clustered particle associations, which enhanced both cohesion and interlocking, whereas the broad GG films promoted adhesion but provided comparatively weaker anchorage to the mineral skeleton. Alginate produced continuous bridge-like and pore-filling phases that improved bonding and compressive response, although their relatively smooth morphology provided less enhancement of frictional resistance. Therefore, Fig. 4 demonstrates that the performance of a biopolymer-treated soil is governed not simply by polymer dosage, but by the geometry, continuity, surface characteristics, and particle-scale anchorage of the polymer–soil bonding network [8].

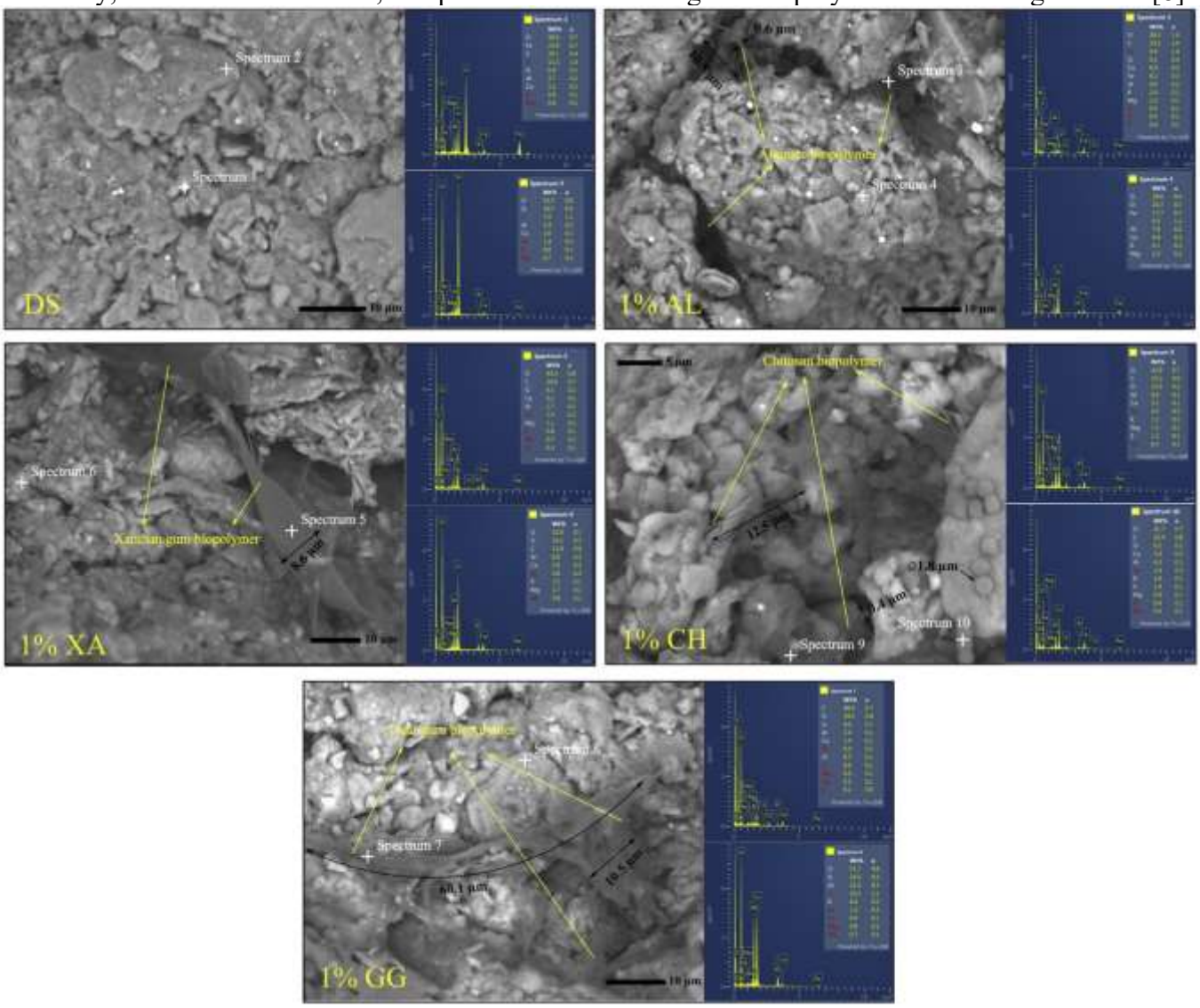


Fig. 4 SEM/BSE microphotographs and EDS spectra of sediment specimens extracted from the UCS samples [8].

The same study showed that 1 wt% XG, alginate, and chitosan increased UCS by approximately 40%, 29%, and 10%, respectively, and that the effect of biopolymers was generally stronger on cohesion than on the friction angle [8]. This observation is consistent with the broader literature, which often reports that biopolymers are particularly effective in creating apparent cohesion through gel networks and particle coating. Durability remains a critical issue. Sesay et al. [13] showed that ionic crosslinking of XG in rammed earth can significantly improve mechanical performance and environmental durability. Likewise, Lee et al. [18] demonstrated that severe weathering cycles may degrade the strength of XG-treated soils, emphasizing that untreated hydrogel networks can be sensitive to repeated wetting and drying. These findings indicate that long-term performance is governed by the stability of the polymer network, not only by the initial strength gain. The ecological application potential of biopolymers is especially clear in slope bioengineering. Gratchev et al. [12] found that lignin treatment increased the strength of several soils and did not show a clear adverse impact on vegetation growth. In sandy soil, lignin contributed to the formation of a firm surface crust, as illustrated in Fig. 5. This result is relevant for erosion control and low-impact surface stabilization.



Fig. 5 A view of the firm surface formed in Sand mixed with (a) 3% and (b) 1% lignin solution [12]

As summarized in Table 1, the effectiveness of biopolymer treatment varies markedly with soil type, polymer chemistry, dosage, and curing conditions. The reported optimum contents should therefore be interpreted as material- and test-specific values rather than universal design recommendations. For example, guar gum exhibited a mechanical optimum of approximately 1.5% after 21 days of curing in aeolian sand [5], whereas approximately 2% XG or GG produced the most favorable UCS response in the investigated sand–bentonite mixtures [6]. In granite residual soil, XG reached its optimum response at approximately 1.5%, while GG maintained a more stable improvement trend up to 2.0% [10]. These differences demonstrate that polymer dosage must be optimized in conjunction with soil mineralogy, particle-size distribution, moisture condition, and the targeted engineering response.

The comparison also indicates that strength enhancement is not governed by a single stabilization mechanism. XG-treated expansive soils exhibited substantial UCS gains accompanied by reduced void space and stronger particle bonding [9], whereas the dredged-sediment study showed that calcium alginate, XG, chitosan, and GG generated different microstructural morphologies and consequently different combinations of cohesion and frictional resistance [8]. In particular, the marked increases in cohesion reported for several biopolymer-treated soils suggest that polymer-induced bonding frequently contributes more strongly to strength improvement than changes in the friction angle. Consequently, comparisons based solely on UCS or peak shear strength may conceal important differences in the underlying soil–polymer interaction mechanisms.

Another important trend is the emergence of hybrid stabilization approaches. The combined use of XG and eggshell powder demonstrates that biopolymers can operate synergistically with mineral or waste-derived additives to improve subgrade performance [11], while biopolymer-assisted biocementation extends this concept directly into biological stabilization [19]. Such hybrid systems are particularly relevant because the polymer network may improve particle retention, fluid distribution, and bonding while biologically induced mineral precipitation provides an additional cementation mechanism. This

suggests that the future role of biopolymers may extend beyond their use as stand-alone binders toward their integration with MICP and other bio-mediated ground-improvement processes.

Table 1. Representative studies on biopolymer-based soil stabilization and their principal findings

Study	Soil type	Biopolymer	Dosage/curing	Main response	Principal finding
Ai et al. [5]	Aeolian sand	Guar gum	Mechanical optimum: 1.5%; 21 d	Direct shear, UCS, tensile strength, permeability	Mechanical optimum at 1.5% after 21 d; marked increase in cohesion and friction angle and approximately 97.5% reduction in hydraulic conductivity.
Al-Jabobi and Bilsel [6]	Sand–bentonite mixture	XG, GG	0.5–3.0%; 1–180 d	UCS, axial strain, FTIR	Approximately 2% produced the most favorable overall mechanical response; XG showed greater long-term stability, whereas GG exhibited biodegradation during extended curing.
Raj et al. [9]	Expansive soils	XG	0.5–2.0%; 3–28 d	UCS, deformation modulus, SEM / FTIR / EDX	UCS increased by about 1.8–3 times, with stronger particle bonding, reduced void space, and evidence of physicochemical interaction.
Liu and Wang [10]	Granite residual soil	XG, GG	0.5–2.0%; 14 d	Direct shear, disintegration, SEM, LF-NMR	XG showed an optimum near 1.5%; GG maintained a more stable improvement trend up to 2.0%. Excess XG caused localized gel-rich zones that weakened interparticle interlocking.
Ghafoori et al. [8]	Dredged sediment	Calcium alginate, XG, chitosan, GG	1 wt% for mechanical tests	UCS, direct shear, SEM/EDS	Polymer-specific bonding morphologies produced different mechanical responses; cohesion gains were generally more pronounced than changes in friction angle, and XG gave the largest UCS increase.
Kalita et al. [11]	Silty sand	XG + eggshell powder	1% XG + 6% ESP	CBR, compaction, SEM/EDX	The combined treatment produced the best overall CBR response among the investigated mixes, demonstrating the potential of hybrid biopolymer–waste-derived stabilization.
Sesay et al. [13]	Low-clay soil / rammed earth	Native and ionically crosslinked XG	0.5–1.5%; 28 d + accelerated environmental cycling	UCS, tensile strength, stiffness, durability, microstructure	Crosslinking markedly enhanced environmental durability; post-cycling UCS loss was about 1.13% for crosslinked XG compared with 19.36% for native XG.
Gratchev et al. [12]	Black clay, silt, and sand used for slope bioeng.	Lignin	1% and 3%; 40 d; mixing/spraying	Penetrometer/vane shear, water retention, vegetation response	Lignin generally increased strength and moisture retention without a clear adverse effect on vegetation; 3% lignin applied by mixing produced the greatest strength gains.
Hang et al. [19]	Sandy soil	XG + biocementation (MICP)	XG used as an additive to biocementation	Biocementation reinforcement	Demonstrated that XG can function as a complementary biopolymer additive that improves the reinforcement effect of biologically cemented sand.

Table 2 highlights that the macroscopic geotechnical response of biopolymer-treated soil originates from fundamentally different microscopic mechanisms. Hydrogel-forming polymers such as XG and GG primarily improve performance through pore filling, particle coating, and interparticle bridging, whereas sodium alginate additionally provides a calcium-responsive crosslinking mechanism. Chitosan relies more strongly on electrostatic interactions and flocculation because of its cationic character, while agar and gellan gum offer stimuli-responsive gelation that can be controlled through temperature or ionic

conditions. These distinctions explain why different biopolymers can produce comparable increases in soil strength while exhibiting markedly different hydraulic, deformation, and durability characteristics.

The practical limitations summarized in Table 2 are equally important as the initial strength gains. XG may exhibit reduced performance when excessive polymer creates localized gel-rich zones, and severe wetting–drying cycles can weaken untreated polymer networks [10], [18]. GG can provide substantial mechanical and hydraulic improvement, but long-term biodegradation may become significant under favorable microbial and moisture conditions [6]. These observations indicate that biodegradability represents both an environmental advantage and an engineering challenge. Crosslinking provides one possible solution, as demonstrated by the markedly improved environmental resistance of ionically crosslinked XG compared with native XG [13].

Accordingly, biopolymer selection should be application-oriented rather than based solely on the maximum laboratory strength. Sodium alginate is particularly attractive where calcium-mediated crosslinking or interaction with biomineralization is desired; agar and gellan gum are promising where controllable gelation and hydraulic or cyclic performance are important; and lignin is especially relevant to surface stabilization and slope bioengineering where soil strengthening must coexist with moisture retention and vegetation development. XG and GG remain versatile general-purpose binders, but their optimum dosage and durability must be verified for the specific soil and exposure condition. This function-oriented selection framework is more appropriate for engineering design than attempting to identify a single universally superior biopolymer.

Table 2. Dominant mechanisms, geotechnical benefits, and practical considerations of commonly used biopolymers.

Biopolymer	Dominant mechanism	Typical geotechnical benefit	Practical consideration	Representative references
Xanthan gum	Hydrogel formation, particle coating, interparticle/fibrous bridging, hydrogen bonding	Higher UCS and cohesion, lower permeability, improved erosion and water stability	Severe wetting–drying can reduce strength; excessive dosage may create localized gel-rich or lubricating zones.	[2], [6], [9], [10], [13], [18], [19]
Guar gum	Hydrogel formation, pore filling, particle encapsulation, flexible interparticle bridging	Strength gain, reduced hydraulic conductivity, improved ductility and subgrade performance	Long-term response can be affected by biodegradation and moisture conditions; performance is strongly soil- and curing-dependent.	[2], [5]–[8], [10]
Sodium alginate	Calcium-mediated ionic crosslinking, film formation, pore filling, microorganism immobilization	Resilient behavior, pore restriction, improved bonding, compatibility with calcium-rich MICP systems	Performance depends on Ca^{2+} availability, crosslinking conditions, solution chemistry, and curing.	[2], [4], [14], [20]
Chitosan	Electrostatic attraction, flocculation, polymer bridging, clustered particle bonding	Shear/UCS enhancement, improved interlocking and reduced hydraulic conductivity	Solution preparation chemistry and pH can strongly influence gelation, adsorption, and stabilization efficiency.	[2], [8], [15]
Gellan gum / agar	Stimuli-responsive gelation controlled by thermal and/or ionic conditions	Potential for hydraulic barriers, injectable treatment, and liquefaction mitigation	Temperature/ionic control and application-specific curing must be managed to achieve adequate injectability and gelation.	[2], [16], [17]
Lignin	Surface adhesion, particle bonding, moisture retention, formation of a stiff surface crust	Slope bioengineering, erosion control, near-surface strengthening, ecological compatibility	Particularly promising for surface and slope applications; long-term field performance requires further validation.	[12]

III. DISCUSSION

The reviewed studies collectively show that biopolymer-assisted biological stabilization cannot be represented by a single universal dosage rule. Instead, the optimum treatment level is controlled by polymer chemistry, soil mineralogy, grain-size distribution, water content, and curing environment. The

contrasting XG and GG trends in granite residual soil provide a clear example: whereas XG displayed an optimum around 1.5%, GG continued to enhance cohesion and friction angle over a wider dosage range [10]. This difference indicates that design strategies should be polymer-specific rather than generic.

A second major outcome is that the interface between biopolymers and microbial stabilization is not merely conceptual; it is mechanistically significant. Biopolymers can increase pore-fluid viscosity, improve bacterial retention, provide nucleation environments, and interact with calcium ions and mineralizing conditions. Sodium alginate is especially relevant because calcium-induced crosslinking can produce a robust calcium-alginate gel network [14], [20]. In MICP-related systems, alginate can additionally function as an immobilization matrix and influence $CaCO_3$ precipitation and mineral formation [4]. The reported use of XG as an additive for improving the reinforcement effect of biocementation further confirms that biopolymers can strengthen biological stabilization processes rather than simply acting as alternative binders [19].

Microstructural observations further clarify why the same dosage can lead to different geotechnical responses. Alginate may create bridge-like and pore-filling films, XG may develop fibrous networks, chitosan may generate rough clustered bonds, and GG may form broader gel films [8]. These distinct morphologies influence not only compressive strength but also whether the treated soil behaves in a cohesion-dominated or friction-dominated manner. Consequently, selecting a biopolymer for hydraulic containment, erosion control, subgrade stabilization, or slope treatment requires a function-oriented approach.

Durability remains the key limitation for practical implementation. Because many biopolymers are hydrophilic and biodegradable, their initial strength gains may not be preserved under severe wetting–drying cycles or prolonged environmental exposure [13], [18]. Crosslinking, hybrid polymer–mineral systems, and polymer-assisted MICP therefore represent some of the most promising research directions. Field-scale validation is another unmet need. Many laboratory studies report strong improvements under controlled mixing and curing conditions, but the translation of these results to in-situ treatment, large-scale mixing, or injection remains insufficiently documented. Overall, the literature indicates that biopolymer additives can function as an effective bridge between biological stabilization and sustainable geotechnical engineering. Their future success will depend on moving from proof-of-concept studies toward performance-based design, durability verification, and realistic field demonstrations.

Another important research need is the development of standardized reporting protocols. Many studies use different curing temperatures, water contents, polymer solution concentrations, mixing energies, and specimen dimensions, which complicates direct comparison. Establishing consistent reporting standards would greatly improve the transfer of laboratory findings into design-oriented practice. The application spectrum of biopolymer stabilization is also broadening. Current studies cover pavement subgrades, hydraulic barriers, dredged-sediment reuse, rammed-earth construction, erosion control, and slope bioengineering [2], [11], [12], [17]. This diversity suggests that biopolymers should not be evaluated solely as substitutes for cementitious binders; in many cases, their value lies in niche functions such as permeability control, surface protection, ecological compatibility, or support to bio-mediated mineralization.

IV. CONCLUSION

This study shows that biopolymer additives represent an important class of bio-based stabilizers and can also function as complementary additives in bio-mediated systems such as MICP.

Biological and bio-mediated soil-improvement approaches include MICP, EICP, bioclogging, and biofilm-mediated stabilization, with MICP remaining one of the most extensively investigated routes for geotechnical strengthening.

Among the most frequently investigated biopolymers in geotechnical engineering are XG, GG, sodium alginate, chitosan, gellan gum, agar, and lignin.

The reviewed studies consistently report increases in strength and reductions in hydraulic conductivity, but the response is strongly dependent on polymer type and dosage. Optimum contents exist, and excessive polymer addition may reduce interparticle interlocking or generate non-uniform gel agglomerations.

Microstructural observations and durability studies indicate that future progress should focus on crosslinking strategies, hybrid polymer–MICP systems, standardized design methods, and field-scale validation.

REFERENCES

[1] J. T. DeJong, B. M. Mortensen, B. C. Martinez, and D. C. Nelson, "Bio-mediated soil improvement," Ecological Engineering, vol. 36, no. 2, pp. 197–210, 2010.

[2] J.-E. Ryou, C. Jang, S. Gang, and J. Jung, "A review of biopolymers in geotechnical engineering: Mechanical properties, dynamic behavior, applications, and mechanisms," Journal of Rock Mechanics and Geotechnical Engineering, vol. 18, pp. 4971–4997, 2026.

[3] V. S. Whiffin, L. A. van Paassen, and M. P. Harkes, "Microbial carbonate precipitation as a soil improvement technique," Geomicrobiology Journal, vol. 24, no. 5, pp. 417–423, 2007.

[4] J. Wu and R. J. Zeng, "Biomimetic regulation of microbially induced calcium carbonate precipitation involving immobilization of Sporosarcina pasteurii by sodium alginate," Crystal Growth & Design, vol. 17, no. 4, pp. 1854–1862, 2017.

[5] Z. Ai et al., "Improvement of aeolian sand using guar gum: Mechanical behavior, plant growth response, and microstructural evaluation," Geotechnical and Geological Engineering, vol. 43, Art. no. 375, 2025.

[6] A. Al-Jabobi and H. Bilsel, "The effects of xanthan gum and guar gum on the mechanical properties of sand–bentonite mixtures," Sustainability, vol. 17, Art. no. 5339, 2025.

[7] T. Pitso, L. M. Lalicata, D. Gallipoli, A. De Paoli, and A. W. Bruno, "Effect of hydrated lime and guar gum treatment on the geotechnical behaviour of a Ligurian sandy silt," Environmental Geotechnics, vol. 12, no. 4, pp. 316–326, 2025.

[8] Y. Ghafoori, P. Ghadir, S. Dolenec, and S. Lenart, "Toward the sustainable stabilization of dredged sediment using biopolymers: A mechanical performance study," Results in Engineering, vol. 28, Art. no. 108233, 2025.

[9] N. Raj, S. Selvakumar, and K. Muthukkumaran, "Effect of xanthan gum biopolymer treatment on the strength enhancement of expansive soils," International Journal of Pavement Research and Technology, 2026.

[10] Y. Liu and Z. Wang, "Hydro-mechanical performance and microstructural evolution of biopolymer-modified granite residual soil," Applied Sciences, vol. 16, Art. no. 4916, 2026.

[11] A. Kalita, N. K. Singh, G. Goswami, S. Basack, and M. Karakouzian, "Microstructural analysis and subgrade improvement of silty sand using xanthan gum biopolymer and eggshell powder," CivilEng, vol. 7, Art. no. 11, 2026.

[12] I. Gratchev, Q. Tang, S. Akosah, and J. Sugawara, "Application of lignin for slope bioengineering: Effect on soil improvement and plant growth," Applied Sciences, vol. 15, Art. no. 4173, 2025.

[13] T. Sesay, Y. Chen, Y. Xie, and J. Xue, "Biopolymer stabilization of rammed earth based on native and ionic crosslinked xanthan gum: Mechanical, microscopic, and environmental durability assessment," Journal of Building Engineering, vol. 128, Art. no. 116505, 2026.

[14] M. G. Arab, R. A. Mousa, A. R. Gabr, A. M. Azam, S. M. El-Badawy, and A. F. Hassan, "Resilient behavior of sodium alginate-treated cohesive soils for pavement applications," Journal of Materials in Civil Engineering, vol. 31, no. 1, Art. no. 04018361, 2019.

[15] N. Hataf, P. Ghadir, and N. Ranjbar, "Investigation of soil stabilization using chitosan biopolymer," Journal of Cleaner Production, vol. 170, pp. 1493–1500, 2018.

[16] C. Jang, B. Yang, W. T. Hong, J. Ahn, and J. Jung, "Soil improvement using agar gum polymer for seismic liquefaction mitigation," Soil Dynamics and Earthquake Engineering, vol. 177, Art. no. 108405, 2024.

[17] A. T. P. Tran, I. Chang, and G. C. Cho, "Gellan gum for strengthening bentonite–sand slurry in vertical cutoff walls," Journal of Materials in Civil Engineering, vol. 34, no. 4, Art. no. 04022023, 2022.

[18] M. Lee, Y.-M. Kwon, D.-Y. Park, I. Chang, and G.-C. Cho, "Durability and strength degradation of xanthan gum-based biopolymer treated soil subjected to severe weathering cycles," Scientific Reports, vol. 12, Art. no. 19453, 2022.

[19] L. Hang, L. A. van Paassen, F. Ehsasi, Y. Chen, R. Li, and J. He, "An experimental investigation on use of xanthan gum biopolymer as additive for improving the reinforcement effect of biocementation on sandy soils," Geomicrobiology Journal, vol. 41, pp. 627–635, 2024.

[20] A. Bakhshizadeh, N. Khayat, and S. Horpibulsuk, "Surface stabilization of clay using sodium alginate," Case Studies in Construction Materials, vol. 16, Art. no. e01006, 2022.

[21] A. Almajed, M. A. Lateef, A. A. B. Moghal, and K. Lemboye, "State-of-the-Art Review of the Applicability and Challenges of Microbial-Induced Calcite Precipitation (MICP) and Enzyme-Induced Calcite Precipitation (EICP) Techniques for Geotechnical and Geoenvironmental Applications," Crystals, vol. 11, no. 4, Art. no. 370, 2021.

[22] L. Averous and N. Boquillon, "Biocomposites based on plasticized starch: Thermal and mechanical behaviours," Carbohydrate Polymers, vol. 56, pp. 111–122, 2004.

[23] T. Gurunathan, S. Mohanty, and S. K. Nayak, "A review of the recent developments in biocomposites based on natural fibres and their application perspectives," Composites Part A: Applied Science and Manufacturing, vol. 77, pp. 1–25, 2015.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Design of a Notched Stepped Antenna with DGS-Enhanced Bandwidth

Burak Dokmetas[1]

[1]*Department of Electrical and Electronics Engineering, Kafkas University, Kars, Turkey*

[1]*(burakd@kafkas.edu.tr)*

Abstract – A wideband antenna based on a notched stepped radiator and a modified ground configuration is investigated in this work. The antenna is modeled on a 0.8-mm-thick PREPERM ABS300 dielectric layer with a relative permittivity of 3, selected for its RF-oriented properties and additive-manufacturing compatibility. The design process begins with a notched rectangular radiator and is progressively improved by introducing stepped transitions along the lower edge of the radiator. Subsequently, a DGS formed by horizontal slots and a circular defect is applied to the ground plane to modify the surface current path and improve impedance matching. After the final ground modification, the -10 dB impedance criterion is satisfied continuously from approximately 3 to 11.2 GHz. A peak simulated gain of about 4.3 dBi is obtained within this frequency interval. The influence of the proposed geometry is further examined through surface current distributions and far-field radiation patterns at selected frequencies. The obtained response indicates that reshaping the radiator together with perturbing the ground plane enables the separated resonances to form a continuous operating band. Owing to its simple geometry, compact planar structure, and compatibility with additive manufacturing, the proposed antenna offers a practical candidate for wideband RF and wireless communication applications.

Keywords –Defected Ground Structure (DGS), wideband antenna, stepped antenna, notched radiator, bandwidth enhancement

I. INTRODUCTION

The increasing demand for high-data-rate and multi-standard wireless communication systems has encouraged the development of antennas capable of operating over wide frequency ranges. Wideband and super-wideband antenna configurations have been extensively investigated because of their ability to support multiple communication services within a single compact structure. Previous studies have broadened antenna operation by reshaping the radiating element, introducing stepped transitions, etching slots, and modifying the ground region [1–3].

In parallel with advances in antenna geometry, additive manufacturing has emerged as a promising fabrication approach for RF and microwave structures. Additive manufacturing provides an alternative route for realizing RF structures with controllable dielectric geometry and short prototype cycles. Its ability to fabricate non-standard shapes directly from a digital model has encouraged the use of 3D-printable dielectrics in microwave and antenna implementations [4–6]. PREPERM ABS300 belongs to a class of dielectric filaments developed specifically for high-frequency applications, where repeatable permittivity and relatively low dielectric loss are important design requirements. These characteristics make it a suitable candidate for the realization of additively manufactured antenna substrates and dielectric RF structures.

Defected Ground Structure (DGS) is another widely used technique for controlling antenna impedance characteristics and enhancing bandwidth. Etching discontinuities in the ground conductor perturbs the return-current distribution and changes the equivalent response seen at the antenna input. Properly positioned defects can therefore shift existing resonances or create additional resonant behavior, allowing neighboring operating regions to approach each other. Different DGS geometries have been successfully employed in planar antennas to obtain wideband, ultra-wideband, and multiband characteristics [7-10].

In this study, a notched stepped antenna combined with a DGS is proposed for wideband operation. The design begins with a notched rectangular radiator and is progressively modified by introducing stepped transitions along the lower section of the radiator. A DGS consisting of horizontal slots and a circular defect is subsequently introduced into the ground plane to improve impedance matching and merge the resonant regions. PREPERM ABS300 with a relative permittivity of 3 and a thickness of 0.8 mm is selected as the substrate material. The final configuration provides an operating frequency range of approximately 3-11.2 GHz, and its performance is evaluated in terms of reflection coefficient, gain, surface current distribution, and far-field radiation characteristics.

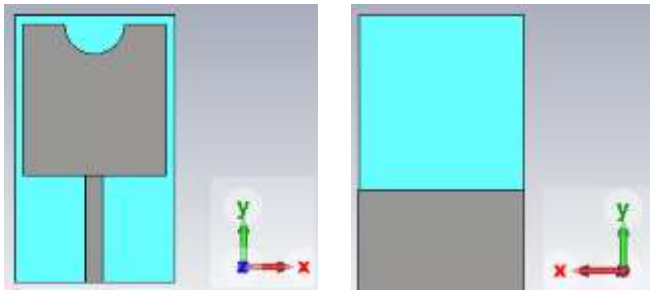
II. Materials and Method

The antenna was developed through three successive configurations. First, a notched planar radiator with a partial ground plane was defined. The lower part of the radiator was then reshaped with stepped transitions, and finally the ground conductor was modified by introducing the proposed DGS. A 0.8-mm-thick PREPERM ABS300 layer with $\varepsilon r = 3$ was used throughout the analysis. Each configuration was numerically examined to identify the influence of these changes on input matching, gain, current behavior, and far-field response.

Figure 1 illustrates the reference geometry used at the beginning of the design procedure. The radiating side contains a feed line connected to a rectangular metallic region whose upper boundary includes a semicircular cut. A partial conductor is placed on the opposite side as the ground plane. A partial ground plane is employed on the back side. This structure was selected as the reference antenna model for the subsequent geometric modifications.

The corresponding input response is given in Fig. 2. A resonance near 4 GHz produces an S11 minimum of approximately −15 dB; however, the −10 dB condition is maintained only over a relatively narrow interval. However, the -10 dB criterion is satisfied only over a limited frequency range, indicating insufficient bandwidth for the initial configuration. Therefore, in the following design stages, modifications were introduced to both the radiator geometry and the ground plane in order to enhance the operating bandwidth.

.



(a) (b)

Fig. 1 Initial antenna configuration: (a) front view and (b) back view

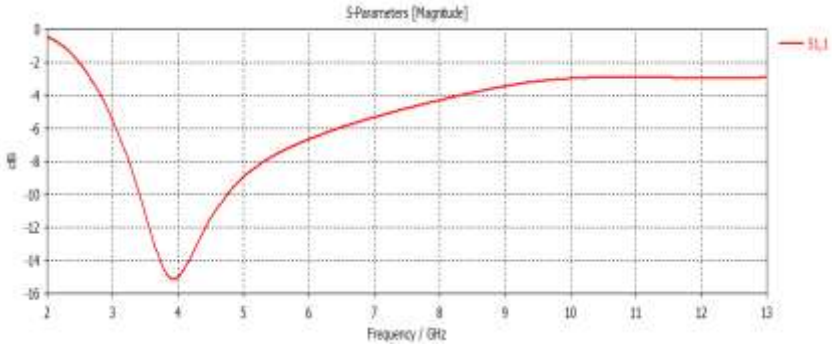
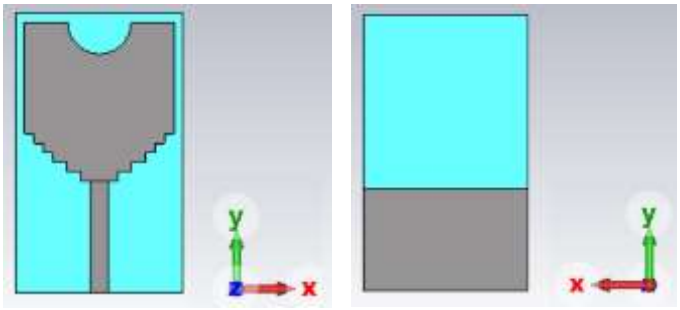


Fig. 2 Simulated reflection coefficient (S11) of the initial antenna structure

At this stage, stepped transitions were introduced along the lower edge of the initial radiator geometry in order to extend the current path and generate additional resonances. The front and back views of the modified structure are shown in Fig. 3(a) and Fig. 3(b), respectively. At this stage, the ground plane was kept unchanged, while only the radiator geometry was modified. As shown in Fig. 4, the stepped structure produces two distinct resonances around 3.5 GHz and 8.5 GHz, with good impedance matching at both frequency regions. However, the reflection coefficient rises above the -10 dB level between these two resonances, indicating that an additional modification is required to achieve continuous wideband operation. Therefore, in the next design stage, a DGS configuration was introduced on the ground plane.



(a) (b)

Fig. 3 Stepped radiator antenna configuration: (a) front view and (b) back view.

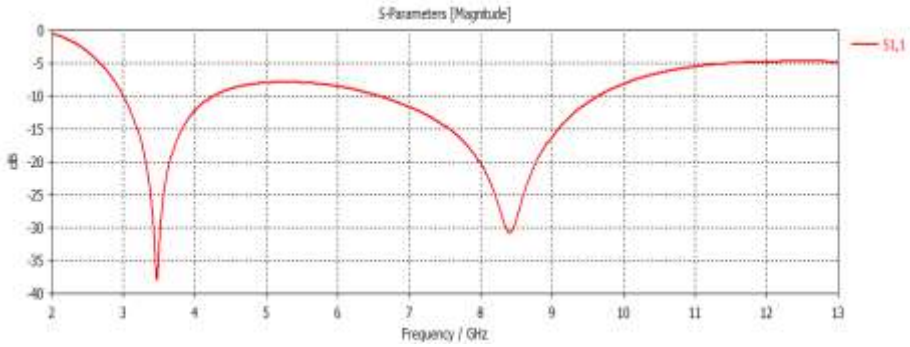
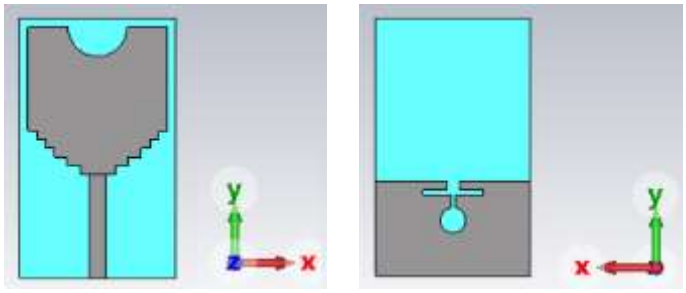


Fig. 4 Simulated reflection coefficient (S11) of the antenna with the stepped radiator geometry

The final antenna configuration was obtained by introducing a DGS into the ground plane to merge the two resonance regions observed in the previous design and achieve continuous wideband operation. The front and back views of the resulting structure are shown in Fig. 5(a) and Fig. 5(b), respectively. While the radiator geometry was preserved, a DGS configuration consisting of horizontal slots and a circular defect was added to the ground plane. This modification altered the surface current path on the ground plane, improving impedance matching and contributing to the merging of the resonance regions. As shown in Fig. 6, the reflection coefficient of the final design remains below -10 dB over approximately 3-11.2 GHz, indicating wideband operation. In particular, the S11 value reaches about -25 dB around 9.8 GHz. According to the maximum gain results presented in Fig. 7, the antenna gain varies approximately between 1.5 and 4.3 dBi depending on frequency, with the highest value obtained around 9.5 GHz.



(a) (b)

Fig. 5 Final antenna configuration with DGS: (a) front view and (b) back view.

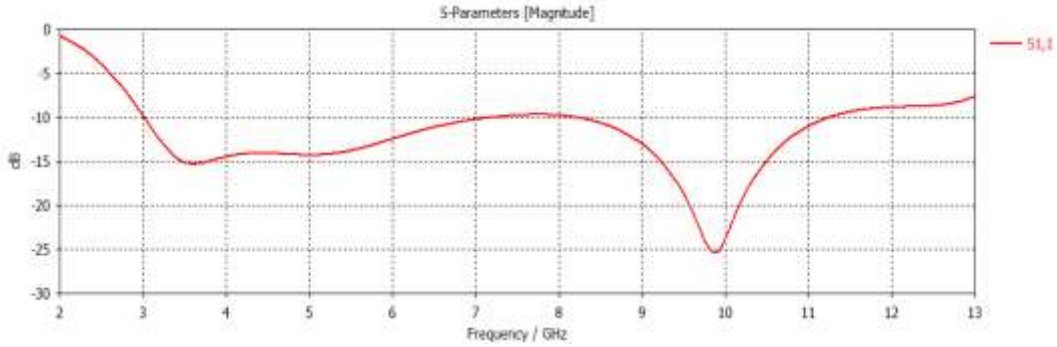


Fig. 6 Simulated reflection coefficient (S11) of the final DGS-based antenna

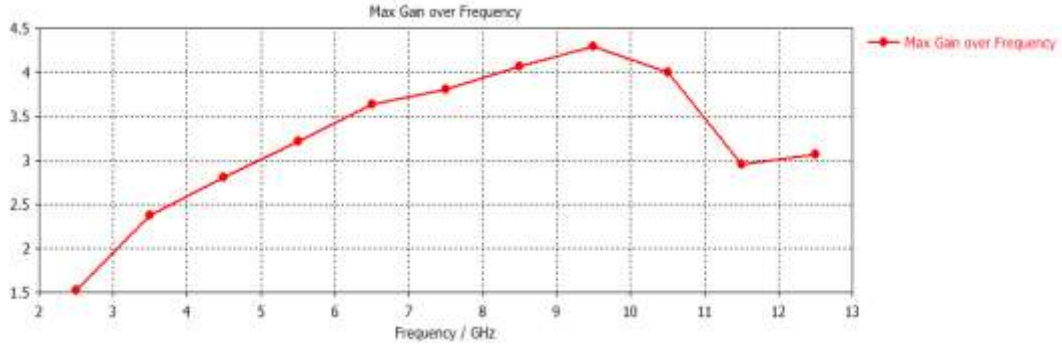
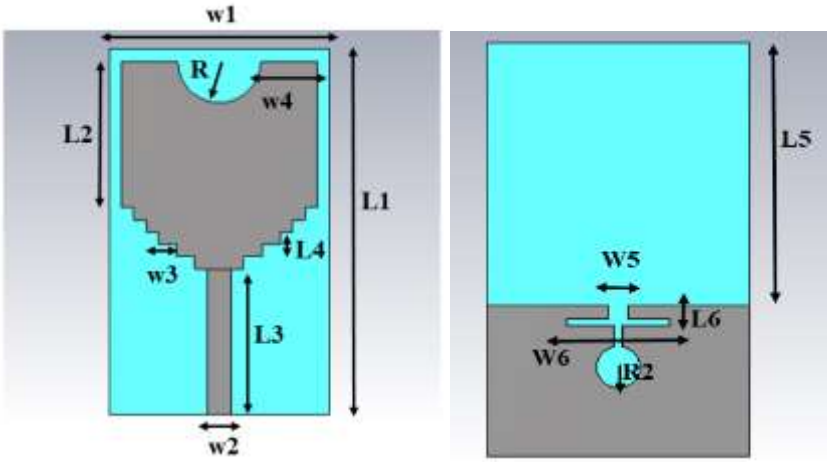


Fig. 7 Simulated maximum gain of the proposed antenna versus frequency

III. RESULTS

This section presents the final geometric configuration and simulation results of the proposed antenna. The final antenna geometry obtained after the radiator and DGS modifications is shown in Fig. 8. The geometrical parameters of the radiator and feed line are illustrated in Fig. 8(a), while the dimensions of the ground plane and DGS structure are defined in Fig. 8(b). The numerical values of all design parameters are listed in Table 1.



(a) (b)

Fig. 8 Geometrical parameters of the proposed antenna: (a) front view and (b) back view

Table 1. Geometrical dimensions of the proposed antenna

Parameter	Value (mm)	Parameter	Value (mm)
W1	18	L1	30
W2	2	L2	12
W3	1	L3	12
W4	4.6	L4	1
W5	1.4	L5	19
W6	7	L6	0.5
R	3.4	R2	1.5

IV. DISCUSSION

This section evaluates the surface current distribution and far-field radiation characteristics of the proposed antenna. The obtained results indicate that the DGS not only contributes to bandwidth enhancement but also modifies the current distribution on the ground plane.

Fig. 9 shows the surface current distribution at 7.5 GHz. The current is mainly concentrated around the feed line, the stepped lower region of the radiator, and the DGS elements. The slots and circular defect introduced in the ground plane alter the surface current path and affect the electromagnetic behavior of the antenna. This distribution confirms the active role of the DGS in achieving the wideband response.

Fig. 10 compares the far-field radiation patterns of the antenna at 5.5 and 9.5 GHz. As shown in Fig. 10(a), the maximum gain at 5.5 GHz is 3.04 dBi with a 3 dB beamwidth of 84.8°. At 9.5 GHz, Fig. 10(b) shows that the maximum gain increases to 4.25 dBi, while the beamwidth decreases to 43.3°. These results indicate that the antenna exhibits a more directive radiation behavior with increasing frequency, accompanied by an increase in gain.

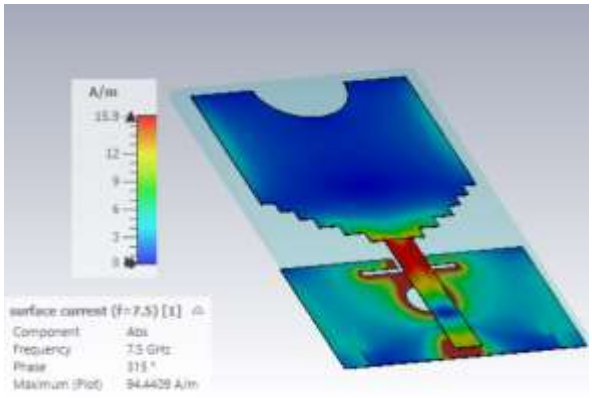
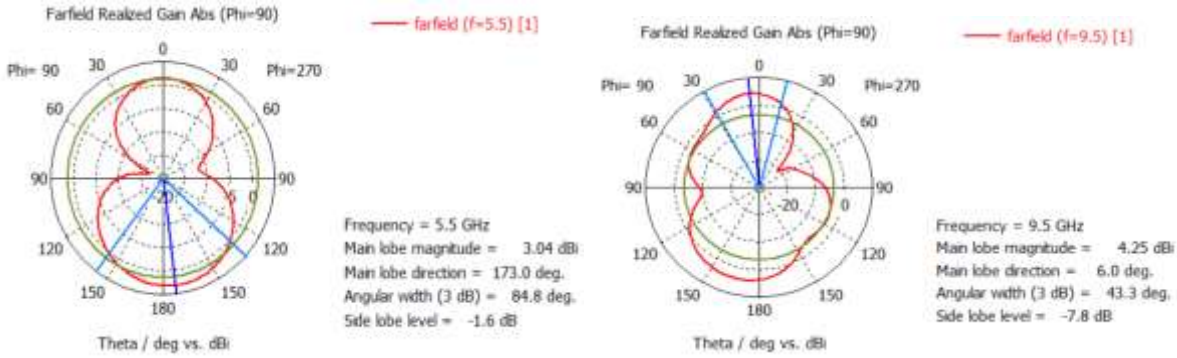


Fig. 9 Simulated surface current distribution of the proposed antenna at 7.5 GHz



(a) (b)

Fig. 10 Simulated far-field radiation patterns of the proposed antenna at (a) 5.5 GHz and (b) 9.5 GHz

V. CONCLUSION

This study presented a notched stepped antenna with a Defected Ground Structure (DGS) for wideband operation. By modifying the radiator geometry and introducing DGS elements on the ground plane, the antenna achieved an operating bandwidth of approximately 3-11.2 GHz with improved impedance matching. The simulated gain varied between about 1.5 and 4.3 dBi across the operating band. Surface current and radiation pattern results also confirmed the contribution of the proposed geometry to the antenna performance. Overall, the proposed design provides a simple and effective approach for wideband antenna applications.

REFERENCES

[1] W. Balani, M. Sarvagya, A. Samasgikar, T. Ali, and P. Kumar, "Design and analysis of super wideband antenna for microwave applications," *Sensors*, vol. 21, no. 2, Art. no. 477, 2021.

[2] S. Dey and N. C. Karmakar, "Design of novel super wide band antenna close to the fundamental dimension limit theory," *Scientific Reports*, vol. 10, Art. no. 16306, 2020.

[3] S. Ullah, C. Ruan, M. S. Sadiq, T. U. Haq, A. K. Fahad, and W. He, "Super wide band, defected ground structure (DGS), and stepped meander line antenna for WLAN/ISM/WiMAX/UWB and other wireless communication applications," *Sensors*, vol. 20, no. 6, Art. no. 1735, 2020.

[4] N. Castro, F. Pizarro, and E. Rajo-Iglesias, "High gain low profile horn array with circular polarization using a 3D printed anisotropic dielectric composite material at 38 GHz," *Scientific Reports*, vol. 12, Art. no. 18944, 2022.

[5] B. Dokmetas, "3D-printed cylindrical dielectric antenna optimized using honey bee mating optimization," *Electronics*, vol. 15, no. 2, Art. no. 393, 2026.

[6] M. Y. Onay and B. Dokmetas, “Experimental characterization of a 3D-printed conformal array antenna for 2.4 GHz WiFi backscatter,” *Electronics*, vol. 15, no. 8, Art. no. 1758, 2026.
[7] A. Singh and S. Singh, “A novel CPW-fed wideband printed monopole antenna with DGS,” *AEÜ—International Journal of Electronics and Communications*, vol. 69, no. 1, pp. 299–306, 2015.
[8] S. S. Bhatia, A. Sahni, and S. B. Rana, “A novel design of compact monopole antenna with defected ground plane for wideband applications,” *Progress In Electromagnetics Research M*, vol. 70, pp. 21–31, 2018.
[9] D. Gopi, A. R. Vadaboyina, and J. R. K. Kumar Dabbakuti, “DGS based monopole circular-shaped patch antenna for UWB applications,” *SN Applied Sciences*, vol. 3, Art. no. 198, 2021.
[10] S. Wang, F. Kong, K. Li, and L. Du, “A planar triple-band monopole antenna loaded with an arc-shaped defected ground plane for WLAN/WiMAX applications,” *International Journal of Microwave and Wireless Technologies*, vol. 13, no. 4, pp. 381–389, 2021.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Membranlı Süreçlerin Ağır Metal İçeren Atıksuların Arıtımında Kullanımı: Güncel Gelişmeler ve Gelecek Perspektifleri

Hatice DEMİR BECEREKLİ[1*], Yalçın ALTUNKAYNAK [2] ve Mutlu CANPOLAT [3]

[1]*Gıda İşleme / Teknik Bilimler Meslek Yüksekokulu, Batman Üniversitesi, Türkiye (ORCID: https://orcid.org/0000-0002-1165-518X)*

[2] *Kimya ve Kimyasal İşleme Teknolojileri / Teknik Bilimler Meslek Yüksekokulu, Batman Üniversitesi, Türkiye (ORCID: https://orcid.org/ https://orcid.org/0000-0003-2562-9297)*

[3] *Kimya ve Kimyasal İşleme Teknolojileri / Teknik Bilimler Meslek Yüksekokulu, Batman Üniversitesi, Türkiye (ORCID: https://orcid.org/ 0000-0002-3771-6737)*

[*]*hatice.becerekli@batman.edu.tr*

Özet **–** Hızlı sanayileşme, nüfus artışı ve teknolojik gelişmeler sonucunda ağır metal içeren atıksuların miktarı her geçen gün artmaktadır. Kadmiyum (Cd), kurşun (Pb), krom (Cr), cıva (Hg), arsenik (As) ve nikel (Ni) gibi toksik ağır metaller biyolojik olarak parçalanamadıkları için sucul ortamlarda uzun süre kalabilmekte ve canlı organizmalar üzerinde ciddi toksik etkilere neden olmaktadır. Geleneksel arıtma yöntemleri birçok uygulamada etkili olmasına rağmen düşük konsantrasyonlardaki metal iyonlarının gideriminde yetersiz kalabilmekte, yüksek kimyasal tüketimi ve fazla çamur oluşumu gibi dezavantajlar göstermektedir. Bu nedenle son yıllarda membran teknolojileri, yüksek ayırma verimi, kompakt tasarımı, otomasyona uygunluğu ve yüksek kaliteli arıtılmış su elde edebilmesi nedeniyle ileri arıtma prosesleri arasında önemli bir yer edinmiştir. Bu derleme çalışmasında mikrofiltrasyon (MF), ultrafiltrasyon (UF), nanofiltrasyon (NF) ve ters ozmoz (RO) süreçlerinin temel çalışma prensipleri, ağır metal giderimindeki performansları, avantajları, sınırlamaları ve son yıllardaki teknolojik gelişmeler kapsamlı olarak değerlendirilmiştir. Ayrıca membran kirlenmesi (fouling), enerji tüketimi, hibrit membran sistemleri ve gelecekte geliştirilebilecek yeni nesil membran malzemeleri de ele alınmıştır. Literatür değerlendirmeleri, özellikle nanofiltrasyon ve ters ozmoz proseslerinin çözünmüş ağır metal iyonlarının gideriminde %95'in üzerinde giderim verimleri sağlayabildiğini göstermektedir. Sonuç olarak membran teknolojilerinin sürdürülebilir su yönetimi ve döngüsel ekonomi hedefleri doğrultusunda gelecekte daha yaygın kullanılacağı öngörülmektedir.

Anahtar Kelimeler – Membran teknolojileri, ağır metal giderimi, nanofiltrasyon, ters ozmoz, ultrafiltrasyon, atıksu arıtımı

I.GİRİŞ

Temiz su kaynaklarının korunması günümüzde sürdürülebilir kalkınmanın temel unsurlarından biri olarak kabul edilmektedir. Artan endüstriyel faaliyetler, madencilik uygulamaları, metal kaplama tesisleri, deri sanayi, tekstil üretimi, akü üretimi ve kimya endüstrisi gibi sektörlerden kaynaklanan atıksular önemli miktarda ağır metal içermektedir. Ağır metaller biyolojik olarak parçalanmadıkları için çevrede uzun süre kalabilmekte ve besin zinciri boyunca biyobirikim ile biyobüyütme süreçlerine neden olmaktadır. Bu durum hem ekosistemlerin sürdürülebilirliğini hem de insan sağlığını tehdit eden önemli çevresel problemlerin ortaya çıkmasına yol açmaktadır [1,2]. Kurşun, kadmiyum, krom, cıva, arsenik ve nikel gibi ağır metaller düşük konsantrasyonlarda dahi toksik etki gösterebilmektedir. Özellikle altı değerlikli krom [Cr(VI)] güçlü oksitleyici özelliği nedeniyle kanserojen olarak sınıflandırılırken, kadmiyum böbrek fonksiyon bozukluklarına, kurşun sinir sistemi hasarına ve cıva ise merkezi sinir sistemi üzerinde kalıcı etkilere neden olabilmektedir. Dünya Sağlık Örgütü (WHO) ve birçok uluslararası çevre kuruluşu içme sularındaki ağır metal konsantrasyonları için oldukça düşük sınır değerler belirlemiştir. Bu nedenle ağır

metal içeren atıksuların çevreye deşarj edilmeden önce etkin biçimde arıtılması zorunlu hale gelmiştir [3]. Ağır metal gideriminde kimyasal çöktürme, iyon değiştirme, koagülasyon-flokülasyon, elektrokimyasal yöntemler, adsorpsiyon ve biyolojik arıtma gibi çok sayıda yöntem kullanılmaktadır. Bununla birlikte bu yöntemlerin önemli bir kısmı yüksek kimyasal tüketimi, fazla çamur oluşumu, düşük seçicilik veya düşük metal konsantrasyonlarında yetersiz giderim performansı gibi çeşitli sınırlamalara sahiptir. Son yıllarda membran ayırma teknolojileri bu dezavantajların önemli bir bölümünü ortadan kaldırabilecek ileri arıtma yöntemleri olarak dikkat çekmektedir [2,3] Membran prosesleri, iki faz arasında seçici geçirgenliğe sahip yarı geçirgen bir bariyer kullanarak çözünmüş veya askıda bulunan kirleticilerin ayrılmasına dayanmaktadır. Basınç, konsantrasyon farkı, elektrik potansiyeli veya sıcaklık farkı gibi itici kuvvetler kullanılarak gerçekleştirilen bu süreçler yüksek ayırma verimi, düşük kimyasal kullanımı, sürekli işletmeye uygunluk ve otomasyona elverişli olmaları nedeniyle hem endüstriyel hem de belediye ölçekli arıtma tesislerinde yaygınlaşmaktadır [4]. Son yıllarda membran malzemelerinde meydana gelen gelişmeler, nanoteknoloji uygulamaları, yüzey modifikasyon teknikleri ve hibrit arıtma sistemlerinin geliştirilmesi sayesinde membran teknolojilerinin performansı önemli ölçüde artmıştır. Özellikle adsorpsiyon–membran, membran biyoreaktör (MBR), ileri oksidasyon–membran ve elektro-membran hibrit sistemleri yüksek arıtma verimleri sunarak sürdürülebilir su yönetiminde yeni fırsatlar ortaya koymaktadır [5,6]. Bu derleme çalışmasının amacı; ağır metal içeren atıksuların arıtımında kullanılan membran teknolojilerinin çalışma prensiplerini, uygulama alanlarını, avantajlarını, sınırlamalarını ve güncel gelişmeleri kapsamlı biçimde değerlendirmek, ayrıca gelecekte bu alanda yapılabilecek araştırmalara yönelik öneriler sunmaktır [1,7].

II. MEMBRAN TEKNOLOJİLERİNİN TEMEL PRENSİPLERİ

Membran teknolojileri, iki faz arasında bulunan yarı geçirgen bir membran aracılığıyla çözünmüş veya askıda bulunan bileşenlerin seçici olarak ayrılmasını sağlayan ileri ayırma prosesleridir. Bu teknolojilerde ayırma işlemi, moleküllerin boyutu, elektriksel yükü, difüzyon katsayısı ve membran ile kirleticiler arasındaki fizikokimyasal etkileşimler temel alınarak gerçekleştirilmektedir. Membran sistemlerinde en yaygın kullanılan itici kuvvet transmembran basıncı olmakla birlikte konsantrasyon farkı, sıcaklık farkı ve elektrik potansiyeli de ayırma mekanizmasını sağlayan diğer önemli kuvvetlerdir [4]. Membran proseslerinin temel çalışma prensibi, besleme çözeltisinin (feed) membran yüzeyine uygulanması ve seçici geçirgenliğe sahip membran aracılığıyla saflaştırılmış suyun (permeat) geçirgen tarafta toplanırken kirleticilerin ise yoğunlaştırılmış akım (retentat veya konsantre) içerisinde tutulmasına dayanmaktadır. Ayırma performansı; membranın gözenek boyutu, yüzey yükü, hidrofobik/hidrofilik karakteri, işletme basıncı, sıcaklık, pH ve besleme çözeltisinin özelliklerinden önemli ölçüde etkilenmektedir [8]. Membranlar üretim materyallerine göre polimerik, seramik, metalik ve kompozit membranlar olarak sınıflandırılmaktadır. Günümüzde polisülfon (PSf), polieter sülfon (PES), poliviniliden florür (PVDF), selüloz asetat (CA) ve poliamid (PA) esaslı polimerik membranlar ekonomik olmaları ve kolay üretilebilmeleri nedeniyle yaygın olarak kullanılmaktadır. Bununla birlikte yüksek sıcaklık, aşındırıcı kimyasallar ve mekanik dayanım gerektiren uygulamalarda alümina, zirkonya ve titanyum dioksit esaslı seramik membranlar tercih edilmektedir [9]. Membran ayırma teknolojileri, gözenek çapı ve ayırma mekanizmasına göre mikrofiltrasyon (MF), ultrafiltrasyon (UF), nanofiltrasyon (NF) ve ters ozmoz (RO) olarak dört temel grupta incelenmektedir [10]. Genel olarak gözenek çapı küçüldükçe daha küçük kirleticilerin giderimi mümkün olmakta; buna karşılık gerekli işletme basıncı ve enerji tüketimi artmaktadır. Bu nedenle uygulanacak membran prosesinin seçiminde yalnızca giderim verimi değil, işletme maliyeti, enerji gereksinimi, membran ömrü ve bakım maliyetleri de dikkate alınmalıdır [11].

A. Mikrofiltrasyon (MF)

Mikrofiltrasyon, yaklaşık 0,05–10 μm gözenek çapına sahip membranların kullanıldığı ve düşük işletme basınçlarında (0,1–2 bar) gerçekleştirilen en temel membran proseslerinden biridir. Bu yöntem esas olarak askıda katı maddelerin, bulanıklığın, alglerin, bakterilerin ve büyük kolloidal parçacıkların

uzaklaştırılmasında kullanılmaktadır [7]. Mikrofiltrasyon membranları çözünmüş ağır metal iyonlarını doğrudan uzaklaştıramamakla birlikte, metal hidroksit çökeleklerinin, metal oksit partiküllerinin ve adsorban yüzeyine bağlanmış metal iyonlarının ayrılmasında oldukça başarılı sonuçlar vermektedir. Bu nedenle MF prosesi çoğu zaman kimyasal çöktürme veya adsorpsiyon gibi ön arıtma yöntemleri ile birlikte kullanılmaktadır [12]. Mikrofiltrasyonun en önemli avantajları düşük enerji tüketimi, yüksek su akısı (flux), kolay işletim ve düşük yatırım maliyetidir. Ancak çözünmüş iyonların gideriminde yetersiz kalması ve membran kirlenmesine karşı hassas olması en önemli dezavantajları arasında yer almaktadır [8].

B. Ultrafiltrasyon (UF)

Ultrafiltrasyon membranları yaklaşık 2–100 nm gözenek çapına sahip olup genellikle 1–10 bar basınç altında çalışmaktadır. UF sistemleri proteinler, makromoleküller, virüsler, kolloidler ve yüksek molekül ağırlıklı organik bileşiklerin uzaklaştırılmasında yaygın olarak kullanılmaktadır [4,8]. Çözünmüş ağır metal iyonları moleküler boyutlarının oldukça küçük olması nedeniyle ultrafiltrasyon membranlarından geçebilmektedir. Bununla birlikte son yıllarda geliştirilen polimer destekli ultrafiltrasyon (Polymer Enhanced Ultrafiltration, PEUF) ve misel destekli ultrafiltrasyon (Micellar Enhanced Ultrafiltration, MEUF) teknikleri sayesinde ağır metal iyonlarının uzaklaştırılmasında önemli ilerlemeler kaydedilmiştir [13]. PEUF yönteminde suda çözünebilen polimerler ağır metal iyonları ile kompleks oluşturarak molekül boyutunu artırmakta ve oluşan kompleksler UF membranı tarafından tutulmaktadır. MEUF yönteminde ise yüzey aktif maddeler kritik misel konsantrasyonunun üzerinde kullanılarak metal iyonlarının misellere bağlanması sağlanmakta ve bu kompleksler ultrafiltrasyon membranları tarafından uzaklaştırılmaktadır [12,13]. Bu hibrit yaklaşımlar özellikle Cd^{2+}, Pb^{2+}, Cu^{2+} ve Ni^{2+} iyonlarının gideriminde yüksek verimler sağlayabilmektedir [1,13].

C. Nanofiltrasyon (NF)

Nanofiltrasyon günümüzde ağır metal gideriminde en yaygın kullanılan membran teknolojilerinden biridir. NF membranlarının etkin gözenek çapı yaklaşık 0,5–2 nm arasında değişmekte olup işletme basınçları genellikle 5–25 bar aralığındadır [3,4]. Nanofiltrasyonda ayırma mekanizması yalnızca molekül boyutuna bağlı değildir. Membran yüzeyinin elektriksel yükü nedeniyle Donnan dışlama mekanizması da önemli rol oynamaktadır. Negatif yüklü membran yüzeyleri özellikle çok değerlikli katyonların geçişini önemli ölçüde sınırlandırmaktadır [15]. Bu nedenle Cd^{2+}, Pb^{2+}, Cr^{3+}, Cu^{2+} ve Zn^{2+} gibi iki ve üç değerlikli metal iyonlarının gideriminde nanofiltrasyon oldukça yüksek performans göstermektedir. Birçok araştırmada uygun işletme koşullarında %90–99 arasında ağır metal giderim verimleri rapor edilmiştir [14]. Nanofiltrasyonun başlıca avantajları şunlardır: Düşük enerji tüketimi (RO'ya göre), Yüksek ağır metal giderim verimi, Organik maddelerin eş zamanlı uzaklaştırılması, Sertlik giderimi, Yüksek su geri kazanımı. Bununla birlikte yüksek konsantrasyonlu besleme sularında membran kirlenmesi ve konsantrasyon polarizasyonu performansı olumsuz etkileyebilmektedir [6,7].

D. Ters Ozmoz (RO)

Ters ozmoz, günümüzde su arıtma teknolojileri içerisinde en yüksek ayırma verimine sahip membran prosesidir. Bu yöntemde yarı geçirgen membran üzerine ozmotik basınçtan daha yüksek bir basınç uygulanarak yalnızca su moleküllerinin membranı geçmesi sağlanmaktadır [16]. RO sistemleri genellikle 15–80 bar arasında değişen yüksek işletme basınçlarında çalışmaktadır. Membranların etkin gözenek çapı yaklaşık 0,1 nm düzeyinde olduğundan çözünmüş iyonların büyük bölümü tutulabilmektedir [17]. Ters ozmoz teknolojisi; deniz suyunun tuzdan arındırılması, ultra saf su üretimi, içme suyu hazırlanması, endüstriyel proses suyu geri kazanımı, ağır metal içeren atıksuların arıtılması gibi birçok alanda başarıyla uygulanmaktadır. Cd^{2+}, Pb^{2+}, Hg^{2+}, Cr^{3+}, Cr(VI), As(V), Ni^{2+} ve Cu^{2+} gibi ağır metaller için bildirilen giderim verimleri çoğunlukla %95–99,5 aralığındadır. Bu yüksek performans sayesinde RO sistemleri özellikle elektronik, metal kaplama ve madencilik endüstrilerinde ileri arıtma basamağı olarak tercih edilmektedir [13,17]. Bununla birlikte yüksek enerji tüketimi, membran maliyetleri, konsantre akımın

bertaraf edilmesi ve membran kirlenmesi ters ozmozun en önemli dezavantajları arasında yer almaktadır. Son yıllarda enerji geri kazanım sistemleri, düşük basınçlı kompozit membranlar ve yüzeyi modifiye edilmiş ince film kompozit membranların geliştirilmesi ile bu dezavantajların önemli ölçüde azaltılması hedeflenmektedir [18].

Tablo 1. Membran proseslerinin genel özelliklerinin karşılaştırılması

Proses	Gözenek Boyutu	İşletme Basıncı	Giderilen Başlıca Kirleticiler
Mikrofiltrasyon (MF)	0,05–10 µm	0,1–2 bar	Askıda katılar, bakteriler
Ultrafiltrasyon (UF)	2–100 nm	1–10 bar	Proteinler, kolloidler, virüsler
Nanofiltrasyon (NF)	0,5–2 nm	5–25 bar	Çok değerlikli iyonlar, organikler
Ters Ozmoz (RO)	≈0,1 nm	15–80 bar	Tuzlar, ağır metaller, çözünmüş iyonlar

III. AĞIR METAL GİDERİMİNDE MEMBRAN TEKNOLOJİLERİNİN UYGULAMALARI

Ağır metallerin çevresel ortamlarda uzun süre kalıcılık göstermeleri, biyolojik olarak parçalanamamaları ve canlı organizmalarda biyobirikime neden olmaları nedeniyle atıksuların etkin şekilde arıtılması büyük önem taşımaktadır. Membran teknolojileri, özellikle düşük konsantrasyonlardaki ağır metal iyonlarının gideriminde yüksek ayırma verimleri sağlamaları nedeniyle son yıllarda geleneksel yöntemlere alternatif veya tamamlayıcı ileri arıtma teknolojileri olarak öne çıkmaktadır [2,3]. Membran proseslerinin ağır metal giderimindeki başarısı; metal iyonunun hidratlanmış çapı, iyon yükü, membran yüzey yükü, işletme basıncı, çözeltinin pH değeri, iyonik kuvvet ve besleme çözeltisinin bileşimine bağlı olarak değişmektedir. Özellikle nanofiltrasyon ve ters ozmoz sistemlerinde hem boyut eleme (size exclusion) hem de elektrostatik dışlama (Donnan exclusion) mekanizmaları birlikte etkili olmaktadır [14].

A. Kurşun (Pb^{2+}) Giderimi

Kurşun; akü üretimi, metal kaplama, boya, seramik ve madencilik endüstrilerinden kaynaklanan en önemli ağır metal kirleticilerinden biridir. İnsanlarda sinir sistemi bozuklukları, böbrek hasarı ve gelişimsel problemlere neden olması nedeniyle içme sularındaki izin verilen konsantrasyonu oldukça düşüktür [1,3]. Nanofiltrasyon membranları, iki değerlikli kurşun iyonlarının uzaklaştırılmasında yüksek performans göstermektedir. Negatif yüklü membran yüzeyi sayesinde Pb^{2+} iyonları elektrostatik olarak uzaklaştırılmakta ve %95–99 seviyelerinde giderim verimleri elde edilmektedir. Ters ozmoz sistemlerinde ise uygun işletme koşullarında %99'un üzerinde kurşun giderimi rapor edilmektedir [13,14]. Son yıllarda grafen oksit, titanyum dioksit (TiO_2), silika nanoparçacıkları ve karbon nanotüpler ile modifiye edilen kompozit membranların kurşun giderim performansını önemli ölçüde artırdığı bildirilmektedir. Bu tür membranlar daha yüksek su geçirgenliği sağlarken aynı zamanda fouling eğilimini de azaltmaktadır [9].

B. Kadmiyum (Cd^{2+}) Giderimi

Kadmiyum; pil üretimi, galvanizleme, fosfat gübreleri ve metal işleme endüstrilerinde yaygın olarak kullanılan toksik bir ağır metaldir. Böbrek fonksiyon bozuklukları ve kemik deformasyonlarına neden olabilmesi nedeniyle çevresel açıdan büyük önem taşımaktadır [1,3]. Nanofiltrasyon membranları kadmiyum iyonlarının uzaklaştırılmasında oldukça başarılı sonuçlar vermektedir. Çalışmalarda uygun pH ve işletme basıncı altında %90–98 arasında değişen giderim verimleri rapor edilmiştir [1,10,14]. Son yıllarda adsorpsiyon ile membran proseslerinin birlikte kullanıldığı hibrit sistemler dikkat çekmektedir. Özellikle biyokömür (biochar), aktif karbon ve biyosorbentler kullanılarak gerçekleştirilen ön adsorpsiyon işlemi, membrana ulaşan metal yükünü azaltmakta ve membran ömrünü uzatmaktadır. Böylece hem daha yüksek giderim verimi hem de daha düşük işletme maliyeti elde edilmektedir [19,20].

C. Krom Giderimi

Krom endüstriyel atıksularda hem üç değerlikli [Cr(III)] hem de altı değerlikli [Cr(VI)] formunda bulunabilmektedir. Cr(VI), yüksek çözünürlüğü ve güçlü oksitleyici özelliği nedeniyle çevresel açıdan en tehlikeli ağır metallerden biridir [3,13]. Nanofiltrasyon membranlarında Cr(VI) iyonlarının giderim mekanizması büyük ölçüde elektrostatik dışlamaya dayanmaktadır. Alkali pH değerlerinde negatif yüklü kromat iyonlarının membran tarafından uzaklaştırılması daha yüksek olmaktadır. Cr(III) iyonlarının gideriminde ise iyon hidrat çapı ve membran gözenek boyutu belirleyici olmaktadır [1,14]. Ters ozmoz sistemleri hem Cr(III) hem de Cr(VI) türlerinin gideriminde %98–99'un üzerinde verim sağlayabilmektedir. Bu nedenle metal kaplama ve deri sanayi atıksularının ileri arıtılmasında yaygın olarak kullanılmaktadır [16].

D. Bakır, Nikel ve Çinko Giderimi

Bakır, nikel ve çinko iyonları; elektronik, metal kaplama, madencilik ve alaşım üretim tesislerinden kaynaklanan atıksularda yaygın olarak bulunmaktadır [3]. Nanofiltrasyon proseslerinde bu metaller için genellikle %90'ın üzerinde giderim verimleri elde edilmektedir. Ancak iyon yarıçapları ve hidratlanma özellikleri nedeniyle giderim sıralaması işletme koşullarına bağlı olarak değişebilmektedir [10]. Çok bileşenli çözeltilerde iyonlar arasında rekabet meydana gelmesi membran performansını etkileyen önemli faktörlerden biridir. Bu nedenle gerçek endüstriyel atıksular üzerinde yapılan çalışmalar laboratuvar ölçekli sentetik çözeltilere göre daha büyük önem taşımaktadır [1,10].

E. Arsenik Giderimi

Arsenik, doğal jeolojik oluşumlar yanında madencilik faaliyetleri ve metalurjik işlemler sonucunda yeraltı sularına karışabilmektedir [1,3,13]. Arseniğin özellikle As(V) formu nanofiltrasyon ve ters ozmoz membranları ile yüksek verimle uzaklaştırılabilmektedir. Ancak As(III) yüksüz yapıda bulunduğundan giderimi daha zordur. Bu nedenle çoğu uygulamada arsenit önce oksidasyon işlemi ile arsenata dönüştürülmekte, ardından membran prosesi uygulanmaktadır [16].

F. Hibrit Membran Sistemleri

Son yıllarda yalnızca membran prosesleri yerine hibrit arıtma sistemlerinin kullanımı hızla artmaktadır. Bu sistemlerde membran teknolojileri adsorpsiyon, iyon değiştirme, ileri oksidasyon prosesleri veya biyolojik arıtma ile birleştirilmektedir [12,19]. En yaygın hibrit sistemler şunlardır: Adsorpsiyon+Nanofiltrasyon, Adsorpsiyon+Ters Ozmoz, Membran Biyoreaktör (MBR), Elektrokoagülasyon+UF, İleri Oksidasyon+UF/NF, Fotokataliz+Membran. Bu sistemlerde ön arıtma basamağı membran üzerine gelen kirletici yükünü azaltarak membran kirlenmesini önemli ölçüde düşürmektedir. Bunun sonucunda membran ömrü uzamakta, enerji tüketimi azalmakta ve işletme maliyetleri düşmektedir [11,18,19]. Özellikle son yıllarda biyokütle kökenli aktif karbonlar, biyokömürler ve tarımsal atıklardan elde edilen adsorbanların membran sistemleriyle birlikte kullanıldığı hibrit uygulamalar sürdürülebilir su arıtımı açısından büyük ilgi görmektedir. Bu yaklaşım, hem atıkların katma değerli malzemelere dönüştürülmesini sağlamakta hem de ağır metal giderim performansını artırmaktadır. Bu nedenle döngüsel ekonomi ve yeşil kimya ilkeleri ile uyumlu bir çözüm olarak değerlendirilmektedir [15,19,20].

Tablo 2. Ağır metal gideriminde membran proseslerinin genel performans karşılaştırması

Ağır Metal	Tercih Edilen Proses	Ortalama Giderim (%)	Başlıca Avantaj
Pb^{2+}	NF / RO	95–99	Yüksek seçicilik
Cd^{2+}	NF / RO	90–98	Düşük konsantrasyonlarda etkili
Cr(III)	NF / RO	95–99	Yüksek geri kazanım
Cr(VI)	NF / RO	97–99	Elektrostatik dışlama etkili
Cu^{2+}	NF	90–98	Düşük enerji tüketimi
Ni^{2+}	NF	90–97	Endüstriyel uygulamalara uygun

Tablo 2. Ağır metal gideriminde membran proseslerinin genel performans karşılaştırması

Ağır Metal	Tercih Edilen Proses	Ortalama Giderim (%)	Başlıca Avantaj
Zn^{2+}	NF	90–98	Yüksek akı değerleri
As(V)	RO / NF	95–99	İçme suyu uygulamalarında başarılı

IV. MEMBRAN KİRLENMESİ (FOULİNG) VE KONTROL YÖNTEMLERİ

Membran teknolojilerinin endüstriyel ölçekte uygulanmasını sınırlandıran en önemli problemlerden biri membran kirlenmesidir (fouling). Fouling, besleme çözeltisinde bulunan partiküllerin, organik maddelerin, inorganik çökeltilerin ve mikroorganizmaların membran yüzeyinde veya gözeneklerinde birikmesi sonucu membran performansının zamanla azalması olarak tanımlanmaktadır. Membran kirlenmesi; permeat akısında azalma, transmembran basıncında artış, enerji tüketiminde yükselme ve membran ömründe kısalma gibi olumsuz sonuçlara yol açmaktadır [11].

Fouling genel olarak dört temel grupta incelenmektedir:

1. Organik fouling: Hümik maddeler, proteinler, yağlar ve doğal organik maddelerin membran yüzeyine adsorplanması sonucu oluşur [18]. **2.**İnorganik fouling (Scaling): Kalsiyum karbonat ($CaCO_3$), kalsiyum sülfat ($CaSO_4$), silika ve metal hidroksitlerin çökelmesi ile meydana gelir [11]. **3.** Biyolojik fouling (Biofouling): Mikroorganizmaların membran yüzeyinde biyofilm oluşturması sonucu gelişir [18]. **4.** Partikül fouling: Askıda katı maddelerin ve kolloidal partiküllerin membran gözeneklerini tıkaması ile oluşur [18].

Özellikle ağır metal içeren endüstriyel atıksularda metal hidroksit çökelekleri ve çözünmeyen metal bileşikleri inorganik fouling oluşumunu hızlandırmaktadır. Bunun yanında organik kirleticiler ile metal iyonlarının kompleks oluşturması membran yüzeyinde daha yoğun bir kirlenme tabakasının oluşmasına neden olabilmektedir [6]. Fouling oluşumunu azaltmak amacıyla günümüzde aşağıdaki yöntemler yaygın olarak uygulanmaktadır: Ön filtrasyon sistemlerinin kullanılması, pH optimizasyonu, Koagülasyon–flokülasyon ön arıtımı, Adsorpsiyon ön arıtımı, Periyodik geri yıkama (backwashing), Kimyasal temizleme (asit, baz ve oksitleyiciler), Hidrofilik yüzey modifikasyonu, Nanomalzeme katkılı antifouling membranlar. Son yıllarda grafen oksit, TiO_2, Ag nanopartikülleri ve ZnO içeren nanokompozit membranların hem antibakteriyel özellik gösterdiği hem de organik kirlenmeye karşı daha dirençli olduğu bildirilmektedir [5,9].

V. MEMBRAN TEKNOLOJİLERİNİN AVANTAJLARI VE DEZAVANTAJLARI

Membran prosesleri, geleneksel su arıtma yöntemlerine kıyasla birçok önemli üstünlüğe sahiptir. Bunların başında yüksek ayırma verimi, düşük kimyasal tüketimi, sürekli işletmeye uygunluk ve yüksek kaliteli arıtılmış su elde edilebilmesi gelmektedir. Ayrıca modüler yapıları sayesinde kapasite artırımı kolaydır ve otomasyon sistemleri ile entegre çalışabilmektedir [4,12]. Bununla birlikte membran teknolojilerinin yaygın kullanımını sınırlayan bazı dezavantajlar da bulunmaktadır. Özellikle yüksek yatırım maliyetleri, enerji tüketimi, membran kirlenmesi ve konsantre akımın yönetimi çözülmesi gereken temel sorunlar arasında yer almaktadır [17].

Tablo 3. Membran teknolojilerinin avantaj ve dezavantajları

Avantajlar	Dezavantajlar
Yüksek giderim verimi	Membran kirlenmesi
Düşük kimyasal tüketimi	Yüksek yatırım maliyeti
Sürekli işletim	Enerji tüketimi
Modüler tasarım	Konsantre atık oluşumu
Su geri kazanımı	Düzenli bakım gereksinimi

Tablo 3. Membran teknolojilerinin avantaj ve dezavantajları	
Avantajlar	Dezavantajlar
Otomasyona uygunluk	Membran değişim maliyeti

Teknolojik gelişmeler sayesinde düşük basınçta çalışan yeni nesil membranlar ve enerji geri kazanım sistemleri geliştirilmekte olup bu dezavantajların önemli ölçüde azaltılması hedeflenmektedir [6,16].

VI. GELECEK PERSPEKTİFLERİ

Membran teknolojileri son yıllarda yalnızca su arıtımı amacıyla değil, aynı zamanda kaynak geri kazanımı, döngüsel ekonomi ve sürdürülebilir üretim yaklaşımlarının önemli bir bileşeni olarak değerlendirilmektedir [2,19]. Gelecekte araştırmaların aşağıdaki alanlarda yoğunlaşması beklenmektedir: Antifouling özellikli yeni nesil membranların geliştirilmesi [6,9], Grafen, MXene ve metal-organik kafes (MOF) tabanlı nanokompozit membranlar [15], Yapay zekâ destekli proses kontrol sistemleri [9], Düşük enerji tüketen membran tasarımları [8], Güneş enerjisi destekli membran sistemleri [17], Atıksulardan değerli metal geri kazanımı [13], Membran–adsorpsiyon hibrit proseslerinin endüstriyel ölçek uygulamaları [15]. Özellikle tarımsal atıklardan elde edilen aktif karbonlar, biyokömürler ve biyokompozit adsorbanların membran sistemleri ile birlikte kullanılması hem ekonomik hem de çevresel açıdan sürdürülebilir çözümler sunmaktadır. Bu yaklaşım, atıkların katma değerli ürünlere dönüştürülmesini sağlayarak döngüsel ekonomi ilkelerine katkıda bulunmaktadır [20]. Bunun yanı sıra, yaşam döngüsü değerlendirmesi (Life Cycle Assessment, LCA) ve tekno-ekonomik analizlerin yeni membran teknolojilerinin geliştirilmesi sürecine entegre edilmesi, bu sistemlerin gerçek endüstriyel koşullardaki uygulanabilirliğinin daha doğru değerlendirilmesine katkı sağlayacaktır [19].

VII. SONUÇLAR

Ağır metal içeren atıksuların etkin şekilde arıtılması çevrenin korunması ve sürdürülebilir su yönetimi açısından kritik öneme sahiptir. Membran teknolojileri, özellikle nanofiltrasyon ve ters ozmoz sistemleri, yüksek ayırma verimleri sayesinde ağır metal gideriminde en etkili ileri arıtma yöntemleri arasında yer almaktadır. Mikrofiltrasyon ve ultrafiltrasyon sistemleri daha çok ön arıtma veya hibrit proseslerde kullanılırken, nanofiltrasyon ve ters ozmoz çözünmüş ağır metal iyonlarının uzaklaştırılmasında üstün performans göstermektedir. Bununla birlikte membran kirlenmesi, enerji tüketimi ve konsantre akımın yönetimi bu teknolojilerin karşı karşıya olduğu temel teknik zorluklardır. Son yıllarda geliştirilen nanokompozit membranlar, yüzey modifikasyon teknikleri ve hibrit membran sistemleri sayesinde bu sınırlamaların önemli ölçüde azaltıldığı görülmektedir. Özellikle adsorpsiyon–membran hibrit sistemleri, hem membran kirlenmesini azaltmaları hem de ağır metal giderim verimini artırmaları bakımından dikkat çekmektedir. Gelecekte çevre dostu membran malzemelerinin geliştirilmesi, enerji tüketiminin azaltılması ve yapay zekâ destekli işletme sistemlerinin yaygınlaştırılması ile membran teknolojilerinin sürdürülebilir su arıtımında çok daha önemli bir konuma ulaşacağı öngörülmektedir. Ayrıca gerçek endüstriyel atıksular üzerinde gerçekleştirilecek uzun süreli pilot ölçekli çalışmaların, laboratuvar ölçeğinde elde edilen sonuçların uygulamaya aktarılmasında kritik rol oynayacağı değerlendirilmektedir.

KAYNAKLAR

[1] Zheng X, et al. Heavy metal removal using membrane processes. *J Water Process Eng*. 2023.

[2] Shannon MA, et al. Science and technology for water purification. *Nature*.

[3] Fu F, Wang Q. Heavy metal removal technologies. *J Environ Manage*.

[4] Baker RW. *Membrane Technology and Applications*. 4th ed. Wiley; 2023.

[5] Wang Z, et al. Graphene oxide modified membranes for water purification. *Chem Eng J*. 2024.

[6] Zhang H, et al. Fouling-resistant nanocomposite membranes. *J Membr Sci*. 2024.

[7] Mulder M. *Basic Principles of Membrane Technology*. Springer.

[8] Singh R. *Membrane Technology and Engineering for Water Purification*. CRC Press.

[9] Goh PS, et al. Nanomaterial-based membranes. *Prog Mater Sci*. 2025.

[10] Mohammad AW, et al. Recent advances in nanofiltration membranes. *J Membr Sci*. 2023.

[11] Zhao S, et al. Antifouling membrane materials. *Chemosphere*. 2023.

[12] Hilal N, et al. Membrane processes in wastewater treatment. *Desalination*.

[13] Lee KP, et al. Membrane technologies for heavy metal removal. *Sep Purif Rev*.

[14] Van der Bruggen B. Nanofiltration for water purification: current status and future perspectives. *Sep Purif Technol*. 2022.

[15] Ali I. New generation adsorbents and membranes. *Chem Rev*.

[16] Li N, et al. Advances in reverse osmosis membranes. *Desalination*. 2024.

[17] Elimelech M, Phillip WA. The future of seawater desalination. *Science*.

[18] Cui Z, et al. Membrane fouling mechanisms. *J Membr Sci*. 2022.

[19] Guo W, et al. Hybrid membrane technologies. *Water Res*. 2022.

[20] Wang J, Chen C. Biosorbents for heavy metals. *Biotechnol Adv*.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Biopolymer Additives for Biological Soil Stabilization: Artificial Neural Network Prediction of Unconfined Compressive Strength

Bayram Ateş [1*], Mohammad Azim Eirgash [1]

[1]Civil Engineering Department, Karadeniz Technical University, Trabzon, Türkiye
[*](bayramates61@hotmail.com)

Abstract – The growing demand for sustainable ground improvement has increased interest in biological and bio-inspired alternatives to conventional soil stabilization. This study develops an artificial neural network (ANN) model for predicting the unconfined compressive strength (UCS) of biopolymer-treated soil using published experimental data. The database contains 38 UCS observations for residual soil treated with xanthan gum (XG), guar gum (GG), beta-1,3/1,6-glucan (BG), chitosan (CHI), and alginate (ALG) at different concentrations and curing times. Biopolymer type, concentration, and curing time were used as model inputs, while UCS was the output. The adopted feed-forward ANN consisted of seven input neurons, one hidden layer with eight neurons, and one output neuron. Cross-validation was used to assess model generalization for the limited dataset. The model achieved an R^2 of 0.929, RMSE of 517 kPa, and MAE of 347 kPa. Garson's algorithm and the Cosine Amplitude Method were also applied to evaluate variable importance. Biopolymer type and curing time were identified as the most influential predictors, whereas concentration showed a comparatively lower contribution. The results demonstrate the potential of ANN-based modeling for interpreting and predicting the strength development of biopolymer-stabilized soils.

Keywords – biological soil stabilization; biopolymer; alginate; xanthan gum; guar gum; beta-glucan; artificial neural network; unconfined compressive strength.

I. INTRODUCTION

Soil improvement is an integral component of geotechnical engineering when naturally occurring soils do not satisfy the strength, stiffness, hydraulic, or deformation requirements imposed by infrastructure systems. Conventional stabilization techniques commonly rely on mechanical modification or chemical binders such as cement and lime. Although these approaches can provide substantial and relatively rapid improvement in engineering performance, their extensive use is accompanied by environmental concerns associated with energy consumption, greenhouse-gas emissions, resource depletion, and potential changes in soil and groundwater chemistry [1]–[4]. These concerns have stimulated growing interest in low-carbon and environmentally compatible ground-improvement technologies. Recent literature consequently identifies biological, bio-mediated, and bio-inspired soil treatment as an increasingly important direction for sustainable geotechnical engineering [1]–[6].

Biological soil stabilization encompasses processes in which biological activity or biologically derived materials are exploited to modify the engineering behavior of soils. Among the most extensively investigated bio-mediated techniques is microbially induced calcium carbonate precipitation (MICP), in which microbial activity promotes carbonate precipitation and generates cementing bonds between soil particles [1], [2], [5]. Related mechanisms include bio-cementation and bio-clogging, which can increase soil strength and reduce hydraulic conductivity by bonding particles and partially filling pore spaces. Despite these advantages, the practical implementation of MICP may be affected by several limitations, including treatment heterogeneity, difficulty in delivering bacterial and cementation solutions uniformly,

sensitivity to soil pore size, process-control requirements, and undesirable chemical by-products associated with some ureolytic pathways [4], [5], [7]. These limitations have encouraged the investigation of biologically derived stabilizers that can be incorporated into soil without requiring sustained in-situ microbial cultivation.

Biopolymers represent one of the most promising alternatives within this broader biological and bio-inspired stabilization framework. They are polymeric materials derived from renewable biological sources such as plants, microorganisms, algae, and animals and include polysaccharides, proteins, and other naturally derived macromolecules [3], [6]. Unlike MICP, direct biopolymer treatment can be implemented without cultivating microorganisms within the soil, which facilitates its application to a wider range of soil gradations, including fine-containing soils. Depending on their molecular structure and interaction with pore water, biopolymers may form viscous hydrogels, coat particle surfaces, bridge neighboring grains, fill pore spaces, and develop hydrogen or electrostatic bonds with mineral surfaces. These processes increase apparent interparticle cohesion and can substantially alter the compressive, tensile, shear, hydraulic, and deformation behavior of treated soils [4], [6]. Recent reviews have consequently highlighted applications ranging from erosion control, pavement and subgrade improvement, and slope stabilization to hydraulic barriers, grouting, and liquefaction mitigation.

Xanthan gum (XG), guar gum (GG), beta-glucan (BG), sodium alginate (SA), chitosan, gellan gum, agar, and lignin are among the biopolymers investigated for geotechnical applications. XG, a microbial polysaccharide produced through fermentation by Xanthomonas campestris, has attracted particular attention because of its high viscosity and ability to establish a stable gel network between soil particles [3], [6]. GG is a plant-derived galactomannan capable of increasing water retention and interparticle bonding, whereas BG is characterized by hydroxyl-rich molecular structures that favor hydrogen bonding and pore-filling interactions [6]. Sodium alginate has also demonstrated considerable potential because of its gel-forming and calcium-binding capacity and has recently been investigated in combination with MICP to improve both mechanical response and treatment uniformity [7].

The influence of curing is similarly complex. Water loss during curing promotes the formation and hardening of biopolymer networks, thereby strengthening interparticle bonds; however, the resulting response varies among polymers and soils [4], [6]. Studies on aeolian sand, silty soils, sand–bentonite mixtures, expansive soils, and granite residual soils have reported significant improvements in compressive or shear strength following biopolymer treatment [8]–[13]. At the same time, long-term behavior may be affected by biodegradation, wetting–drying cycles, prolonged saturation, and polymer hydrophilicity. Thus, both the magnitude and persistence of strength enhancement depend on interacting treatment and environmental variables. For example, Ai et al. [10] demonstrated significant mechanical enhancement of aeolian sand treated with GG and identified distinct optimum conditions for mechanical performance. Liu and Wang [11] showed that XG and GG modified the shear strength and water stability of granite residual soil through interparticle bridging and film formation, while excessive biopolymer dosage promoted localized gel accumulation and reduced the effectiveness of particle interlocking. Pitso et al. [12] further demonstrated that GG treatment influences not only shear resistance but also compressibility and hydraulic conductivity of sandy silt. These findings collectively emphasize that the response of biopolymer-treated soils is inherently nonlinear and cannot always be represented adequately by simple linear or single-parameter empirical correlations.

This nonlinear behavior has created a natural opportunity for data-driven modeling. Artificial neural networks (ANNs) are particularly attractive for geotechnical applications because they can approximate complex relationships between multiple input variables and engineering responses without requiring an a priori mathematical formulation. ANN-based models have previously been applied to the prediction of UCS, California Bearing Ratio, swelling behavior, and other nonlinear geotechnical responses [15]–[17]. More recently, PaeinAfrakoti and Ghazavi [14] compiled 248 UCS tests on XG-stabilized soils and demonstrated the feasibility of machine-learning-based strength prediction using soil properties, XG content, moisture conditions, and curing variables as predictors. Their study illustrates the growing role of data-driven approaches in biopolymer-based ground improvement.

Nevertheless, the experimental studies considered in the present literature survey also indicate that the simultaneous effects of biopolymer identity, concentration, and curing duration warrant further

quantitative examination. The experimental database reported by Soldo et al. [4] provides a suitable basis for data-driven modeling because all observations originate from the same base soil and laboratory investigation. In the present study, all five investigated biopolymers, XG, GG, BG, CHI, and ALG, were considered. Although the original experimental design was not fully balanced for CHI and ALG, using observations from a single experimental program minimizes the inter-study variability associated with differences in soil type, specimen geometry, preparation procedure, and testing protocol. Only experimentally reported UCS observations were retained, and missing combinations were not imputed. Accordingly, the objective of the present study is to develop an ANN-based predictive model for the unconfined compressive strength of biopolymer-stabilized soil using biopolymer type, biopolymer concentration, and curing time as explanatory variables.

II. Materials and Method

A. Source Experimental Database

The experimental data used in this study were obtained from the laboratory investigation reported by Soldo et al. [4]. No new laboratory experiments were performed in the present study; instead, the published experimental results were reorganized into a structured database for ANN modeling. This distinction is maintained throughout the study to ensure clear attribution between the original experimental work and the data-driven modeling conducted herein. The base material employed in the source investigation was residual Piedmont soil. The soil had coefficients of uniformity and curvature of approximately 10 and 1.5, respectively. The fine fraction exhibited a liquid limit of 49%, a plastic limit of 29%, and a plasticity index of 20%. According to the Unified Soil Classification System (USCS), the overall soil was classified as SW–SM (well-graded sand with silt).

Five biopolymers were originally investigated by Soldo et al. [4]: xanthan gum (XG), guar gum (GG), beta-1,3/1,6-glucan (BG), chitosan (CHI), and alginate (ALG). With the exception of ALG, the biopolymers were dry-mixed with the base soil, and water corresponding to 16.5% of the dry soil mass was subsequently added. This water content corresponded to the optimum water content selected for specimen preparation. The principal biopolymer concentrations were 1%, 2%, and 4% by dry mass of soil. ALG required a different preparation procedure involving calcium chloride and a lower water content, and the 4% ALG condition was omitted because of workability problems. For the unconfined compression tests, cylindrical specimens with a diameter of approximately 33 mm and a height of 71 mm were prepared and compacted in five layers. The specimens were tested after curing periods of 1 h, 5 days, and 30 days. The unconfined compression test was performed in accordance with ASTM D2166/D2166M using an axial strain rate of 1.5%/min, and compressive strength was calculated from the maximum axial load and the corresponding specimen cross-sectional area [4], [18]. For the present ANN analysis, all five biopolymers investigated in the source study, XG, GG, BG, CHI, and ALG, were retained. The original experimental program was unbalanced because CHI was not investigated at 30 days, while ALG was not tested at 4% concentration and was prepared using a distinct $CaCl_2$-assisted procedure. Therefore, only experimentally available treatment combinations were included in the predictive database, and no missing UCS observations were estimated or imputed. A total of 38 numerical UCS observations were available for ANN modeling. The 1 h–2% XG condition was experimentally investigated in the source study; however, its exact numerical UCS value was not explicitly labelled in the published figure and was therefore excluded from the predictive database rather than estimated graphically. Untreated plain-soil observations were retained only as reference values and were not used for ANN training.

Table 1. Variables used in the ANN database

Variable	Symbol	Values / Range	ANN role
Biopolymer type	B_T	XG, GG, BG, CHI, ALG	Categorical input
Biopolymer concentration	B_C	1, 2, 4%	Numerical input
Curing time	t_c	1 h, 5 days, 30 days	Numerical input
Unconfined compressive strength	q_u	92–7087 kPa	Output

The predictor–response structure summarized in Table 1 was constructed from the unconfined compression test results reported by Soldo et al. [4]. The selected database captures three principal treatment variables governing the measured response: biopolymer type, biopolymer concentration, and curing time. Considerable variation exists within the experimental response, with UCS values ranging from 92 to 7087 kPa across the available biopolymer-treated specimens. This wide range reflects the combined influence of polymer chemistry, dosage, and curing duration and provides a suitable nonlinear response space for ANN-based modeling. The detailed experimental observations used to construct the database are presented in Table 2. Unlike a fully balanced factorial design, the source experimental program does not contain every possible combination of biopolymer type, concentration, and curing period. Complete concentration–curing combinations are available for BG and GG, whereas CHI was not investigated at 30 days and ALG was not tested at a concentration of 4%. In addition, ALG specimens were prepared using a distinct $CaCl_2$-assisted procedure with a lower initial water content than the other biopolymer-treated specimens [4]. These experimentally imposed differences were retained in the database rather than corrected or artificially completed. Only experimentally available numerical UCS observations were therefore used for ANN development, and no missing values were generated through interpolation or imputation. The 1 h–2% XG condition was experimentally investigated in the original study, but its exact numerical UCS value was not explicitly reported in the published figure; consequently, this observation was excluded from model training. This procedure resulted in 38 usable biopolymer-treated observations. The untreated plain-soil values shown in Table 2 were retained solely as reference data for comparison with the treated specimens and were not included as ANN training observations.

Table 2. Experimental UCS data for biopolymer-treated and untreated reference soil reported by Soldo et al. [4]

Curing time	Material	0%	1%	2%	4%
1 h	Plain soil	181	-	-	-
	CHI	-	219	192	243
	ALG	-	116	92	-
	BG	-	153	177	180
	GG	-	174	190	207
	XG	-	148	-	162
5 days	Plain soil	1197	-	-	-
	CHI	-	1,261	1,401	1,059
	ALG	-	1,784	1,678	-
	BG	-	1,625	2,064	2,513
	GG	-	2,819	2,933	3,130
	XG	-	3,348	5,989	6,185
30 days	Plain soil	1169	-	-	-
	CHI	-	-	-	-
	ALG	-	1,946	2,044	-
	BG	-	1,705	2,047	2,467
	GG	-	2,944	2,999	3,277
	XG	-	3,258	5,939	7,087

Note: All compressive-strength values are in kPa. XG = xanthan gum; GG = guar gum; BG = beta-1,3/1,6-glucan; CHI = chitosan; ALG = alginate. ALG was tested only at 1% and 2%. CHI was available only for 1 h and 5 days. Missing treatment combinations were not imputed.

B. Artificial Neural Network Model

The ANN architecture shown in Figure 1 consists of an input layer, one hidden layer containing eight neurons, and a single-neuron output layer. The input layer represents the experimental predictors used for UCS prediction, including the biopolymer type encoded as categorical variables, biopolymer concentration (BC), and curing duration (t_c). The input variables are fully connected to the eight neurons in the hidden layer, where weighted combinations of the inputs are processed through the selected activation function. The resulting hidden-layer outputs are then transmitted to the single output neuron, which provides the predicted unconfined compressive strength, q_u. This architecture enables the ANN to

capture nonlinear relationships between biopolymer characteristics, concentration, curing duration, and UCS. Biopolymer type was represented using one-hot encoding, resulting in five binary variables (XG, GG, BG, CHI, and ALG). Together with biopolymer concentration and curing time, seven predictor variables were supplied to the ANN [19].

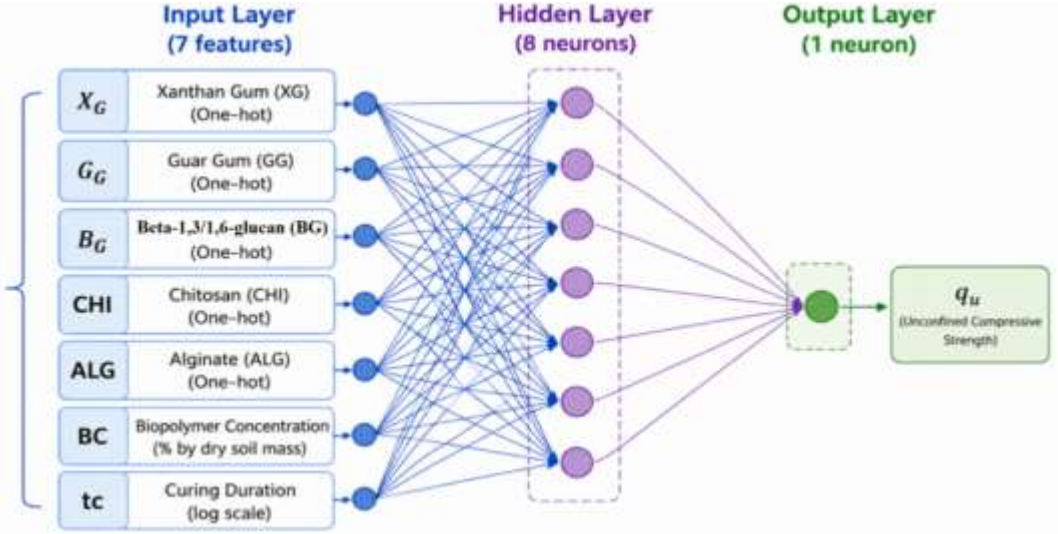


Fig. 1 Architecture of the ANN model for predicting UCS

C. *Model Validation and Performance Criteria*

Because of the limited number of observations, a conventional single random division into training, validation, and testing subsets was not adopted. Instead, *k*-fold cross-validation was employed to maximize the use of the available data and obtain a more reliable estimate of model generalization performance. In each fold, the model was trained on the majority of the observations and evaluated on the remaining subset, with the process repeated until every observation had been used for validation. Model performance was assessed using R^2, RMSE, and MAE, which collectively measure the goodness of fit, overall prediction error, and average absolute prediction error, respectively. This validation strategy reduces the dependence of the results on a single random data split and is particularly appropriate for small experimental datasets.

Logarithmic transformation of variables is performed to reduce skewness and improve the distributional characteristics of the variables; logarithmic transformation was applied to UCS and curing time. The base-10 logarithmic transformations were defined as $t_{c,log} = \log_{10}(t_c)$ and $q_{u,log} = \log_{10}(q_u)$, where t_c represents curing time and q_u denotes unconfined compressive strength. These transformed variables were subsequently used consistently in the statistical correlation and sensitivity analyses. Throughout the manuscript, $t_{c,log}$ and $q_{u,log}$ are used to denote the logarithmically transformed curing time and UCS, respectively.

III. Criteria for Model Performance Assessment

A. *Statistical Evaluation of Dataset Properties*

Table 3 summarizes the distribution of the main experimental variables. The mean UCS (q_u) was 1993.55, with a relatively high standard deviation of 1878.86, indicating substantial variability in strength. The raw UCS exhibited positive skewness (1.084), while logarithmic transformation reduced the skewness to −0.428, indicating a more balanced distribution. Curing time also showed considerable variability, ranging from 1 to 720 h, with a mean of 252.97 h. The BC values ranged from 1% to 4%, with a mean of 2.21%.

Table 3. Statistical Insights into the Experimental Dataset

Variable	Mean	Median	SD	Skewness	Kurtosis	Min	Max
BC (%)	2.211	2.000	1.234	0.555	−1.288	1.000	4.000
t_c (hours)	252.974	120.000	306.337	0.838	−1.172	1.000	720.000
t_c (log hours)	3.668	4.787	2.779	−0.448	−1.526	0.000	6.579
q_u (kPa)	1993.553	1744.500	1878.864	1.084	0.646	92.000	7087.000
log(q_u)	6.923	7.464	1.383	−0.428	−1.386	4.522	8.866

Table 4 evaluates whether BT significantly affects UCS. The one-way ANOVA indicated a statistically significant difference among the biopolymer groups (F = 4.82, p = 0.00356). However, the Kruskal–Wallis test produced H = 8.65 and p = 0.07049, which is not significant at the 5% level. Therefore, considering the non-normal distribution of UCS and the relatively small sample size, the effect of biopolymer type should be interpreted with caution. The parametric ANOVA indicated statistically significant differences, whereas the nonparametric Kruskal–Wallis test did not provide sufficient evidence at $\alpha = 0.05$.

Table 4. Effect of Biopolymer Type on UCS

Statistical test	Statistic	*p-value*	Decision at $\alpha = 0.05$
One-way ANOVA	F = 4.82	0.00356	Significant
Kruskal–Wallis	H = 8.65	0.07049	Not significant

Curing time was logarithmically transformed because of its wide range and positively skewed distribution. The transformation reduced the influence of extreme curing-time values and provided a more suitable representation of its relationship with UCS. The resulting $t_{c,log}$ exhibited a strong positive correlation with $q_{u,log}$ (r = 0.934).

Furthermore, Figure 2 provides a comprehensive statistical overview of the biopolymer UCS dataset and helps identify the main characteristics and relationships among the variables. Figure 2(a) shows a positively skewed distribution of UCS, with a mean of 1994 kPa and a median of 1748 kPa, indicating considerable variability and the presence of high-strength observations. Figure 2(b) illustrates differences in UCS among the biopolymer types, with XG showing the highest median strength and overall strength range. Figure 2(c) presents the relationship between biopolymer content (BC) and UCS, colored according to curing time, and indicates that increasing BC alone does not produce a consistent increase in UCS. Figure 2(d) demonstrates the substantial influence of curing time, with UCS increasing markedly from 1 h to 5 and 30 days. Figure 2(e) shows the Q–Q plot of log-transformed UCS, indicating that logarithmic transformation improves the distributional characteristics of the data. Finally, Figure 2(f) presents the correlation heatmap, showing a very strong positive correlation between $t_{c,log}$ and $q_{u,log}$ (r = 0.934), while BC exhibits only a weak correlation with $q_{u,log}$ (r = 0.079). Overall, the figure highlights the importance of curing time and biopolymer type in UCS development and provides an important basis for subsequent ANN modeling and sensitivity analysis.

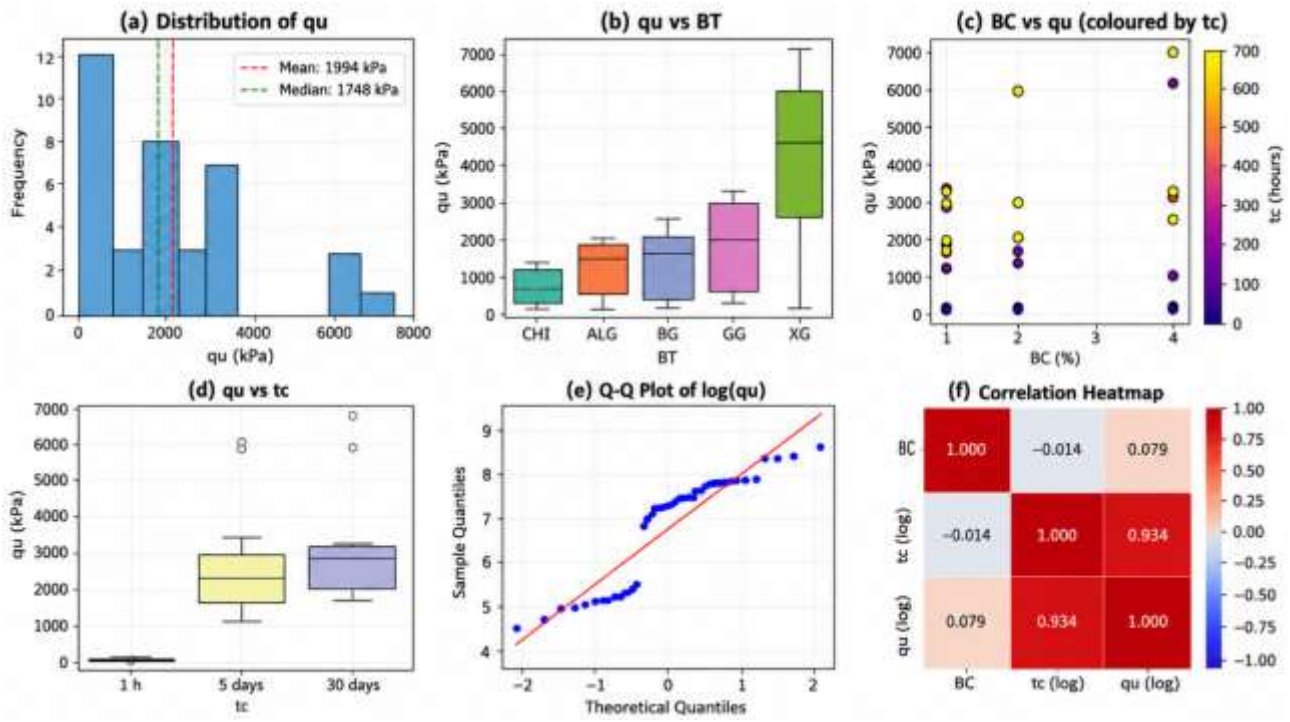


Fig. 2 Statistical distribution analysis of the biopolymer UCS dataset: (a) histogram of UCS values, (b) boxplots by biopolymer type and curing time, (c) scatter plot of UCS vs. biopolymer content, (d) Q-Q plot, (e) Q-Q plot showing log-transformed UCS distribution, and (f) correlation heatmap of input variables

IV. RESULTS

A. *ANN Configuration and Predictive Performance*

As outlined earlier, the main objective of this study is to develop a suitably configured ANN model for the accurate prediction of soil UCS. Table 5 summarizes the principal configuration parameters and settings adopted for the developed ANN model.

Table 5. ANN architecture and training hyperparameters

Algorithm	Hyperparameter Value	Description / Effect
Artificial Neural Network (ANN)		
Connection alpha	0.1	Transparency of connections
Learning Rate	0.001	Controls how quickly the model learns
Node radius	0.3	Radius of neuron circles
Input neurons	7	Represent the seven predictor variables used by the ANN.
Hidden layers	1	A single hidden layer captures the nonlinear relationships in the data.
Hidden neurons	8	Eight neurons provide sufficient capacity to learn complex patterns.
Output neurons	1	Produces the predicted UCS (q_u).
Hidden-layer activation function	Sigmoid	Introduces nonlinearity, allowing the ANN to model complex relationships between the inputs and UCS.

Figure 3 illustrates the predictive performance of the developed ANN model. In Figure 3(a), the predicted UCS values show close agreement with the measured values, with most points distributed near the perfect-fit line. Figure 3(b) presents the model performance metrics, yielding an R2 of 0.929, RMSE of 517, and MAE of 347, indicating strong predictive accuracy. Figure 3(c) shows that the residuals are mainly concentrated around zero, suggesting limited systematic bias and satisfactory model reliability. Overall, Figure 3 confirms the ANN's capability to capture the nonlinear relationship between the input variables and UCS.

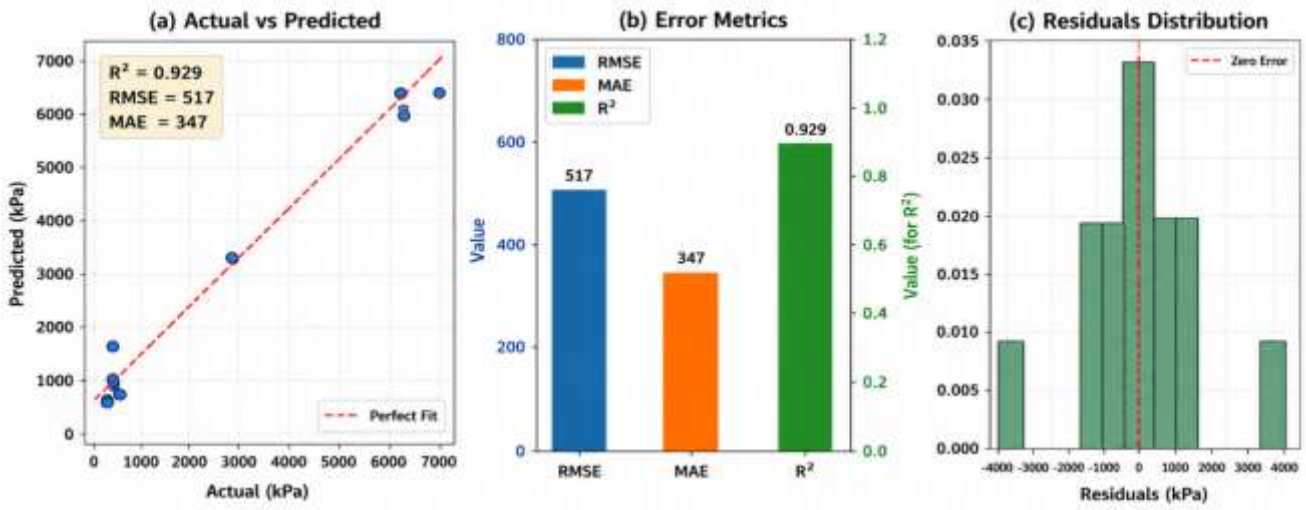


Fig. 3 Performance evaluation of the ANN model for UCS prediction: (*a*) actual versus predicted UCS values, (*b*) error metrics, and (*c*) residual distribution

Table 6 presents the actual and predicted UCS values for the best-performing cross-validation fold. The results show generally close agreement between the actual and predicted values, with relatively small absolute errors for most samples. The errors range from 1 to 180 kPa, indicating that the ANN model provides satisfactory prediction accuracy across different UCS levels. Notably, the selected predictions show good agreement between measured and predicted UCS values across both relatively low- and high-strength observations.

Table 6. Detailed Predictions for the Best-Performing Cross-Validation Fold

Sample	Actual qu (kPa)	Predicted q_u (kPa)	Absolute Error (kPa)
1	3348	3231	117
2	2944	2788	156
3	243	149	94
4	162	342	180
5	180	234	54
6	153	143	10
7	2999	2875	124
8	116	115	1

B. *Sensitivity Assessment of Input Variables*

A sensitivity analysis was performed using the Cosine Amplitude Method (CAM) to evaluate the relative influence of each input variable on the model output. CAM measures the degree of similarity between input and output variables based on the cosine of the angle between their corresponding data vectors [20].

$$\text{CAM}\,(X,\,Y) = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2} * \sqrt{\sum_{i=1}^{n} y_i^2}} \tag{1}$$

Where: x_i, y_i = data values of X and Y, Numerator = dot product of the two vectors, Denominator = product of their magnitudes.

Figure 4 presents the relative importance of the treatment/input variables determined using the CAM and Garson's algorithm. The Garson and CAM analyses produced different feature-importance rankings because they rely on different attribution mechanisms. Garson's algorithm estimates feature contributions based on the connection weights of the ANN, whereas CAM evaluates the degree of similarity between the input and output variables based on their corresponding data vectors. Accordingly, Garson identified biopolymer type (BT) as the dominant predictor (64.9%), whereas CAM assigned the highest sensitivity to curing time ($t_{c,log}$, 56.0%). These differences should not be interpreted as contradictory results but

rather as complementary perspectives on the learned ANN relationships. Importantly, both approaches identified BT and $t_{c,log}$ as the two principal predictors, while BC consistently exhibited a comparatively lower contribution. For grouped interpretation, the relative importance values of the five one-hot encoded biopolymer variables were summed and reported as the combined contribution of biopolymer type (BT).

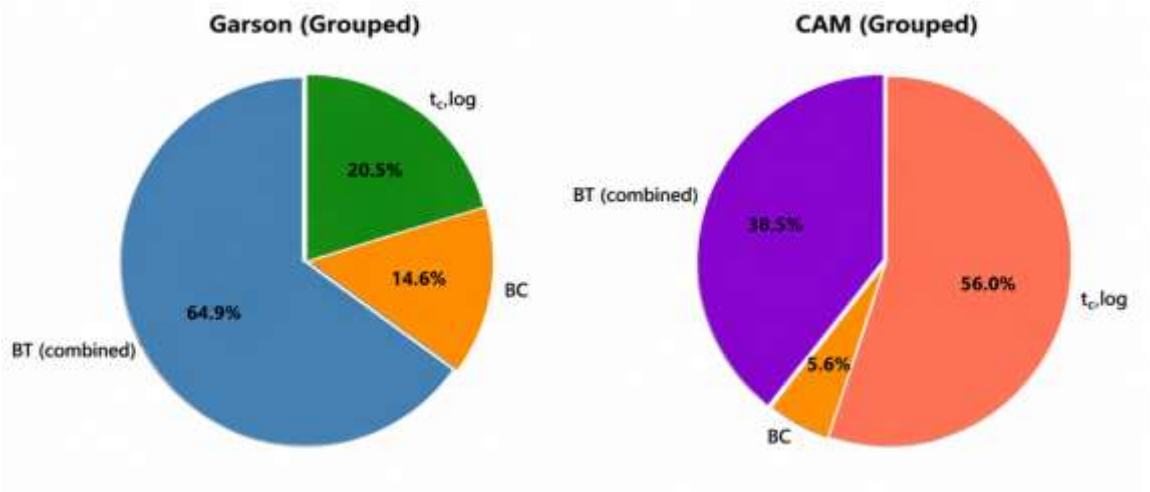


Fig. 4 Comparison of grouped feature importance obtained using Garson's algorithm and CAM analysis for the ANN model

V. DISCUSSION

The results demonstrate that the developed ANN model can effectively capture the nonlinear relationship between biopolymer treatment conditions and the UCS of treated soil. The dataset exhibits substantial variability, with UCS values ranging from 92 to 7087 kPa, reflecting the combined effects of biopolymer type, concentration, and curing duration. Such variability represents a challenging prediction problem; therefore, the ability of the ANN to achieve a high coefficient of determination (R^2=0.929) with RMSE and MAE values of 517 and 347, respectively, indicates a strong predictive capability.

The actual-versus-predicted results further confirm the reliability of the ANN. Most observations are located close to the perfect-fit line, indicating good agreement between measured and predicted UCS values. In addition, the residuals are distributed predominantly around zero, suggesting that the model does not exhibit substantial systematic overprediction or underprediction. This is particularly important for the present dataset because UCS contains both low-strength and high-strength observations. The detailed predictions in Table 6 support this observation, with absolute errors ranging from only 1 to 180 for the reported samples. These results indicate that the ANN is capable of reproducing the response across a relatively broad UCS range rather than being accurate only for the dominant middle portion of the dataset.

The statistical analysis provides further insight into the factors governing UCS. Curing time exhibits the strongest relationship with UCS, with a Pearson correlation coefficient of (r = 0.934) between log-transformed curing time and log-transformed UCS. In contrast, the direct correlation between biopolymer concentration and UCS is very weak (r =0.079). This finding suggests that strength development is strongly associated with curing duration, while the influence of concentration cannot be adequately represented through a simple linear relationship. From a practical perspective, this highlights the importance of allowing sufficient curing time for biopolymer-treated soils to develop their strength and also supports the use of nonlinear modeling approaches such as ANN.

The influence of biopolymer type is also evident from the experimental data, although its statistical significance should be interpreted cautiously. One-way ANOVA indicated significant differences among biopolymer groups (p = 0.00356), whereas the Kruskal–Wallis test did not reach the 5% significance level (p = 0.07049). The difference between these tests is important because the UCS data are non-normally distributed and the available sample size is relatively small. Thus, the results should not be interpreted as definitive evidence that biopolymer type alone determines UCS. Instead, they indicate that

polymer type likely contributes substantially to the response and should be retained as an input variable in the predictive framework.

The sensitivity analysis strengthens this interpretation. Garson's algorithm identified biopolymer type as the most influential predictor, accounting for approximately 64.9% of the estimated contribution, whereas CAM assigned the greatest sensitivity to curing time, with ($\mathbf{t}_{c,log}$) contributing approximately 56.0%. Both approaches consistently identified biopolymer type and curing time as the two most important predictors, while biopolymer concentration had a comparatively lower contribution. The difference between the two rankings is not necessarily contradictory because the methods quantify importance from different perspectives. Rather, their agreement on the dominant role of biopolymer type and curing time provides greater confidence in the interpretation of the ANN model.

An important strength of the present framework is that the model was developed using the experimentally available observations without artificially generating missing treatment combinations. The original database is unbalanced, with several combinations of polymer type, concentration, and curing duration unavailable experimentally. Retaining the original experimental structure avoids introducing potentially unrealistic synthetic observations. Furthermore, because only 38 usable treated-soil observations were available, k-fold cross-validation was adopted rather than relying exclusively on a single random train–test division. This approach improves the utilization of the limited dataset and provides a more robust assessment of generalization performance.

Nevertheless, the results should be interpreted within the limitations of the database. The relatively small sample size, unbalanced experimental design, and missing treatment combinations restrict the extent to which the ANN can be generalized to biopolymer types, concentrations, and curing conditions outside the investigated experimental domain. In addition, the strong association between curing time and UCS should not be interpreted as evidence of causality based solely on correlation analysis. Future studies should therefore incorporate larger and more systematically designed experimental datasets, additional soil and environmental variables, and independent external datasets for validation.

Overall, the findings demonstrate that the ANN provides a promising data-driven approach for UCS prediction of biopolymer-treated soils. The combination of satisfactory predictive accuracy, small prediction errors, residual analysis, statistical evaluation, and sensitivity assessment provides a comprehensive evaluation of model performance. More importantly, the interpretability results indicate that BT and curing time are the principal factors governing the ANN predictions, while biopolymer concentration has a comparatively smaller direct contribution. The developed framework can therefore support preliminary assessment of biopolymer-treated soils and reduce the need for extensive UCS testing during early-stage experimental and engineering investigations.

VI. CONCLUSION

This study successfully developed a promising artificial neural network (ANN) model for predicting the unconfined compressive strength (UCS) of biopolymer-treated soils. The adopted ANN architecture, consisting of a single hidden layer with eight neurons and a sigmoid activation function, achieved strong predictive performance, with an R^2 of 0.929, RMSE of 517, and MAE of 347. The close agreement between actual and predicted UCS values over the experimental range of 92–7087 kPa demonstrates the ability of the ANN to capture the nonlinear relationships governing biopolymer-treated soils. The use of k-fold cross-validation was particularly appropriate for the limited dataset of 38 observations, providing a more reliable assessment of model generalization and reducing dependence on a single data split.

The sensitivity analyses based on Garson's algorithm and the CAM identified biopolymer type (BT) and curing time ($t_{c,log}$) as the most influential predictors, while BC showed a comparatively lower contribution. This finding is consistent with the statistical analysis, which revealed a very strong positive correlation between logarithmic curing time and UCS (r = 0.934), compared with the weak correlation between BC and UCS (r = 0.079). From an engineering perspective, the findings indicate that UCS improvement depends more strongly on the selection of an appropriate biopolymer type and sufficient curing duration than simply increasing the biopolymer concentration within the investigated 1–4% range. Although ANOVA indicated significant differences among biopolymer groups (p=0.00356), the Kruskal–

Wallis test was not significant at the 5% level (p=0.07049). Therefore, the effect of biopolymer type should be interpreted cautiously considering the non-normal distribution and limited sample size.

Despite the promising predictive performance, several limitations should be acknowledged. The small dataset (n = 38) and unbalanced experimental design, including the absence of certain treatment combinations such as chitosan at 30 days and alginate at 4%, limit the ability of the model to fully capture all possible interactions among the input variables. In addition, the distinct $CaCl_2$-assisted preparation procedure and lower initial water content used for alginate introduce experimental heterogeneity that may affect direct comparisons among biopolymer types. Furthermore, the database represents a single soil type under controlled laboratory conditions; therefore, extrapolation to different soil types, environmental conditions, and field-scale applications should be undertaken with caution.

Future research should focus on integrating metaheuristic optimization algorithms with ANN and deep learning models to achieve more systematic optimization of network architecture and hyperparameters. Algorithms such as Grey Wolf Optimizer (GWO), and Whale Optimization Algorithm (WOA) could be used to optimize the number of hidden neurons, learning rate, activation functions, and initial connection weights. Such hybrid approaches may improve predictive accuracy and robustness compared with manually selected ANN configurations. As larger experimental databases become available, deeper architectures such as Deep Neural Networks (DNNs) could also be investigated. In addition, advanced explainability methods such as SHAP could complement Garson and CAM analyses by providing more detailed insights into the individual and interactive effects of biopolymer type, concentration, and curing time. Ultimately, these developments could contribute to a more accurate, interpretable, and generalizable intelligent framework for sustainable biopolymer-based ground improvement.

REFERENCES

[1] J. T. DeJong, B. M. Mortensen, B. C. Martinez, and D. C. Nelson, "Bio-mediated soil improvement," Ecological Engineering, vol. 36, no. 2, pp. 197–210, 2010.

[2] M. Umar, K. A. Kassim, and K. T. P. Chiet, "Biological process of soil improvement in civil engineering: A review," Journal of Rock Mechanics and Geotechnical Engineering, vol. 8, pp. 767–774, 2016.

[3] I. Chang, J. Im, and G.-C. Cho, "Introduction of microbial biopolymers in soil treatment for future environmentally-friendly and sustainable geotechnical engineering," Sustainability, vol. 8, no. 3, Art. no. 251, 2016.

[4] A. Soldo, M. Miletić, and M. L. Auad, "Biopolymers as a sustainable solution for the enhancement of soil mechanical properties," Scientific Reports, vol. 10, Art. no. 267, 2020, doi: 10.1038/s41598-019-57135-x.

[5] S.-G. Choi, I. Chang, M. Lee, J.-H. Lee, J.-T. Han, and T.-H. Kwon, "Review on geotechnical engineering properties of sands treated by microbially induced calcium carbonate precipitation (MICP) and biopolymers," Construction and Building Materials, vol. 246, Art. no. 118415, 2020.

[6] J.-E. Ryou, C. Jang, S. Gang, and J. Jung, "A review of biopolymers in geotechnical engineering: Mechanical properties, dynamic behavior, applications, and mechanisms," Journal of Rock Mechanics and Geotechnical Engineering, vol. 18, pp. 4971–4997, 2026, doi: 10.1016/j.jrmge.2025.10.046.

[7] M. Şimşek, S. Çelik, and H. Akoğuz, "Comparative study of mechanical and microstructural properties of biocemented sandy soils enhanced with biopolymer: Evaluation of mixing and injection treatment methods," Applied Sciences, vol. 15, Art. no. 8090, 2025, doi: 10.3390/app15148090.

[8] A. Al-Jabobi and H. Bilsel, "The effects of xanthan gum and guar gum on the mechanical properties of sand–bentonite mixtures," Sustainability, vol. 17, Art. no. 5339, 2025, doi: 10.3390/su17125339.

[9] H. O. Kara and M. U. Yılmazoğlu, "Investigation of the effects of two different biopolymers on the strength parameters of silty soil," Gazi University Journal of Science Part C: Design and Technology, 2025, doi: 10.29109/gujsc.1628900.

[10] Z. Ai et al., "Improvement of aeolian sand using guar gum: Mechanical behavior, plant growth response, and microstructural evaluation," Geotechnical and Geological Engineering, vol. 43, Art. no. 375, 2025, doi: 10.1007/s10706-025-03337-2.

[11] Y. Liu and Z. Wang, "Hydro-mechanical performance and microstructural evolution of biopolymer-modified granite residual soil," Applied Sciences, vol. 16, Art. no. 4916, 2026, doi: 10.3390/app16104916.

[12] T. Pitso, L. M. Lalicata, D. Gallipoli, A. De Paoli, and A. W. Bruno, "Effect of hydrated lime and guar gum treatment on the geotechnical behaviour of a Ligurian sandy silt," Environmental Geotechnics, vol. 12, no. 4, pp. 316–326, 2025, doi: 10.1680/jenge.24.00081.

[13] N. Raj, S. Selvakumar, and K. Muthukkumaran, "Effect of xanthan gum biopolymer treatment on the strength enhancement of expansive soils," International Journal of Pavement Research and Technology, 2026, doi: 10.1007/s42947-026-00860-0.

[14] M. Taslimi PaeinAfrakoti and M. Ghazavi, “Unified model for prediction of the strength of soils stabilized with xanthan gum using white-box machine learning,” Scientific Reports, vol. 16, Art. no. 22136, 2026, doi: 10.1038/s41598-026-51877-1.
[15] H. Javdanian and S. Lee, “Evaluating unconfined compressive strength of cohesive soils stabilized with geopolymer: A computational intelligence approach,” Engineering with Computers, vol. 35, no. 1, pp. 191–199, 2019.
[16] A. K. Sabat, “Prediction of California bearing ratio of a soil stabilized with lime and quarry dust using artificial neural network,” Electronic Journal of Geotechnical Engineering, vol. 18, pp. 3261–3272, 2013.
[17] R. K. Dutta, A. Singh, and T. Gnananandarao, “Prediction of free swell index for the expansive soil using artificial neural networks,” Journal of Soft Computing in Civil Engineering, vol. 3, no. 1, pp. 47–62, 2019.
[18] ASTM D2166/D2166M-13, “Standard Test Method for Unconfined Compressive Strength of Cohesive Soil,” ASTM International, 2016.
[19] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” Nature, vol. 323, pp. 533–536, 1986.
[20] H. Hasanzadehshoili, A. Lakirouhani, and J. Medzvieckas, “Superiority of artificial neural networks over statistical methods in prediction of the optimal length of rock bolts,” Journal of Civil Engineering and Management, vol. 18, no. 5, pp. 655–661, 2012.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Gyroid ve Split-P TPMS Yapılarının Geometrik Tasarım Kriterlerinin Laminar Akış Karakteristikleri ve Sayısal Analiz Sonuçları Üzerindeki Etkilerinin İncelenmesi

İsmail Tur[1], Erkan Bahçe[2]

[1]*Malatya Turgut Özal Üniversitesi, Doğanşehir Vahap Küçük Meslek Yüksekokulu, Mülkiyet Koruma ve Güvenlik Bölümü, Sivil Savunma ve İtfaiyecilik Programı, Malatya, Türkiye.*
[2] *İnönü Üniversitesi Mühendislik Fakültesi Makine Mühendisliği Bölümü, Malatya, Türkiye.*
[*]*(ismail.tur@ozal.edu.tr)*

Özet – Üçlü Periyodik Minimal Yüzey (TPMS) geometrileri, periyodik, pürüzsüz ve kesintisiz iç kanal yapıları sayesinde akışkanlar mekaniği uygulamalarında önemli hidrodinamik avantajlar sunmaktadır. Bu çalışmada, TPMS yapılarının temel geometrik tasarım kriterlerinin iç akış karakteristikleri ve hesaplamalı akışkanlar dinamiği (CFD) performans metrikleri üzerindeki etkileri incelenmiştir.Bu kapsamda literatürde yaygın olarak kullanılan Gyroid ve Split-P topolojileri karşılaştırmalı olarak değerlendirilmiştir. Akışkan bölgeleri, CAD yazılımında implicit modelleme yöntemi kullanılarak, 5 × 5 × 5 mm birim hücre boyutunda ve %80 gözeneklilik oranında oluşturulmuştur. Sayısal akış analizleri (CFD), Poly-Hexcore ağ yapısı ile duvar yakınındaki viskoz etkileri hassas biçimde çözümleyen sınır tabaka (inflation layer) yaklaşımı kullanılarak gerçekleştirilmiştir. Simülasyonlar, akışkanın (havanın) hızı $v = 0.39,\ 0.78,\ 1.95$ ve 3.89 m/s değerlerinde değiştirilerek, daimi, sıkıştırılamaz ve laminar akış koşulları altında yürütülmüştür. Elde edilen sayısal sonuçlar doğrultusunda, geometrik parametrelerin akış direnci üzerindeki etkileri değerlendirilmiştir. Bunun yanı sıra, topolojik tasarımın akış homojenliği üzerindeki etkisini ortaya koymak amacıyla, yapı içerisindeki yerel hız gradyanları nicel olarak karşılaştırılmıştır.Elde edilen bulgular, TPMS yapılarında topoloji türü ve gözeneklilik gibi geometrik tasarım değişkenlerinin akış fiziğini mikroskobik ölçekte nasıl etkilediğini ortaya koymaktadır. Böylece çalışma, gelecekte geliştirilecek yüksek performanslı akışkan sistemlerinin tasarımı ve optimizasyonu için önemli bilimsel veriler sunmaktadır.

Anahtar Kelimeler – Eklemeli İmalat, Lattice Yapılar, TPMS, Hesaplamalı Akışkanlar Dinamiği,Akış Karakteristiği.

I. GİRİŞ

Eklemeli imalat teknolojileri, geleneksel üretim yöntemlerinin kısıtlamalarını aşarak, karmaşık geometrilerin ve yüksek performanslı bileşenlerin katman katman inşa edilmesine olanak tanıyan yenilikçi bir üretim modeli sunmaktadır. Bu ileri imalat yaklaşımının sağladığı tasarım özgürlüğü sayesinde, üstün mekanik özelliklere ve yüksek yüzey alanı/hacim oranına sahip karmaşık kafes (lattice) ve TPMS yapıları yüksek hassasiyetle üretilebilmektedir. Hafiflik, yüksek enerji sönümleme kapasitesi ve artırılmış ısı transfer yüzeyi gibi çok yönlü avantajlar sunan bu hücresel yapılar; havacılık, uzay, biyomedikal, otomotiv endüstrileri ve yeni nesil ısı değiştirici sistemlerinde giderek genişleyen bir kullanım alanı bulmaktadır.[1] Söz konusu karmaşık topolojilerin içinden geçen akışkan davranışının ve ısı transferi karakteristiklerinin anlaşılmasında Hesaplamalı Akışkanlar Dinamiği (CFD) analizlerinin gerçekleştirilmesi sistem optimizasyonu için kritik bir öneme sahip olup [2]; alandaki çalışmalar bu yapıların termal akış uygulamaları performansını artırma motivasyonuyla hız kazansa da, TPMS gibi yüzey alanı fazla olan kafes yapılardaki etkileri detaylı çalışılmamıştır. Literatürde bu alanda yapılan

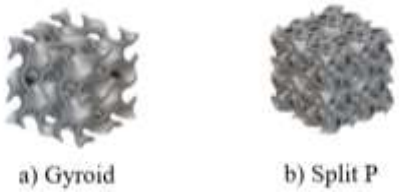
çalışmalar incelendiğinde eklemeli imalat teknolojilerindeki gelişmeler lattice ve TPMS yapılarının termal akış uygulamalarındaki kullanımını ivmelendirmiş ve bu alandaki Hesaplamalı Akışkanlar Dinamiği (CFD) analizlerinin temelini oluşturmuştur [3-5]. Bu dönemin ilk çalışmalarında araştırmacılar, geleneksel ısı değiştirici tasarımlarının yerini alabilecek hücresel yapıların akış direnci ve ısı transferi potansiyellerini anlamaya odaklanmıştır. Ji vd. (2025), farklı kafes yapılarının akış ve ısı transferi karakteristiklerini CFD yöntemleri ile inceleyerek bu yapıların geleneksel tasarımlara kıyasla çok daha yüksek bir yüzey alanı/hacim oranı sunduğunu ve bu durumun termal performansı önemli ölçüde artırdığını ortaya koymuştur [6]. Benzer şekilde polimer ve metal bazlı gyroid yapıların içinden geçen akışkanın basınç düşüş davranışının sayısal olarak modellendiği ve hücresel yapıların porozite oranlarının makroskopik akışkanlar dinamiği üzerindeki etkisinin araştırıldığı çalışmalar da literatürde mevcuttur [7]. Bu alanda olan ilk çalışmalar, TPMS ve lattice yapıların temel CFD modellerinin doğrulanması ve potansiyel mühendislik uygulamalarının belirlenmesi açısından kritik bir teorik zemin hazırlamıştır. Son yıllarda yapılan çalışmalar ise temel yapısal incelemelerden ziyade farklı sınır koşulları altında TPMS ve kafes yapıların hidrodinamik olarak optimize edilmesine ve eşlenik ısı transferi gibi daha karmaşık akış senaryolarının CFD ile modellenmesine yönelmiştir[8].Moradmand vd.(2024),Gyroid ve Schwarz yapılarının düşük ve yüksek Reynolds sayılarındaki termal-hidrolik performanslarını üç boyutlu CFD analizleriyle karşılaştırmış; Gyroid yapısının yüksek Reynolds sayılarında daha üstün ısı transfer katsayısı sergilediğini bildirmiştir[9]. Ayrıca Han vd. (2026), eklemeli imalat ile üretilen metalik TPMS yapılarında zorlanmış taşınım mekanizmalarını inceleyerek, akış yönü ile yapısal oryantasyon arasındaki ilişkinin yerel akışkan hızı dağılımları ve girdap oluşumu üzerindeki etkilerini sayısal olarak haritalandırmıştır [10]. Güncel literatür incelendiğinde, TPMS yapılarının yalnızca temel akış ve ısı transferi performanslarının değerlendirilmesiyle sınırlı kalınmadığı, aynı zamanda fonksiyonel derecelendirilmiş geometriler, üretim kaynaklı yüzey özellikleri ve çok amaçlı optimizasyon yaklaşımlarının da araştırmalara dahil edildiği görülmektedir. Özellikle gözeneklilik dağılımının akış doğrultusunda değiştirildiği fonksiyonel derecelendirilmiş TPMS yapılarının, basınç kaybı ile ısı transferi arasındaki dengeyi iyileştirmede önemli bir tasarım parametresi olduğu gösterilmiştir. Bu çalışmalar, geometrik parametrelerin yalnızca mekanik performansı değil, aynı zamanda akış alanındaki hız dağılımı ve sıcaklık gradyanlarını da doğrudan etkilediğini ortaya koymuştur [10].Bununla birlikte, TPMS tabanlı kompakt ısı değiştiriciler üzerinde gerçekleştirilen deneysel ve sayısal çalışmalar, Gyroid, Diamond ve IWP gibi farklı topolojilerin aynı çalışma koşullarında belirgin şekilde farklı hidrodinamik davranışlar sergilediğini göstermektedir. Özellikle yüksek özgül yüzey alanına sahip sürekli yüzey geometrilerinin, klasik yapılara göre daha yüksek taşınımla ısı transferi sağlarken, buna karşılık akış direncinde de artış meydana getirdiği rapor edilmiştir. Bu nedenle güncel araştırmalar, maksimum ısı transferi ile minimum basınç kaybı arasında optimum dengeyi sağlayacak TPMS topolojilerinin belirlenmesine odaklanmaktadır [11].Eklemeli imalat yöntemlerinin gelişmesiyle birlikte araştırmalar yalnızca ideal CAD modelleri üzerinden yürütülmemekte, üretim sürecinde oluşan geometrik sapmaların da akış performansına etkisi incelenmektedir. Özellikle Lazer Toz Yatağı Füzyonu (LPBF) yöntemiyle üretilen TPMS yapılarında yüzey pürüzlülüğü, duvar kalınlığı değişimleri ve üretim toleranslarının akış direncini doğrudan etkilediği belirlenmiştir. Deneysel karakterizasyon ve sayısal analizlerin birlikte kullanıldığı çalışmalar, gerçek üretim geometrilerinin ideal modellere göre daha yüksek basınç kaybı oluşturabileceğini göstermektedir [12].Diğer taraftan, CFD çözücülerinin optimizasyon algoritmalarıyla birlikte kullanılması TPMS tasarımında yeni bir araştırma alanı oluşturmuştur. Çok amaçlı optimizasyon teknikleri sayesinde gözeneklilik, birim hücre boyutu ve topoloji türü gibi tasarım değişkenleri aynı anda değerlendirilerek basınç kaybı, geçirgenlik ve ısı transfer katsayısı gibi performans kriterleri birlikte optimize edilmektedir. Böylece uygulamaya özgü yüksek performanslı TPMS geometrilerinin tasarlanması mümkün hale

gelmiştir [13].Son yıllarda yayımlanan kapsamlı derleme çalışmaları da TPMS yapılarında gelecekteki araştırmaların yalnızca mekanik veya termal performansa odaklanmaması gerektiğini, üretim süreci, çok ölçekli modelleme, akış karakterizasyonu ve veri tabanlı tasarım yöntemlerinin birlikte ele alınmasının önemini vurgulamaktadır. Özellikle yapay zekâ destekli tasarım optimizasyonu, fonksiyonel derecelendirilmiş yapılar ve çok disiplinli optimizasyon yaklaşımlarının TPMS araştırmalarının gelecek yönelimleri arasında yer aldığı belirtilmektedir [14].Mevcut literatür değerlendirildiğinde, Gyroid topolojisi üzerine gerçekleştirilen CFD çalışmalarının sayısı oldukça fazla olmasına rağmen Split-P topolojisinin hidrodinamik davranışını inceleyen çalışmaların sınırlı olduğu görülmektedir. Ayrıca mevcut araştırmaların önemli bir bölümü ısı transferi odaklı olup, yalnızca akış karakteristikleri, geçirgenlik parametreleri, tortuosity, hız dağılımları ve hidrodinamik ölü bölge oluşumlarını birlikte değerlendiren parametrik çalışmalar yeterli düzeyde değildir. Bu durum, farklı TPMS topolojilerinin aynı geometrik koşullar altında sistematik olarak karşılaştırılmasına yönelik yeni çalışmalara ihtiyaç olduğunu göstermektedir [15].Bu çalışmada, farklı iç mimari tasarım kriterlerine sahip Gyroid ve Split-P TPMS topolojilerinin, %80 gözeneklilik oranlarında oluşturulan akış kanalları içerisindeki laminar akış karakteristikleri Hesaplamalı akışkanlar dinamiği yazılımı kullanılarak parametrik olarak incelenmiştir. Elde edilen CFD sonuçları doğrultusunda yapısal değişimlerin tortuosity, geçirgenlik, basınç gradyanları ve lokal hız dağılımları üzerindeki etkileri ayrıntılı olarak değerlendirilmiştir. Böylece çalışma, farklı TPMS topolojilerinin hidrodinamik performanslarının karşılaştırmalı olarak ortaya konulmasına katkı sağlamayı amaçlamaktadır.

II. MATERYAL VE YÖNTEM

A. ***Lattice Yapılarının Tasarımı ve CFD Analizi***

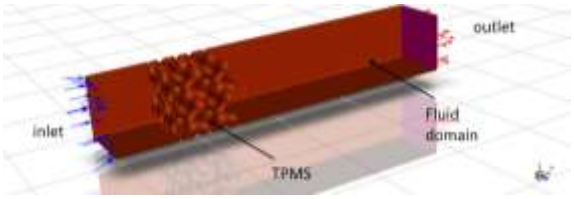
Bu çalışmada, Üçlü Periyodik Minimal Yüzey (TPMS) tabanlı hücresel yapıların iç akış karakteristikleri; geometrik modelleme ve Hesaplamalı Akışkanlar Dinamiği (CFD) analizlerini kapsayan bir metodoloji ile incelenmiştir. Geometrik modellerin oluşturulması aşamasında Gyroid ve Split-P topolojileri kullanılmış ve CAD ortamında tasarlanmıştır.Tüm modellerde birim hücre boyutu 5 × 5 × 5 mm olarak belirlenmiş; topoloji türü ve %80 gözeneklilik oranlarının akış fiziği üzerindeki etkilerini izole edebilmek amacıyla sistematik bir tasarım matrisi oluşturulmuştur.Analizlerde katı fazı temsil eden yapısal malzeme olarak Polilaktik Asit (PLA) referans alınmıştır. PLA, düşük yoğunluğu (yaklaşık 1,24 g/cm³), yüksek boyutsal kararlılığı ve karmaşık geometrilerin yüksek hassasiyetle modellenmesine imkan tanıyan pürüzsüz yüzey karakteristiği nedeniyle termofluidik araştırmalarda sıklıkla tercih edilen bir polimerdir[16]. Tasarlanan TPMS yapılara ait görsel Şekil.1'de gösterilmiştir.



Şekil.1 Tasarlanan TPMS yapılara ait CAD görselleri.

Sayısal akış analizlerinin gerçekleştirilmesi amacıyla, tasarlanan TPMS geometrilerinden akışkanın hareket edeceği iç hacimler çıkarılarak hesaplama bölgeleri tanımlanmıştır. Duvar yüzeylerine yakın bölgelerde viskoz etkileri ve hız gradyanlarını yüksek doğrulukla yakalamak amacıyla çok katmanlı sınır tabaka yaklaşımı uygulanmış; ağ bağımsızlığı çalışmalarıyla doğrulanan optimum çözümleme ağları tüm modeller için standart hale getirilmiştir.Sayısal simülasyonlar CFD yazılımında; hava fazında,

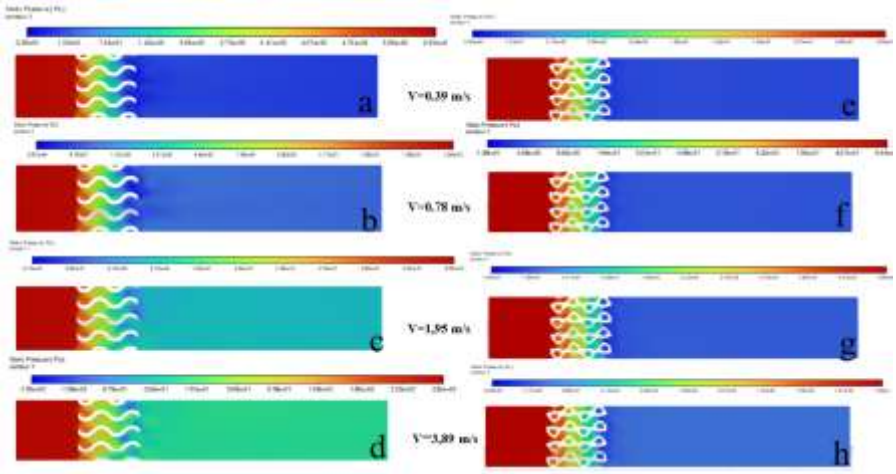
sıkıştırılamaz, daimi ve laminar akış koşulları altında yürütülmüştür. Akışkan olarak havanın termofiziksel özellikleri (yoğunluk ρ=1,225kg/m^3 =, dinamik viskozite μ=1,789×10^{-5}kg/ms) baz alınmıştır. Analizler, akışkanın (havanın) hızı $v = 0.39$, 0.78, 1.95 ve 3.89 m/s değerlerini kapsayacak şekilde giriş hızları uygulanarak gerçekleştirilmiştir. Simülasyon sürecinde hız girişi (Velocity Inlet) ve basınç çıkışı (Pressure Outlet) sınır şartları tanımlanmış, TPMS yüzeylerinde ise kaymama (No-slip) koşulu benimsenmiştir. Ayrıca, topolojik tasarımların iç akış homojenliği üzerindeki etkilerini belirlemek amacıyla, yapı içerisindeki lokal hız gradyanları ve hidrodinamik ölü bölge hacim yüzdeleri iso-surface yöntemi kullanılarak olarak karşılaştırılmıştır.Bu metodolojik yaklaşım, TPMS yapılarının tasarım değişkenlerinin hava akış fiziğini nasıl değiştirdiğinin bütünsel bir şekilde ortaya konulmasını sağlamıştır. CFD analizinde TPMS yapılara ait görsel Şekil.2'de gösterilmiştir.



Şekil.2 CFD Analizinde Kullanılan Örnek TPMS Yapısı.

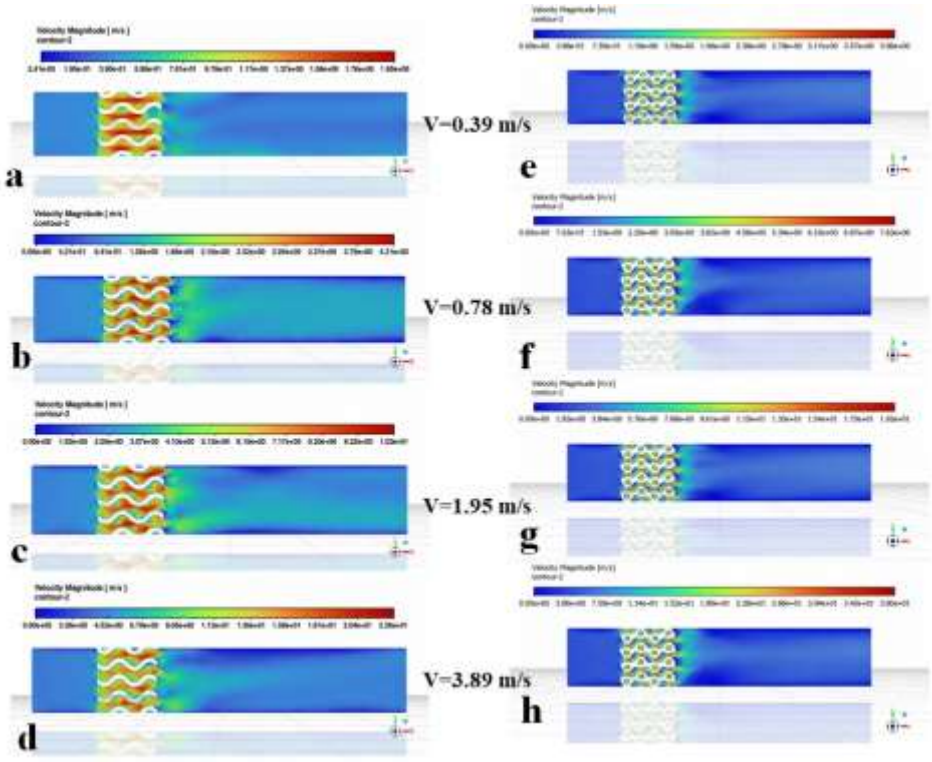
III. BULGULAR

Bu çalışmada, eklemeli imalat teknolojileri kullanılarak üretilme potansiyeline sahip, %80 gözenekliliğe (porozite) sahip Gyroid ve Split-P TPMS yapılarının hidrodinamik performansları CFD yazılımı kullanılarak sayısal olarak incelenmiştir. Akışkanın (havanın) hızı $v = 0.39$, 0.78, 1.95 ve 3.89 m/s değerlerinde değiştirilerek giriş ile çıkış arasındaki basınç düşüşü (ΔP) ve hız konturları analiz edilmiştir. Sürekli yüzey yapısına sahip TPMS geometrilerinde akış performansının yalnızca gözeneklilik oranına değil, aynı zamanda iç kanal bağlantılarının sürekliliğine ve kanal kıvrımlılığına (tortuosity) da bağlı olduğu literatürde yaygın olarak bildirilmektedir [17]. İlk olarak, her iki TPMS topolojisine ait geometrik modeller değerlendirilmiş ve CFD analizlerinde kullanılan akışkan bölgeleri oluşturulmuştur. Ağ yapısının oluşturulmasının ardından farklı Reynolds sayılarında gerçekleştirilen analizler sonucunda TPMS geometrileri boyunca oluşan basınç dağılımları incelenmiştir. Basınç konturları, girişten çıkışa doğru meydana gelen enerji kayıplarının görsel olarak değerlendirilmesine olanak sağlamaktadır. Sayısal analizlerden elde edilen verilere göre, giriş hızı ile birim uzunluk başına basınç düşüşü ($\Delta P/L$) arasında parabolik bir ilişki mevcuttur. Forchheimer-Darcy eşitliğine de uygun olarak, bu ilişki $y = ax^2 + bx$ formatında 2. dereceden polinom denklemleri ile yüksek doğrulukta ($R^2 > 0.995$) modellenmiştir. Analiz sonuçları karşılaştırıldığında, aynı giriş hızlarında ve aynı gözeneklilik oranında (%80) Split-P geometrisinin, Gyroid geometrisine kıyasla çok daha yüksek bir basınç düşüşüne neden olduğu tespit edilmiştir. Örneğin, en düşük hız olan 0.39 m/s'de Gyroid yapısında basınç düşüşü yalnızca 296.02 Pa/m iken, Split-P yapısında bu değer 1442.97 Pa/m seviyesindedir (Yaklaşık 4.8 kat). En yüksek giriş hızı olan 3.89 m/s'de ise Gyroid için basınç düşüşü 12704.21 Pa/m olarak hesaplanırken, Split-P yapısında 30344.30 Pa/m seviyesine (Yaklaşık 2.4 kat) ulaşmıştır. Bu durum, Gyroid yapısının sıfır ortalama eğrilik (zero mean curvature) özelliğinin hidrodinamik direnci minimize ettiğini, Split-P yapısının ise daha keskin/dirençli iç topolojisi nedeniyle akışa karşı çok daha yüksek bir sürtünme oluşturduğunu göstermektedir. Şekil-3 'de verilen statik basınç konturları da bu durumu desteklemektedir; yapı girişindeki yüksek basınç bölgesi, Gyroid yapısı boyunca görece daha kademeli bir düşüş sergilerken, Split-P yapısı çok daha yüksek bir yığma basıncına yol açmaktadır.



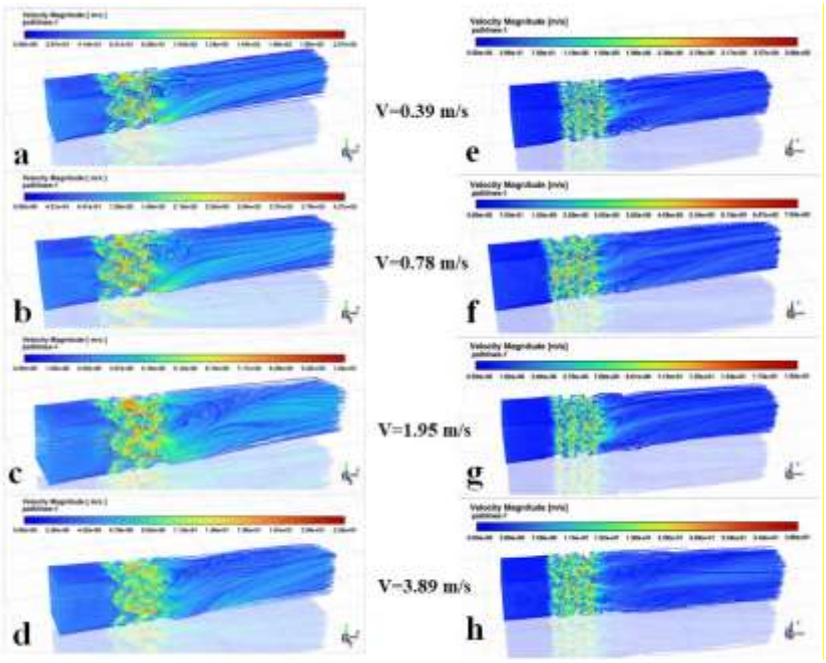
Şekil 3. Gyroid ve Split-P Topolojisine Ait Basınç Konturları.

Basınç dağılımlarının ardından hız alanları değerlendirilmiştir. TPMS yapılarının sürekli yüzey karakteri nedeniyle akışın kanal boyunca yön değiştirmesi, yerel hız büyüklüklerinin topolojiye bağlı olarak değişmesine neden olmaktadır. Bu nedenle hız konturlarının incelenmesi, akışın homojenliği ve yüksek hız bölgelerinin belirlenmesi açısından önem taşımaktadır. Hızın $v = 0.39$ m/s (a) ve $v = 0.78$ m/s (b) olduğu durumlara ait hız konturları ve akış çizgileri incelendiğinde (Şekil-3), Gyroid yapısının düzgün ve sürekli kavisli (smooth) yüzey topolojisinin akışın yapı içerisine homojen bir şekilde nüfuz etmesine olanak tanıdığı görülmektedir. Gyroid geometrisinde, katı yüzeyler etrafında akış ayrılması nispeten daha gecikmeli gerçekleşmekte ve akış kanalları içinde düzgün bir ivmelenme gözlemlenmektedir. Buna karşın, artan giriş hızlarında (örneğin c, $v = 1.95$ m/s), kesit daralmalarından kaynaklanan lokal hızlanmalar ve yapı ardında oluşan girdaplı (wake) bölgeler daha belirgin hale gelmektedir. Akış çizgilerinden (Pathlines) görüleceği üzere (Şekil.3), gözenekli yapının içindeki karmaşık 3B yollar, akışkanın yönünü sürekli değiştirerek yanal karışımı (lateral mixing) artırmaktadır. Bu durum, TPMS yapıların ısı değiştiricilerde (heat exchangers) veya karışım odalarında neden yüksek performans sergilediğini fiziksel olarak doğrulamaktadır.

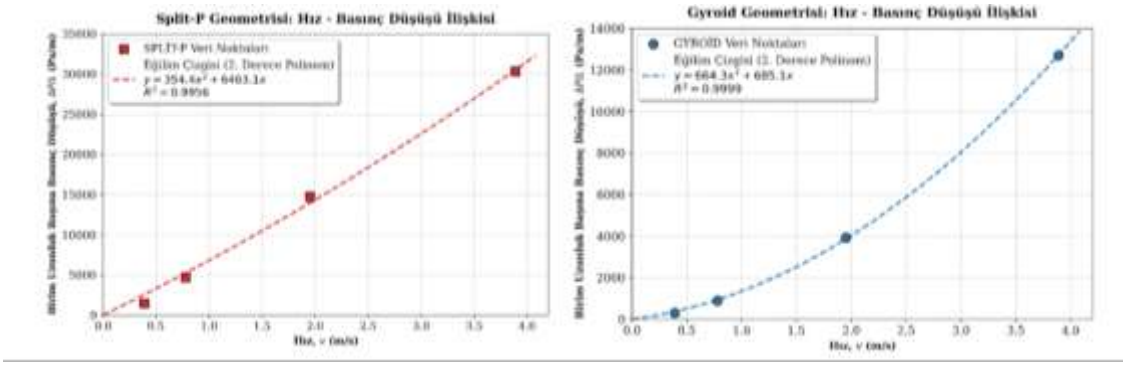


Şekil 4. Gyroid ve Split-P Topolojisine ait hız konturları.

Hız konturlarının ardından akış çizgileri incelenerek akış yollarının sürekliliği değerlendirilmiştir. Akış çizgileri, kanal geometrisinin akış yönünü nasıl şekillendirdiğini ve olası geri akış veya resirkülasyon bölgelerini görsel olarak ortaya koymaktadır.Akış alanlarının değerlendirilmesini takiben basınç kaybı verileri kullanılarak Darcy–Forchheimer modeli oluşturulmuştur. Böylece gözenekli yapının geçirgenlik (K) ve atalet katsayısı (C_2) hesaplanmış ve makroskobik akış direnci nicel olarak belirlenmiştir. Darcy–Forchheimer yaklaşımı, TPMS yapılarının eşdeğer gözenekli ortam olarak modellenmesinde yaygın biçimde kullanılan bir yöntemdir [18].



Şekil 5. Gyroid ve Split-P Topolojisindeki Akış Çizgileri.



Şekil 6. Gyroid ve Split-P geometrilerinde akış hızı (v) ile birim uzunluk başına basınç düşüşü ($\Delta P/L$)

Elde edilen tüm hidrodinamik parametreler birlikte değerlendirildiğinde, TPMS geometrisinin ve gözeneklilik oranının iç akış karakteristikleri üzerinde belirleyici olduğu görülmektedir. Basınç kaybı, hız dağılımı, geçirgenlik ve akış homojenliği gibi performans kriterlerinin birlikte değerlendirilmesi, farklı TPMS topolojilerinin mühendislik uygulamalarına yönelik tasarımında kapsamlı bir karşılaştırma yapılmasına olanak sağlamaktadır. Bu nedenle çalışma kapsamında elde edilen bulguların, özellikle gözenekli ortam akışları, kompakt ısı değiştiricileri ve hafif akış bileşenlerinin tasarımına yönelik gelecekteki optimizasyon çalışmalarına katkı sağlayacağı düşünülmektedir.

IV. SONUÇLAR

Bu çalışmada elde edilen sayısal sonuçlar şu şekilde özetlenebilir:

- ✓ Akış Rejimi ve Basınç Kaybı İlişkisi: Hem Gyroid hem de Split-P TPMS yapılarında, akış hızı ile basınç düşüşü arasındaki ilişki Forchheimer kanununa uygun şekilde 2. derece polinom davranışı ($y = ax^2 + bx$) sergilemiştir. Modelin uyum başarısı Gyroid için $R^2 = 0.9999$, Split-P için ise $R^2 = 0.9956$ olarak belirlenmiştir.Topolojinin Hidrodinamik Dirence Etkisi: Aynı %80 gözeneklilik (porozite) oranına sahip olmalarına rağmen, geometrik yapı akış direncini önemli ölçüde değiştirmektedir. Tüm hız değerlerinde Split-P geometrisi, Gyroid geometrisinden belirgin şekilde daha yüksek basınç kaybı yaratmıştır.
- ✓ Tasarım ve Optimizasyon Önerisi: Enerji verimliliğinin (düşük pompalama gücünün) kritik olduğu mühendislik uygulamalarında (örneğin filtreler, biyomedikal implantlar veya düşük basınçlı ısı değiştiriciler), pürüzsüz kavisli yüzeyi ile daha düşük form sürtünmesi sunan Gyroid geometrisinin tercih edilmesi daha uygun bulunmuştur. Buna karşın, basınç kaybından ziyade aşırı yüksek türbülans, yoğun ısı transferi veya karışım oranının arzu edildiği özel durumlarda Split-P topolojisi bir alternatif olarak değerlendirilebilir.CFD Analizinin Etkinliği: Kontur ve akış çizgisi sonuçları, TPMS yapılarının 3B karmaşık yapısının ard izi (wake), akış ayrılması ve anafor (vortex) oluşumu gibi etkenleri nasıl tetiklediğini açıkça ortaya koymuştur.

KAYNAKLAR

[1]. Yeo, S. J., Oh, M. J., & Yoo, P. J. (2019). Structurally controlled cellular architectures for high-performance ultra-lightweight materials. *Advanced Materials*, *31*(34), 1803670.

[2]. Sarkar, S. (2024). Computational Thermo-Fluid Dynamics Modeling For Process Optimization In Hydrogen-Integrated Industrial Heat Systems. *Journal of Sustainable Development and Policy*, *3*(03), 87-133.

[3]. Al-Ketan, O., Ali, M., Khalil, M., Rowshan, R., Khan, K. A., & Abu Al-Rub, R. K. (2021). Forced convection computational fluid dynamics analysis of architected and three-dimensional printable heat sinks based on triply periodic minimal surfaces. *Journal of Thermal Science and Engineering Applications*, *13*(2), 021010.

[4]. Thumthae, C., Aye, N. N., Reamkrathok, M., Sukjit, E., & Hemsuwan, W. (2026). Case study on a compact and effective additively manufactured octahedral lattice heat exchanger: experimental and CFD validation. *Applied Thermal Engineering*, 130959.

[5]. Ji, H., Liu, Z., Zhao, J., & Wang, B. (2025). Enhanced cooling performance in cutting tools using TPMS-integrated toolholders: A CFD-based thermal-fluidic study. *Modelling*, *6*(3), 73.

[6]. Ji, H., Liu, Z., Zhao, J., & Wang, B. (2025). Enhanced cooling performance in cutting tools using TPMS-integrated toolholders: A CFD-based thermal-fluidic study. *Modelling*, *6*(3), 73.

[7]. Sun, Q., Zhi, G., Zhou, S., Dong, X., Shen, Q., Tao, R., & Qi, J. (2024). Advanced design and manufacturing approaches for structures with enhanced thermal management performance: a review. *Advanced materials technologies*, *9*(15), 2400263.

[8]. Reynolds, B. W. (2020). Simulation of flow and heat transfer in 3D printable triply periodic minimal surface heat exchangers. *University of Canterbury: Christchurch, New Zealand*.

[9]. Moradmand, M. M., & Sohankar, A. (2024). Numerical and experimental investigations on the thermal-hydraulic performance of heat exchangers with Schwarz-P and gyroid structures. *International Journal of Thermal Sciences*, *197*, 108748.

[10]. Han, G., Zhang, Q., Xiong, W., Ye, Z., Wei, L., Xiao, C., ... & Liu, F. (2026). Decoupling Thermo-Fluidic Trade-Offs in High-Reynolds-Number Flows with a Damage-Tolerant, Starfish-Inspired Dual-Channel Microlattice via Metal Additive Manufacturing. *3D Printing and Additive Manufacturing*, 23297662261443751.

[11]. Tang, W., Zhou, H., Zeng, Y., Yan, M., Jiang, C., Yang, P., ... & Zhao, Y. (2023). Analysis on the convective heat transfer process and performance evaluation of Triply Periodic Minimal Surface (TPMS) based on Diamond, Gyroid and Iwp. *International Journal of Heat and Mass Transfer*, *201*, 123642.

[12]. Chouhan, G., & Bidare, P. (2025). Manufacturability of A20X printed lattice heat sinks. *Progress in Additive Manufacturing*, *10*(8), 5541-5556.

[13]. Anacreonte, V. A., Iasiello, M., Mauro, G. M., Andreozzi, A., Bianco, N., & Chiu, W. K. (2024). A THERMAL ANALYSIS OF A FUNCTIONALLY-GRADED GYROID AS A HEAT SINK. In *ICHMT DIGITAL LIBRARY ONLINE*. Begel House Inc..

[14]. Li, J., Wang, L., He, X., Liang, J., & Dong, C. (2025). Additively manufactured metallic TPMS lattice structures: design strategies, fabrication, multifunctional properties, and applications. *npj Advanced Manufacturing*, *2*(1), 45.

[15]. Reynolds, B. W., Fee, C. J., Morison, K. R., & Holland, D. J. (2023). Characterisation of heat transfer within 3D printed TPMS heat exchangers. *International Journal of Heat and Mass Transfer*, *212*, 124264.

[16]. Ali, S. (2025). Sustainable Bioplastic Blends for 3D Printing: Formulation, Process Optimization, and Application Development.

[17]. Della Torre, A., Montenegro, G., Tabor, G. R., & Wears, M. L. (2014). CFD characterization of flow regimes inside open cell foam substrates. *International Journal of Heat and Fluid Flow*, *50*, 72-82.

[18]. Piedra, S., Gómez-Ortega, A., & Pérez-Barrera, J. (2023). Prediction of flow properties of porous triply periodic minimal surface (TPMS) structures. *Fluids*, *8*(12), 312.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Predicting Urban Traffic Congestion from Pre-Event Contextual Features: A Comparative Evaluation of Classical Machine Learning Algorithms

Cafer Yazicioglu [1*], Ammar Aslan [2]

[1] *Department of Construction, Vocational School of Technical Sciences, Batman University, Batman, Turkey*

[2] *Vocational School of Technical Sciences, Batman University, Batman, Turkey*

Corresponding Author: Cafer Yazicioglu, Phone:+905345725423, e-mail:cafer.yazicioglu@batman.edu.tr

Abstract – Predicting traffic congestion plays a critical role for proactive traffic management and driving planning in terms of Intelligent Transportation Systems (ITS). This study aims to classify traffic congestion levels via machine learning methods by using multi-source smart city database. The problem is formulated to forecast the congestion through previously known contextual attributes (hour, day, holiday, weather, and location). Real-time traffic measurements (speed, occupancy, number of vehicles) were excluded from the model to prevent data leakage. The target variable contains two classes: low (71%) and medium (29%) congestion. Logistic Regression, Naïve Bayes, Decision Tree, Random Forest, Gradient Boosting and XGBoost algorithms are compared via accuracy, precision, sensitivity, F1, and ROC-AUC metrics The Random Forest model showed the highest performance in terms of macro-F1, which was determined as the primary metric (Accuracy = 0.875; Macro-F1 = 0.841). In terms of threshold-independent discrimination capability, Gradient Boosting achieved the highest ROC-AUC value (0.950). Ensemble and tree-based models were found to be significantly superior to linear and probabilistic base models; the strongest predictor of congestion was found to be the time of day (peak time pattern).

Keywords – Traffic congestion, classification, machine learning, smart city, Random Forest

I. INTRODUCTION

Traffic prediction is a core function of Intelligent Transportation Systems (ITS), providing the basis for adaptive signal control and proactive congestion management [1]. Traffic state can be expressed either as a continuous variable, such as flow or speed, or as a discrete nominal class. The nominal representation is easier for travelers to interpret, since it can be displayed on traffic maps with color codes, and it supports departure-time and route choices as well as the operational decisions of transportation agencies [2]. These studies significantly differ regarding the categorization of traffic statuses. For example, Arabiat et al. developed a dual classification problem (normal and congested) for an urban arterial in Amman, Jordan [3]. Other researchers expanded this approach to three-class structures and classified the traffic in rural roads as light, semi-heavy, and heavy [2]. To represent urban traffic in more detail, some studies adopted four-class or five-class systems. Cheng et al. divided urban expressway traffic into five levels ranging from fluent traffic to severe congestion [4], while Zafar and Ul Haq used a similar five-level congestion scale based on Estimated Time of Arrival (ETA), ranging from smooth traffic to blockage [5]. Increasing the number of classes provides a more detailed representation of traffic conditions but also makes the classification task more challenging.

The most common approach to solve these two- or multi-class classification problems is to comparatively evaluate traditional machine learning algorithms such as Random Forest (RF) [6], Gradient Boosting (GB)

[7], Extreme Gradient Boosting (XGBoost) [8], Decision Tree (DT), Support Vector Machine (SVM) [9], K-Nearest Neighbors (KNN) [10], Logistic Regression (LR), and Naive Bayes (NB) [5]**.** Comparative studies in recent years report that tree-based ensemble methods often achieve the strongest performance in traffic congestion classification [5], [11], [12]**.** For example, Zafar and Ul Haq (2020) showed that Random Forest achieved the highest accuracy (92%) in ETA-based traffic congestion prediction, outperforming XGBoost (91%), KNN (91%), Gradient Boosting (83%), Logistic Regression (50%), and SVM (39%). [5]. Similarly, Abu Sara et al. (2026) reported that Random Forest achieved 96.07% accuracy, surpassing Decision Tree (95.67%) and SVM (91.47%) in multi-class traffic congestion classification on a dataset combining cyber-physical and social network data. [12].

Similar findings have been reported in traffic flow prediction studies. Lu et al. showed that the proposed BO-XGBoost model outperformed the Bi-LSTM model [13], while Tsalikidis et al. reported that ensemble tree-based methods, particularly LGBM, achieved higher predictive performance than LSTM- and GRU-based models on their traffic prediction datasets [14]. These findings suggest that ensemble methods remain a competitive and practical choice for structured traffic datasets, particularly when the prediction task is based on tabular contextual variables.

The classification success of these algorithms is directly linked to input variable selection and feature engineering. Traditional macroscopic traffic variables such as speed, traffic volume, and occupancy rate continue to form the basis of traffic congestion classification models [15]. The literature also shows that temporal variables and derived attributes play an important role in model performance [16]. In a Random Forest travel time prediction model, the travel time observed 15 minutes earlier and the time of day (TOD) ranked as the two most influential predictors, with relative importance values of 34.85% and 30.12%, respectively [17]. To better represent the complex nature of urban mobility, researchers also include derived and exogenous features in addition to standard traffic variables. Cheng et al. proposed the ample degree, defined as the ratio of residual flow to maximum flow, as an additional traffic state classification indicator alongside traffic flow, speed, and occupancy [4]. Similarly, Rasaizadi et al. applied a Genetic Algorithm (GA) to reduce 32 candidate features to 25 for hourly traffic state classification, eliminating the three road-blockage variables together with four holiday-related variables [2].

Other studies have reported that multi-source feature sets, where weather data are combined with Google Maps ETA information [5], can improve classification performance. However, integrating multiple data sources generally increases data collection and preprocessing complexity while requiring more sophisticated data fusion strategies. Therefore, recent studies have shown growing interest in approaches that achieve high classification performance using a smaller number of easily accessible traffic variables.

Despite these developments, there are significant methodological differences in the literature in terms of dataset size, validation strategies, and performance assessments. The datasets used vary considerably, ranging from local studies of approximately 5,000 observations to large-scale datasets encompassing over one million observations. Validation approaches also differ across studies, ranging from hold-out training-test splits to stratified data partitioning and cross-validation. Model performance is often assessed through standard classification metrics such as accuracy, precision, recall, and F1-score [4], [11], [12], [18]. However, Bokaba et al. demonstrated that assessments based solely on statistical accuracy may be insufficient by applying a cost-sensitive misclassification matrix that considers the operational impacts of prediction errors [11].

These methodological differences reveal several research gaps in the literature. Firstly, many studies rely on real-time traffic measurements or multi-source data integration to achieve high classification performance [5], [18]. Moreover, many existing approaches use traffic variables measured simultaneously with the traffic state, making them more suitable for traffic state estimation than for prediction and increasing the risk of information leakage.. In addition, some studies depend on infrastructure-intensive traffic measurements collected through fixed roadside monitoring systems [4], [18]. While these approaches achieve successful results, they require more extensive data acquisition and integration than models based on a small number of easily obtainable traffic variables.

Additionally, the majority of current studies are conducted in a single city, at a single intersection, or on a single highway, and the generalizability of the models across different urban settings has not been extensively investigated. Therefore, there is a need for approaches that evaluate generalizability across

different cities while providing reliable classification performance using easily accessible contextual variables. Considering the accuracy, interpretability, and computational efficiency of classical machine learning algorithms on structured datasets, evaluating these algorithms on multi-city datasets can make a significant contribution to the literatüre [19].

This study investigates traffic congestion classification using a multi-source smart city dataset. Unlike existing studies, the problem is formulated as a classification task that predicts traffic congestion using contextual attributes (time, calendar, weather, and location) that are available before congestion occurs. Accordingly, real-time traffic variables directly related to congestion (speed, occupancy, and vehicle count) were excluded from the model to prevent information leakage. In this context, a context-based, leakage-controlled classification framework was developed by integrating multiple data sources through shared keys, and the performance of six classical machine learning algorithms was compared using multiple evaluation metrics.

II. MATERIAL AND METHOD

A. Dataset

This study uses Smart City Traffic Flow Prediction database openly shared on Kaggle [20]. The dataset consists of multiple interrelated files, such as traffic_sensor_data.csv (containing main sensor measurements, 218,400 observations), weather_conditions.csv (containing weather logs), and sensor_locations.json (containing sensor location/hardware information). The target variable is the categorical variable congestion_level, which specifies the level of congestion for each observation. In the dataset, this variable contains two classes: low congestion (Low, 155,409 observations; 71.2%) and medium congestion (Medium, 62,991 observations; 28.8%). Therefore, the problem is a binary classification problem with an unbalanced class distribution.

B. Data Integration and Pre-Processing

Information from different files has been added to the main table using a left join method via common key columns. Weather records are combined via timestamp, and sensor information via sensor ID. Temporal attributes such as hour, day of the week, month, and weekend indicator are derived from the timestamp; sine and cosine transformations (hour_sin, hour_cos) are applied to preserve the cyclical nature of the clock [21].

The problem in this study is formulated as a forecasting scenario. Therefore, instantaneous traffic variables (number and sub-counts of vehicles, average speed, and occupancy rate) that are measured simultaneously with the tag and cannot be known before the event occurs have been removed from the model. This step both prevents data leakage and forces the model to predict congestion only from contextual information [22] that can be known in advance (time, calendar, weather, and location). In addition, sensor hardware/telemetry metadata (manufacturer, model, software version, communication protocol, etc.) and ID columns, which lack predictive power, have been eliminated.

For numerical features, median-based padding and standardization; and for categorical features, most frequent padding and one-hot encoding were implemented within a pipeline structure. To address class imbalance, balanced class weights were used in supporting models, while balanced sample weights were used in Gradient Boosting and XGBoost models.

C. Used Classification Model

To classify the traffic congestion, supervised learning algorithms that are commonly used in the literature are compared. Logistic Regression, Naive Bayes, Decision Tree, Random Forest, Gradient Boosting, and XGBoost. Ensemble methods like Random Forest and Gradient Boosting/XGBoost stand out due to their ability to capture nonlinear and complex relationships by combining numerous decision trees. Support Vector Machine (SVM) and K-Nearest Neighbor (KNN) algorithms, which have high computational costs, were not included in this comparison due to the size of the dataset (218,400 observations).

D. Assessment Metrics

Considering the unbalanced class distribution, the models were evaluated not only for accuracy but also for macro-averaged precision, recall, and F1 score, as well as weighted F1 and area under the ROC curve (ROC-AUC) metrics. Accuracy represents how many of the predicted positives are actually positive; sensitivity represents how many true positives are captured, and the F1 score represents the harmonic mean of the two. The macro mean reflects the performance of the minority class by giving equal weight to each class; therefore, it is more informative than accuracy in unbalanced datasets [23]. The macro-F1 score was adopted as the primary criterion in model selection. The rationale for this is that the operational objective of the study is to produce a discrete decision regarding the level of congestion, and the macro mean gives equal weight to the minority class as to the majority class. ROC-AUC, being a threshold-independent ranking criterion, was reported as supplementary information and was not considered a determining factor in model ranking.

E. Experimental Setup

The dataset was stratified to 80% training and 20% testing (174,720 training, 43,680 testing; random_state = 42). All studies were performed in Python using PyCharm with scikit-learn, xgboost, pandas, and numpy libraries; visualizations were generated using matplotlib and seaborn at 300 DPI resolution.

III. RESULTS

A. Exploratory Data Analysis

The distribution of the target variable is shown in Figure 1. The dataset is clearly imbalanced, with approximately 71% of the observations belonging to the low congestion class and the remaining 29% to the medium congestion class. Because of this imbalance, macro-averaged evaluation metrics and class weighting were used during model training to ensure that the minority class received appropriate consideration.

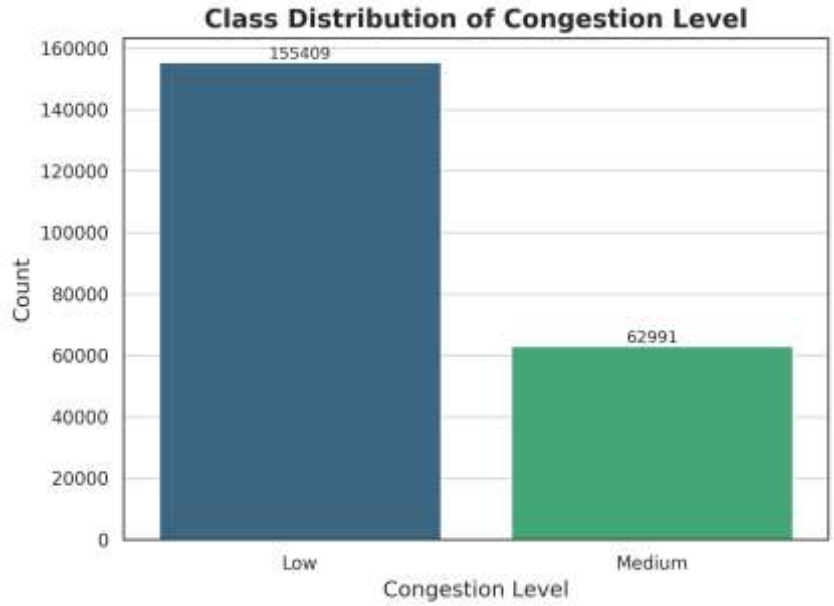


Fig 1. Distribution of congestion level classes.

Figure 2 illustrates how congestion levels vary across the days of the week. Moderate congestion is more common on weekdays, accounting for roughly 33% of observations, whereas its frequency drops noticeably on Saturdays and Sundays. As a result, low congestion becomes the dominant class during weekends. This pattern suggests that calendar-related variables, such as the day of the week, provide useful information for predicting traffic congestion.

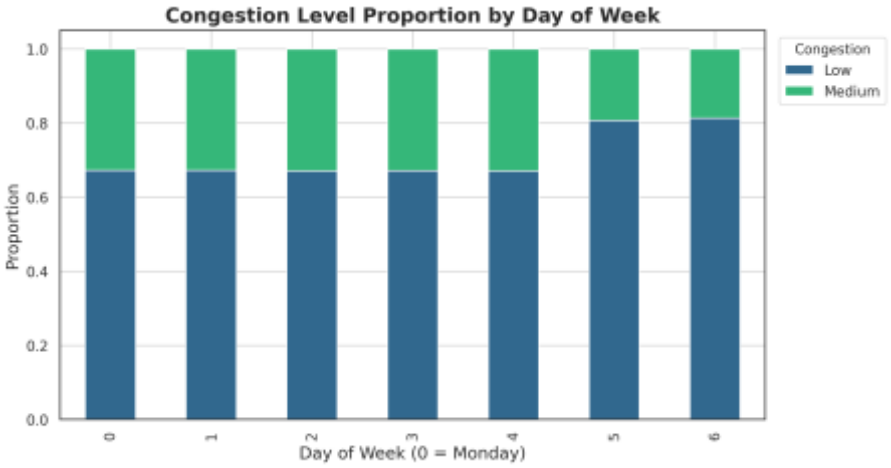


Fig 2. Proportions of congestion levels by day of the week.

The correlation matrix of the numerical contextual features is presented in Figure 3. Overall, the results indicate that there are no strong linear relationships between the weather variables and the calendar-related features. The highest correlations are found among time-based variables that are naturally related, such as hour, hour_sin, and hour_cos. These findings suggest that the contextual features, when considered individually, do not exhibit strong linear associations with the target variable, reinforcing the need for models capable of capturing more complex, nonlinear relationships.

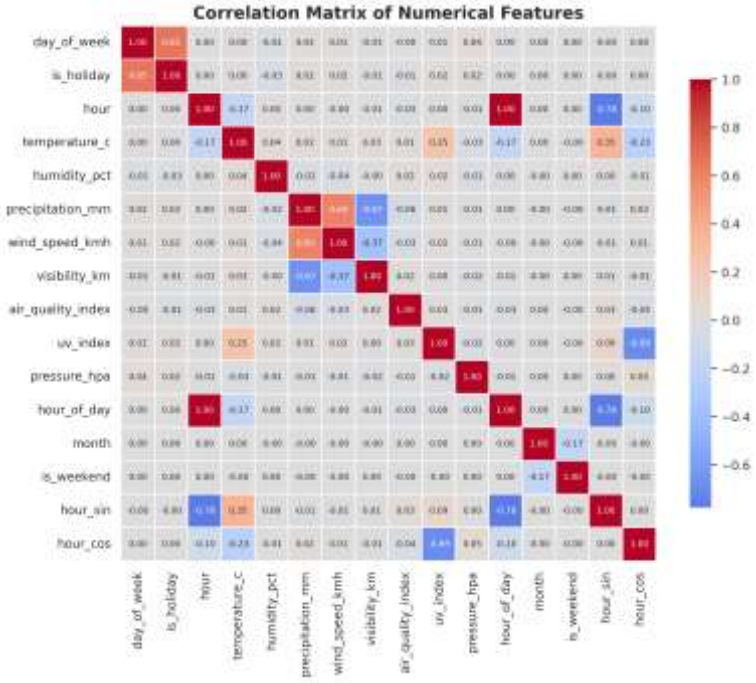


Fig 3. Correlation matrix of numerical features.

B. Model Comparison Results

Six classification algorithms are trained and evaluated based on leakage-controlled contextual attribute dataset and the results are summarized in Table 1. The highest performance for the primary measurement macro-F1 is observed in Random Forest (Macro-F1 = 0.841; Accuracy = 0.875). It was followed by Gradient Boosting (0.832), Decision Tree (0.826), and XGBoost (0.820). The ranking is different in terms of ROC-AUC where the highest value is Gradient Boosting (0.950) followed by XGBoost (0.948), Random Forest and Decision Tree (0.945). The fact that the ROC-AUC values of the four tree-based and ensemble-based models clustered in a narrow range of 0.945–0.950 indicates that these models are practically equivalent in their ability to rank classes; the difference between them stems primarily from their classification performance at the decision threshold. However, Logistic Regression (0.726) and Naïve

Bayes (0.707) show significantly lower performances. The performance of these two linear/probabilistic models is close to the baseline (approximately 71%), always choosing the majority class. However, tree-based models significantly exceed this baseline. This difference shows that the relationship between contextual attributes and congestion level is not linear and ensemble methods can effectively capture this relationship.

Table 1. Performance comparison of classification models on the test set.

Model	Accuracy	Macro-F1	Weighted F1	ROC-AUC
Random Forest	0.875	0.841	0.872	0.945
Gradient Boosting	0.855	0.832	0.858	0.950
Decision Tree	0.848	0.826	0.852	0.945
XGBoost	0.842	0.820	0.847	0.948
Logistic Regression	0.726	0.708	0.739	0.769
Naive Bayes	0.707	0.671	0.717	0.782

The accuracy and macro-F1 comparison of the models is visualized in Figure 4. The superiority of four tree and ensemble-based models over two linear/probabilistic models can be clearly seen.

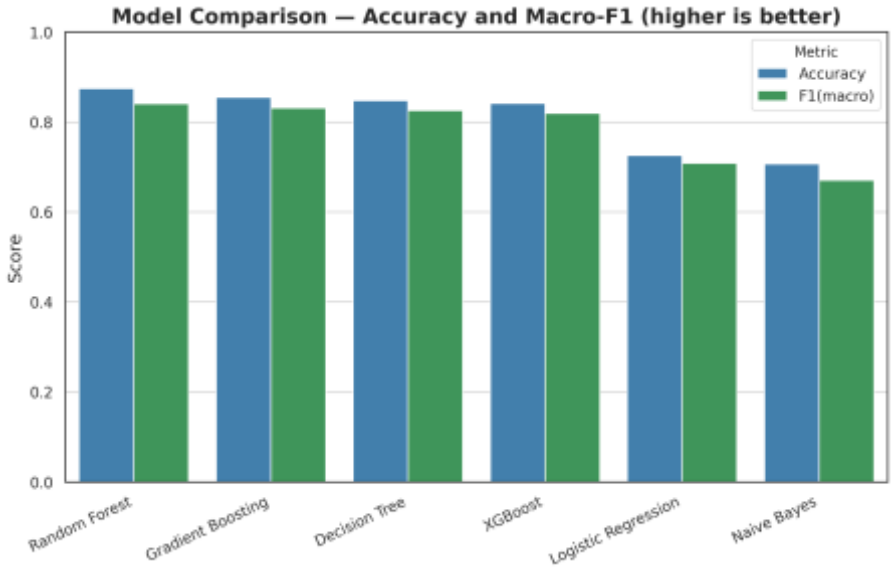


Fig 4. Comparison of model accuracy and macro-F1 score.

The confusion matrix of Random Forest, the best model in terms of Macro-F1, is presented in Figure 5, and its row-normalized version is shown in Figure 6. The model distinguishes the low congestion class with high accuracy (94% sensitivity); and achieves 71% sensitivity in the minority class of medium congestion. Thanks to class weighting, the overlooking of the minority class is largely prevented.

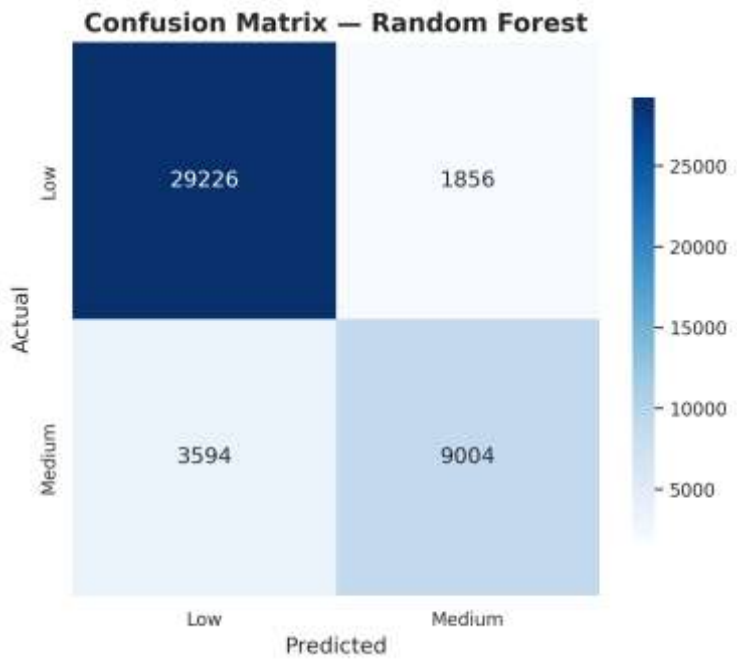


Fig 5. Confusion matrix for the Random Forest model (number of observations).

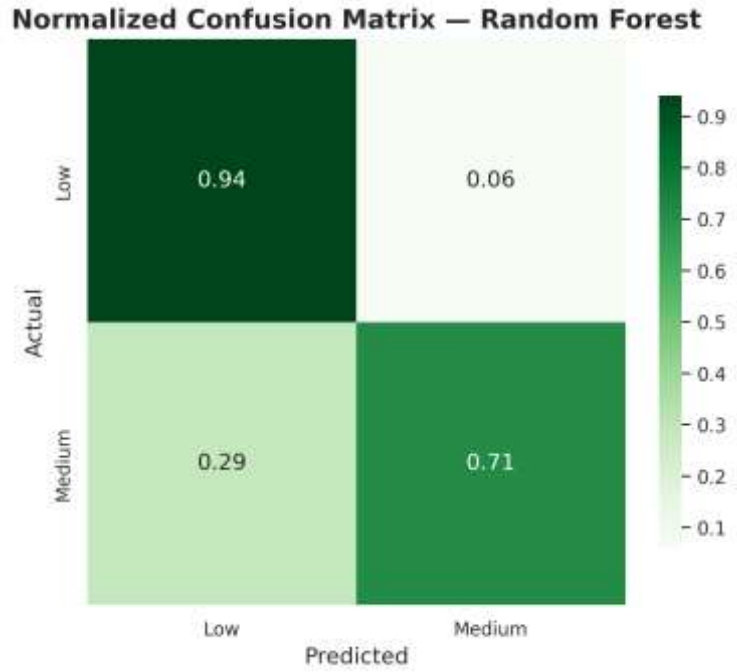


Fig 6. Normalized confusion matrix for the Random Forest model.

The ROC curve for the Random Forest model is shown in Figure 7. The model achieved an area under the curve (AUC) of 0.945, indicating a strong ability to distinguish between the two classes despite the class imbalance in the dataset. This performance is only slightly lower than that of the Gradient Boosting model, which achieved the highest AUC of 0.950. The small difference of 0.005 suggests that both tree-based ensemble methods provide comparable ranking performance.

Figure 8 presents the class-specific F1 scores. The model performed very well on the majority class (low congestion), achieving an F1 score of 0.915. However, performance on the minority class (medium congestion) was lower, with an F1 score of 0.768. This gap is expected because minority classes in imbalanced datasets typically contain fewer training examples, making them more difficult for the model to classify accurately.

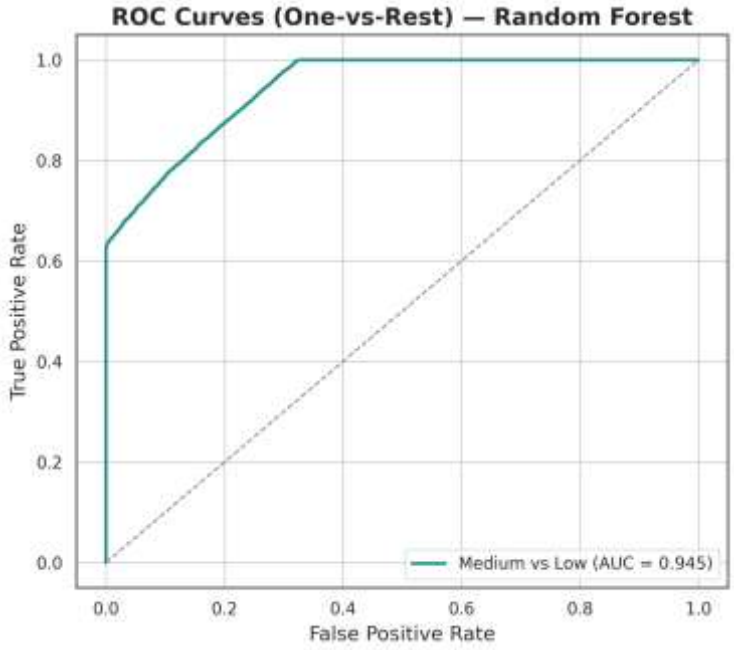


Fig 7. ROC curves for the Random Forest model.

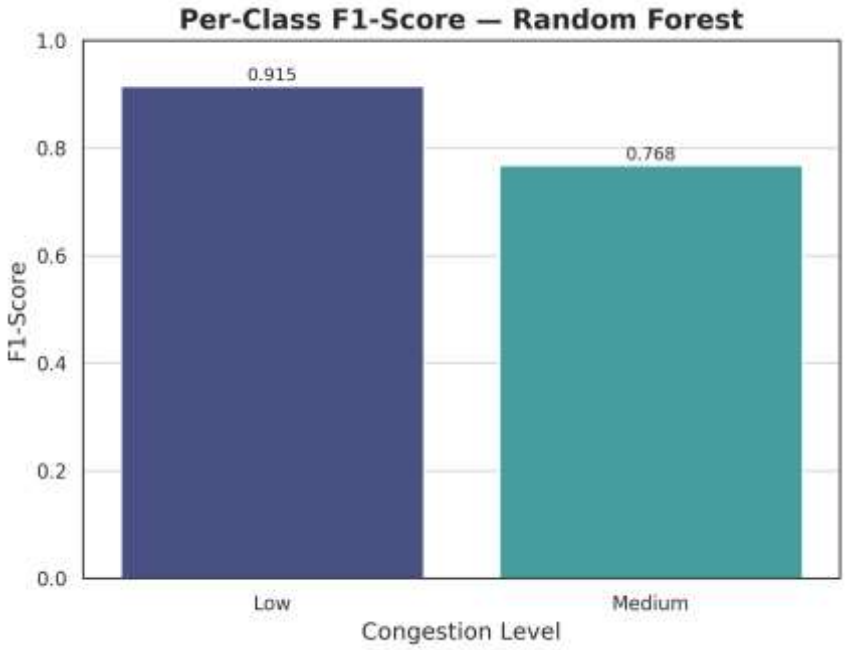


Fig 8. Class-wise F1 scores for the Random Forest model.

The feature importance results for the best-performing model are shown in Figure 9. Variables related to the time of day—particularly hour_cos, hour, hour_of_day, and hour_sin—were by far the most influential predictors of congestion. This indicates that traffic congestion is primarily driven by daily rush-hour patterns, while weather conditions and the remaining variables have a comparatively smaller impact. The dominance of these time-based features also highlights the nonlinear relationship between time and congestion, where traffic tends to peak during specific periods of the day. This characteristic helps explain why linear models are less effective than tree-based methods for this prediction task.

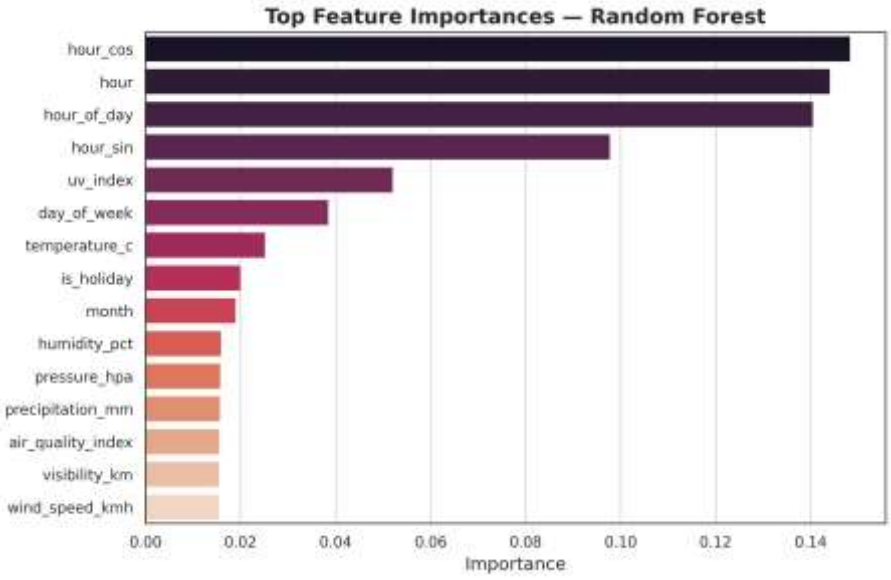


Fig 9. Feature importance ranking for the Random Forest model.

IV. CONCLUSION

In this study, traffic congestion levels were classified using machine learning based on only previously known contextual attributes (time, calendar, weather, and location) from a multi-source smart city dataset. The problem is considered as imbalanced dual classification (low/middle congestion). Real-time traffic measurements are excluded to avoid data leakage. Among the six algorithms compared, Random Forest showed the highest performance in terms of macro-F1, which was set as the primary metric (Accuracy = 0.875; Macro-F1 = 0.841). Gradient Boosting had the highest ROC-AUC value (0.950).

The most important finding is the significant superiority of tree and ensemble-based models on linear and probabilistic models. While Logistic Regression and Naive Bayes remain close to the majority baseline, Random Forest and Gradient Boosting significantly exceed this baseline. This is because, as feature significance analysis also shows, congestion exhibits a largely nonlinear pattern dependent on the time of day. This result is consistent with the literature reporting the superiority of ensemble methods in traffic condition estimation [3, 4, 6].

This study has several limitations that also point to opportunities for future research. First, the proposed approach predicts traffic congestion using only contextual information. Incorporating real-time sensor measurements, such as speed and occupancy, into a state estimation framework could further improve prediction accuracy. Second, including temporal lag features that capture the historical states of the same sensor may enhance the model's ability to identify traffic patterns over time. Third, the performance of the minority class could be improved by applying class-balancing techniques, such as SMOTE [24], along with systematic hyperparameter optimization. Finally, future studies could compare the proposed approach with deep learning models, such as Long Short-Term Memory (LSTM) networks and Graph Neural Networks (GNNs), which are designed to capture temporal and spatial dependencies more effectively [25].

ACKNOWLEDGMENT

The authors declare that no specific funding was received for this research.

REFERENCES

[1] A. M. Rao and K. R. Rao, “Measuring urban traffic congestion-a review.,” *International Journal for Traffic & Transport Engineering*, vol. 2, no. 4, p. 286, 2012.

[2] A. Rasaizadi, S. Seyedabrishami, and M. Saniee Abadeh, “Short-Term Prediction of Traffic State for a Rural Road Applying Ensemble Learning Process,” *Journal of Advanced Transportation*, vol. 2021, no. 1, p. 3334810, 2021, doi: 10.1155/2021/3334810.

[3] A. Arabiat and M. Altayeb, “Assessing the effectiveness of data mining tools in classifying and predicting road traffic congestion,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 34, p. 1295, May 2024, doi: 10.11591/ijeecs.v34.i2.pp1295-1303.

[4] Z. Cheng, W. Wang, J. Lu, and X. Xing, "Classifying the traffic state of urban expressways: A machine-learning approach," *Transportation Research Part A: Policy and Practice*, vol. 137, pp. 411–428, Jul. 2020, doi: 10.1016/j.tra.2018.10.035.
[5] N. Zafar and I. U. Haq, "Traffic congestion prediction based on Estimated Time of Arrival," *PLOS ONE*, vol. 15, no. 12, p. e0238200, Dec. 2020, doi: 10.1371/journal.pone.0238200.
[6] Y. Liu and H. Wu, "Prediction of Road Traffic Congestion Based on Random Forest," in *2017 10th International Symposium on Computational Intelligence and Design (ISCID)*, Dec. 2017, pp. 361–364. doi: 10.1109/ISCID.2017.216.
[7] S. Yang, J. Wu, Y. Du, Y. He, and X. Chen, "Ensemble Learning for Short-Term Traffic Prediction Based on Gradient Boosting Machine," *Journal of Sensors*, vol. 2017, no. 1, p. 7074143, 2017, doi: 10.1155/2017/7074143.
[8] B. Lartey, A. Homaifar, A. Girma, A. Karimoddini, and D. Opoku, "XGBoost: a tree-based approach for traffic volume prediction," in *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Oct. 2021, pp. 1280–1286. doi: 10.1109/SMC52423.2021.9658959.
[9] T. D. Toan and V.-H. Truong, "Support Vector Machine for Short-Term Traffic Flow Prediction and Improvement of Its Model Training using Nearest Neighbor Approach," *Transportation Research Record*, vol. 2675, no. 4, pp. 362–373, Apr. 2021, doi: 10.1177/0361198120980432.
[10] L. Zhang, Q. Liu, W. Yang, N. Wei, and D. Dong, "An Improved *K*-nearest Neighbor Model for Short-term Traffic Flow Prediction," *Procedia - Social and Behavioral Sciences*, vol. 96, pp. 653–662, Nov. 2013, doi: 10.1016/j.sbspro.2013.08.076.
[11] T. Bokaba, W. Doorsamy, and B. S. Paul, "A Comparative Study of Ensemble Models for Predicting Road Traffic Congestion," *Applied Sciences*, vol. 12, no. 3, p. 1337, Jan. 2022, doi: 10.3390/app12031337.
[12] M. R. A. Sara, M. Zeer, J. H. Alkhateeb, and K. Musa, "Machine Learning Classifier Comparative Performance Analysis of Smart Mobility: Multiclass Traffic Congestion Prediction," in *2026 IEEE 15th International Conference on Communication Systems and Network Technologies (CSNT)*, Apr. 2026, pp. 385–389. doi: 10.1109/CSNT69054.2026.11502201.
[13] X. Lu, C. Chen, R. Gao, and Z. Xing, "Prediction of High-Speed Traffic Flow around City Based on BO-XGBoost Model," *Symmetry*, vol. 15, no. 7, p. 1453, Jul. 2023, doi: 10.3390/sym15071453.
[14] N. Tsalikidis *et al.*, "Urban Traffic Congestion Prediction: A Multi-Step Approach Utilizing Sensor Data and Weather Information," *Smart Cities*, vol. 7, no. 1, pp. 233–253, Feb. 2024, doi: 10.3390/smartcities7010010.
[15] Q. Shang, Y. Yu, and T. Xie, "A Hybrid Method for Traffic State Classification Using K-Medoids Clustering and Self-Tuning Spectral Clustering," *Sustainability*, vol. 14, no. 17, p. 11068, Jan. 2022, doi: 10.3390/su141711068.
[16] C. Yazıcıoğlu and A. P. Akgüngör, "Modeling Short-Term Passenger Flows in Metro and Bus Systems Using Meteorological Data: Deep Learning Model Comparisons," *Applied Sciences*, vol. 15, no. 11, p. 6260, Jan. 2025, doi: 10.3390/app15116260.
[17] B. Qiu and W. (David) Fan, "Machine Learning Based Short-Term Travel Time Prediction: Numerical Results and Comparative Analyses," *Sustainability*, vol. 13, no. 13, p. 7454, Jan. 2021, doi: 10.3390/su13137454.
[18] T. S. Tamir *et al.*, "Traffic Congestion Prediction using Decision Tree, Logistic Regression and Neural Networks," *IFAC-PapersOnLine*, vol. 53, no. 5, pp. 512–517, Jan. 2020, doi: 10.1016/j.ifacol.2021.04.138.
[19] S. B. Çelebi, A. Aslan, and C. Yazicioglu, "Short-term wind power forecasting using KNN, SVR, and linear regression with high-resolution SCADA data," *Wind Engineering*, p. 0309524X261417206, Jan. 2026, doi: 10.1177/0309524X261417206.
[20] "Smart City Traffic Flow Prediction Dataset." Accessed: Aug. 04, 2026. [Online]. Available: https://www.kaggle.com/datasets/rauffauzanrambe/smart-city-traffic-flow-prediction-dataset
[21] L. B. Rashid, S. Z. Shuja, and S. Rehman, "Machine Learning Forecasting of Direct Solar Radiation: A Multi-Model Evaluation with Trigonometric Cyclical Encoding," *Forecasting*, vol. 7, no. 4, p. 58, Dec. 2025, doi: 10.3390/forecast7040058.
[22] S. Kapoor and A. Narayanan, "Leakage and the reproducibility crisis in machine-learning-based science," *Patterns*, vol. 4, no. 9, Sep. 2023, doi: 10.1016/j.patter.2023.100804.
[23] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, p. 6, Jan. 2020, doi: 10.1186/s12864-019-6413-7.
[24] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic Minority Over-sampling Technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, Jun. 2002, doi: 10.1613/jair.953.
[25] W. Jiang and J. Luo, "Graph neural network for traffic forecasting: A survey," *Expert Systems with Applications*, vol. 207, p. 117921, Nov. 2022, doi: 10.1016/j.eswa.2022.117921.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Building Extraction from Remote Sensing Imagery: Quantifying Label Efficiency and the Effect of Transfer Learning on Annotation Cost

Ammar Aslan [1], Cafer Yazicioglu [2*]

[1] *Vocational School of Technical Sciences, Batman University, Batman, Turkey*

[2] *Department of Construction, Vocational School of Technical Sciences, Batman University, Batman, Turkey*

Corresponding Author: Cafer Yazicioglu, Phone:+905345725423, e-mail:cafer.yazicioglu@batman.edu.tr

Abstract – Deep learning based extraction of building footprints from very-high-resolution remote sensing imagery is constrained less by network design than by the cost of pixel-level annotation. This study quantifies how much labelled data building segmentation actually requires, and how much of that requirement ImageNet pretraining removes. A two-factor full factorial design was set up on the aerial subset of the WHU Building Dataset (8,189 tiles of 512 × 512 pixels at 0.3 m GSD). The label fraction was varied over six levels (1%, 5%, 10%, 25%, 50%, 100%) and the encoder was initialised either from ImageNet weights or from scratch. Architecture (FPN with a ResNet-18 encoder), loss function, augmentation and evaluation protocol were held constant, and each of the twenty-four runs received an identical budget of 2,500 gradient steps, so that data quantity is not confounded with training duration or compute. Performance saturates early: 57 annotated tiles reach 79.81% IoU, 97.0% of the fully labelled baseline (82.31%), and 287 tiles reach 99.0%, while the subsequent twentyfold increase in annotation adds only 0.84 points. The gain from pretraining did not narrow as the label budget grew, it remained between 4.00 and 4.46 points across a hundredfold range. The two curves never overlap, so no data equivalence coefficient could be computed; a pretrained model trained on 57 tiles exceeded a from-scratch model trained on 5,732 tiles by 1.83 points. Pretraining also reduced seed-to-seed variability in the low-label regime by a factor of roughly three.

Keywords – Building extraction; remote sensing; aerial imagery; transfer learning; semantic segmentation

I. INTRODUCTION

Extracting building footprints automatically and precisely from very high resolution remote sensing imagery is a basic requirement for smart city planning, cadastral and map updating, building inventory compilation, and post-disaster damage assessment. The labour demand and time cost of conventional pixel-based classification and manual digitisation have pushed the field towards deep learning based semantic segmentation models. In Türkiye, building footprint data is compiled within an established photogrammetric workflow in which orthophotos are vectorised by trained operators [1], and this national base is being extended by an active research effort on building extraction from high-resolution imagery, following both conventional image processing [2] and deep learning [1], [3] lines.

The WHU Building Dataset, released by Wuhan University to support the evaluation of algorithmic progress in this area, has become a widely used benchmark. Ji et al. [4] published it with both raster and vector labels. It comprises an aerial subset covering 450 km² in the Christchurch region of New Zealand and containing more than 220,000 individual buildings, together with two satellite subsets: the first compiled from 204 images of cities worldwide, and the second consisting of six adjacent images covering 550 km² in East Asia. A cropped portion of the aerial subset, containing 187,000 buildings, was downsampled from a GSD of 0.075 m to 0.3 m and divided into 8,189 non-overlapping 512 × 512 tiles.

The same study proposed SiU-Net, which processes the original image and a downsampled copy in two weight-sharing branches; on the aerial subset, IoU rose from 86.8% for U-Net to 88.4%, indicating that the scale problem posed by large buildings can be mitigated. Liu et al. [5] reported reference values for the standard baselines on the same subset in their evaluation of DE-Net, obtaining F1 scores between 92.80% and 93.91% for SegNet, DeepLabV3+ and U-Net.

Subsequent architectural work has concentrated on three axes. Wang and Miao [6] proposed RU-Net, in which modified residual blocks placed in the U-Net backbone reduce network degradation and an atrous spatial pyramid pooling (ASPP) unit bridging the encoder and the decoder captures multi-scale context; on the WHU aerial subset it achieved 97.23% F1 and 94.61% IoU, ahead of U-Net (96.87% F1, 93.94% IoU), SegNet (96.72%, 93.65%) and DeepLabV3+ (88.52%, 79.41%) as reproduced by the same authors under an identical protocol. The same study reported that the contribution of Focal Loss depends on the dataset: IoU fell by 0.06 points on WHU, whose labels are highly accurate, but rose by 1.73 points on Inria, where label quality is lower. Yang et al. [7] proposed the edge-aware EANet, which supervises building boundaries through a separate edge-detection network and the EALoss term; under the same experimental conditions it obtained 93.33% IoU on the WHU validation set, compared with 91.96% for DeepLabV3+ and 91.82% for PSPNet. On the transformer side, Chang et al. [8] developed CTANet, which combines local convolutional features with global contextual representations; on the WHU aerial subset it reached 95.74% F1 and 91.84% IoU, ahead of BuildFormer (95.53%, 91.44%) and Swin-Unet (94.46%, 89.50%) at lower computational cost. In ablation experiments, the IoU of the baseline ConvNeXt-Tiny rose from 81.51% to 82.68% on Inria and from 75.24% to 76.52% on Massachusetts.

This accumulation on the architectural side has not removed the actual bottleneck in practice. Supervised segmentation requires pixel-level annotated imagery, so the binding constraint is the labels rather than the model. Zheng et al. [9] proposed the weakly supervised BAC-AR-Net, which identifies shared building features by clustering features extracted from many building instances and uses them to expand the building regions in class activation maps, reaching IoU values of 85.56%, 81.63% and 77.97% on ISPRS Potsdam, Vaihingen and their own WHU-RRDSD dataset using image-level labels alone. Hua et al. [10] complemented the supervised loss on sparse point-, scribble- and polygon-level annotations with FESTA, which regularises spatial and feature relations, and showed that annotation cost can be reduced several-fold while partly closing the performance gap with a densely labelled model. Xia et al. [11] used a small number of labelled samples together with abundant unlabelled imagery in SDLED, a semi-supervised edge detection approach built on D-LinkNet, and with only half the labelled samples obtained improvements of at least 6.47% line IoU and 7.49% F1 over the fully supervised approach in generalisation tests on neighbouring districts.

A second line of attack on annotation cost changes the initial weights rather than the method. Liu et al. [5] evaluated DE-Net, an architecture designed for building extraction, on the official WHU split. They reported that SegNet (86.58% IoU), RefineNet (88.03%) and DeepLabV3+ (88.00%), all with ImageNet-pretrained encoders, were surpassed by U-Net (88.52%) and DE-Net (90.12%) trained from scratch, and argued that where training data are sufficient, training from scratch can outperform fine-tuning a model pretrained on an unrelated dataset. Architecture, parameter count and learning rate were not held constant across conditions in that comparison, however: the pretrained models were trained with an initial learning rate of 1×10^{-4} and those trained from scratch with 1×10^{-2}. The observed difference therefore cannot be attributed to initialisation alone. Neupane et al. [12] compared three transfer learning strategies between WHU (TR1) and a Melbourne dataset of their own construction (TR2), reporting that their surgical fine-tuning approach preserved accuracy while reducing the number of trainable parameters by a factor of 300 and training time by a factor of 2.5. Vats et al. [13] learned in-domain self-supervised representations from LiDAR-derived elevation models through a pretext task defined as DTM estimation from DSM, and with only 1% of the labels (25 annotated samples) exceeded ImageNet pretraining by 0.022 points in IoU while falling 0.134 points behind in boundary IoU. At a 10% label fraction the advantage extended to all three metrics, whereas under the fully labelled regime the difference between the two approaches largely disappeared and ImageNet retained a very small edge. P. Liu et al. [14] likewise reported that a two-stage framework, following general-purpose pretraining (GPPT) with domain-adapted pretraining (DAPT),

outperformed the ImageNet-initialised paradigm on Potsdam, and confirmed the contribution of the DAPT stage on four remote sensing datasets with differing spectral and spatial characteristics, the gain being largest at a 1/100 label ratio and narrowing overall as the amount of labelled data increased.

The interaction between pretraining and label budget has been tested directly in the computer vision literature. Zoph et al. [15] showed on COCO object detection that, under the strongest data augmentation setting, ImageNet pretraining gave a gain of 2.6 AP when one fifth of the labels was used but lowered accuracy by 1.0 AP in the fully labelled condition. In the same experiments the benefit of pretraining fell from +0.3 AP to −1.0 AP as augmentation strength increased. On the PASCAL segmentation set, which contains only 1,500 images, pretraining instead raised mIoU from 41.5 to 81.4 under the same strong augmentation, showing that the effect of pretraining depends on the task and on the scale of the data. Kornblith et al. [16] examined sixteen networks across twelve datasets and found a strong relationship between ImageNet performance and transfer performance ($r = 0.99$ for fixed features and $r = 0.96$ for fine-tuning), while reporting that the gain from pretraining remained negligible on two small fine-grained datasets whose classes are poorly represented in ImageNet. These studies show that the gain from pretraining can narrow as the amount of data grows, but they are limited to natural-image datasets; in the literature surveyed, no design was found in which the same question is tested directly for building segmentation from high-resolution aerial orthoimagery. In this study, all label fractions are therefore evaluated under the same number of gradient steps, so that the comparison is not affected by the compute budget.

In this paper we set up a two-factor full factorial design on the aerial subset of the WHU Building Dataset. The first factor is the label fraction, the second the initialisation of the encoder. The model and the training protocol are held constant across all conditions, and the number of gradient steps rather than the number of epochs is equalised, which isolates the effect of data quantity from that of training duration and computational effort. The lowest label fractions are repeated with several seeds, so that the variability arising from subset selection is measured as well. The contributions of the study are: (i) quantifying the saturation point of the relationship between label quantity and performance in building segmentation; (ii) measuring, under an equal compute budget, how the gain from ImageNet pretraining behaves across the label budget; and (iii) reporting the effect of initialisation on outcome variability in the low-label regime.

II. MATERIALS AND METHODS

A. Dataset

The study uses the aerial subset of the openly available WHU Building Dataset, compiled for building extraction from aerial imagery [4]. The dataset was produced from orthophotos with a 0.3 m ground sampling distance and consists of 8,188 tiles of 512 × 512 pixels. The labels are binary masks.

The dataset was obtained through a version distributed via a public repository that splits the tiles 70% training (5,732 tiles), 15% validation (1,228 tiles) and 15% test (1,228 tiles). This differs from the split defined in the original distribution (4,736 / 1,036 / 2,416), so the absolute performance values reported here should not be compared directly with publications using the original split. Since the focus of the study is relative comparison across conditions, and the split is identical in every condition, this does not affect the internal validity of the results. The validation set was used for model selection only, and the test set was not seen at any stage of the training process.

B. Research Question and Experimental Design

Annotating building footprints is labour intensive and costly. The question a decision maker faces in practice is less which architecture performs best than how many tiles must be annotated to reach an acceptable level of performance. The study sets out to answer two questions quantitatively: (i) how the relationship between the amount of labelled data and segmentation performance behaves, and where it saturates; and (ii) to what extent ImageNet pretraining reduces the annotation requirement.

A two-factor full factorial design was set up for this purpose. The first factor is the labelled data fraction (1%, 5%, 10%, 25%, 50%, 100%), the second the initialisation of the encoder (ImageNet pretrained or from

scratch). The model is held constant across all conditions, which ensures that observed differences can arise only from these two factors.

Which tiles are sampled can affect the outcome markedly in the low label regime, so fractions of 10% and below were repeated with three different seeds. The seed controls both subset selection and model initialisation, so that the question "what if a different subset had been annotated" is modelled correctly. Twenty-four runs were carried out in total (Table 1).

Table 1. Experimental grid. Each condition was trained separately with two initialization schemes (ImageNet / from scratch)

Label fraction (%)	Labeled tiles	Seeds	Dataset passes	Number of runs
1	57	3	833.3	6
5	287	3	147.1	6
10	573	3	71.4	6
25	1,433	1	28.1	2
50	2,866	1	14	2
100	5,732	1	7	2
Total	—	—	—	24

C. Fixed Gradient-Step Budget

The methodological backbone of the design is that the number of gradient steps, not the number of epochs, is held fixed across conditions. Every condition received 2,500 weight updates. Had epochs been fixed instead, the 1% condition would have seen one hundredth of the updates of the 100% condition, leaving the low-label conditions systematically undertrained; the measured difference would then have confounded the effect of data quantity with the effect of training duration.

A fixed budget makes the number of passes over the labelled subset vary widely across conditions: the 1% condition sees its 57-tile subset roughly 833 times, while the 100% condition sees its 5,732 tiles only 7 times (Table 1). These values were logged in the output file of each run.

A second and less obvious benefit of this choice is that it equalises computational cost as well. Run times for the twenty-four runs varied between 42.4 and 45.2 minutes, averaging 43.1 ± 0.6 minutes. The comparison is therefore fair not only in the number of updates but in the compute expended; no condition consumed more compute than any other.

D. Model and Training Settings

All conditions use a Feature Pyramid Network (FPN) [17] with a ResNet-18 encoder [18]. The model has 13.05 million trainable parameters. It was kept deliberately plain: the question addressed here is label efficiency rather than architectural performance, and a complex architecture would make the measurement needlessly expensive.

Training used the AdamW optimiser (weight decay 1×10^{-4}), a learning rate decayed on a cosine schedule after 150 iterations of linear warm-up, a batch size of 16, and mixed precision. The base learning rate is 3×10^{-4}. In the pretrained condition the encoder received 0.3 times this value, which prevents early corruption of the transferred features; in the from-scratch condition a discriminative learning rate would be meaningless, so a single parameter group was used. The gradient norm was clipped at 1.0.

Training used random 256 × 256 crops, with horizontal and vertical flips, rotation in multiples of 90°, an affine transform comprising ±15° rotation and 0.90–1.15 scaling, and colour jitter. Image and mask were processed with the same border padding strategy in the affine transform, which prevents label noise. Evaluation was carried out on full 512 × 512 tiles without augmentation.

Because building pixels form the minority class, the loss function is an equally weighted combination of binary cross-entropy and a Dice term:

$$L = 0.5L_{BCE} + 0.5L_{Dice}$$

E. Plain Evaluation Protocol

Weight averaging (EMA), test-time augmentation (TTA) and decision threshold tuning were deliberately not used. The decision threshold was fixed at 0.50 in every condition. Refinements of this kind at the inference stage raise performance somewhat, but their effect can differ in magnitude across conditions and blur the measurement of label efficiency. A plain protocol ensures that observed differences arise from the experimental factors alone.

F. Evaluation Metrics

Performance was evaluated with intersection over union (IoU), precision, and recall. Rather than being computed per tile and averaged, the metrics were computed from a confusion matrix accumulated over the whole test set:

$$\text{IoU} = \frac{TP}{TP + FP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

Here, TP, FP, and FN denote the numbers of true-positive, false-positive, and false-negative pixels, respectively. For conditions repeated with three seeds, the mean and standard deviation are reported.

G. Hardware and Software Environment

The experiments were run on an NVIDIA Tesla T4 (15 GB) GPU. The twenty-four runs took approximately 17.2 hours in total. The implementation was developed in Python using PyTorch and segmentation-models-pytorch (version 0.5.0).

III. RESULTS

A. Label Efficiency Curve

Test set performance for all conditions is given in Table 2, and its variation with the number of labelled tiles in Figure 1. The relative performance column is computed as a percentage of the fully labelled pretrained condition (82.31% IoU).

Table 2. Test-set performance by label budget and initialization scheme. Standard deviations are reported only for conditions repeated with three seeds; the "—" symbol indicates a single-seed run and does not denote zero variability. Relative performance is computed with respect to the fully labeled pretrained condition (82.31%).

Label ratio (%)	Labeled tiles	Initialization	IoU (%)	Std	Relative to baseline (%)
1	57	ImageNet	79.81	0.38	**97.0**
5	287	ImageNet	81.47	0.21	**99.0**
10	573	ImageNet	81.84	0.12	99.4
25	1,433	ImageNet	82.16	—	99.8
50	2,866	ImageNet	82.15	—	99.8
100	**5,732**	**ImageNet**	**82.31**	**—**	**100.0**
1	57	From scratch	75.48	1.13	91.7
5	287	From scratch	77.47	0.09	94.1

10	573	From scratch	77.82	0.22	94.6
25	1,433	From scratch	**77.70**	—	94.4
50	2,866	From scratch	78.11	—	94.9
100	5,732	From scratch	77.98	—	94.7

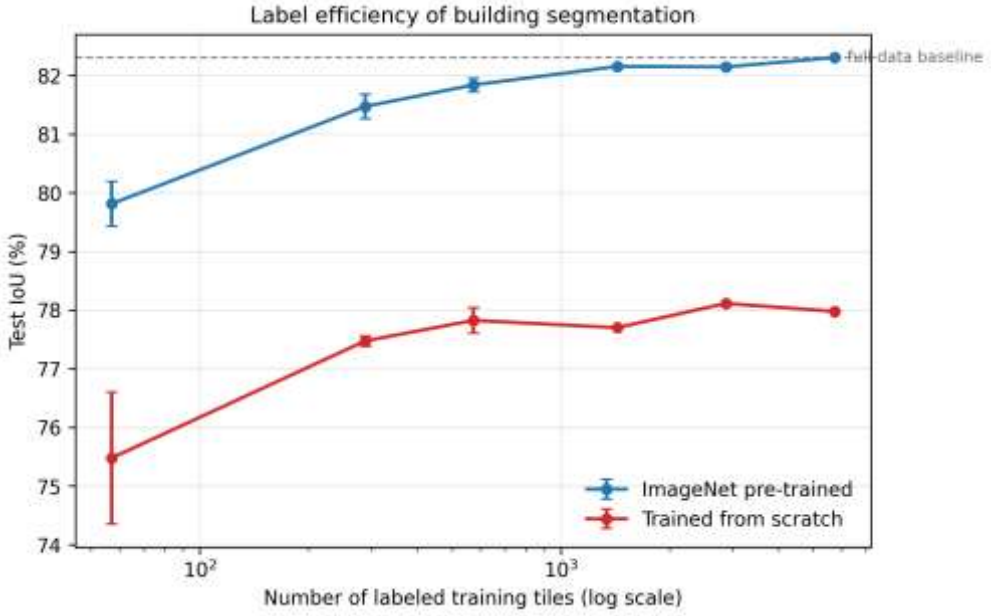


Fig 1. Test IoU as a function of the number of labeled tiles. Error bars indicate the standard deviation across three seeds; no bars are shown for single-seed conditions.

The most salient feature of Figure 1 is that both curves saturate early. The pretrained condition reaches 79.81% IoU with only 57 annotated tiles, which corresponds to 97.0% of the fully labelled pretrained condition. Raising the number of labelled tiles to 287 (a fivefold increase) brings performance to 99.0% of that level, and 573 tiles to 99.4%. Beyond this point, the additional gain remains below 0.5 IoU points. Marginal returns fall sharply. Raising the number of labelled tiles from 57 to 287 (fivefold) yields 1.66 IoU points, while raising it from 287 to 5,732 (twentyfold) yields only 0.84 points. In practical terms, the highest return comes from spending the annotation budget on the first few hundred tiles, and investment beyond that yields rapidly diminishing returns.

The curve for the from-scratch condition has a similar overall shape, but saturates at a markedly lower level. The highest value reached in this condition is 78.11%, obtained with 2,866 tiles, and even the full label budget does not exceed it (77.98%).

B. Pretraining Gain Across the Label Budget

The behaviour commonly expected in the transfer learning literature is that the gain from pretraining narrows as the label budget grows and disappears once the label budget is large enough. The findings of this study do not support that expectation (Table 3, Figure 2).

Table 3. Absolute IoU gain from ImageNet pretraining as a function of label budget.

Label fraction (%)	Labeled tiles	IoU ImageNet (%)	IoU from scratch (%)	Gain (points)
1	57	79.81	75.48	**4.34**
5	287	81.47	77.47	**4.00**
10	573	81.84	77.82	**4.02**
25	1,433	82.16	77.70	**4.46**
50	2,866	82.15	78.11	**4.04**
100	5,732	82.31	77.98	**4.33**
Mean ± Std	—	—	—	**4.20 ± 0.20**

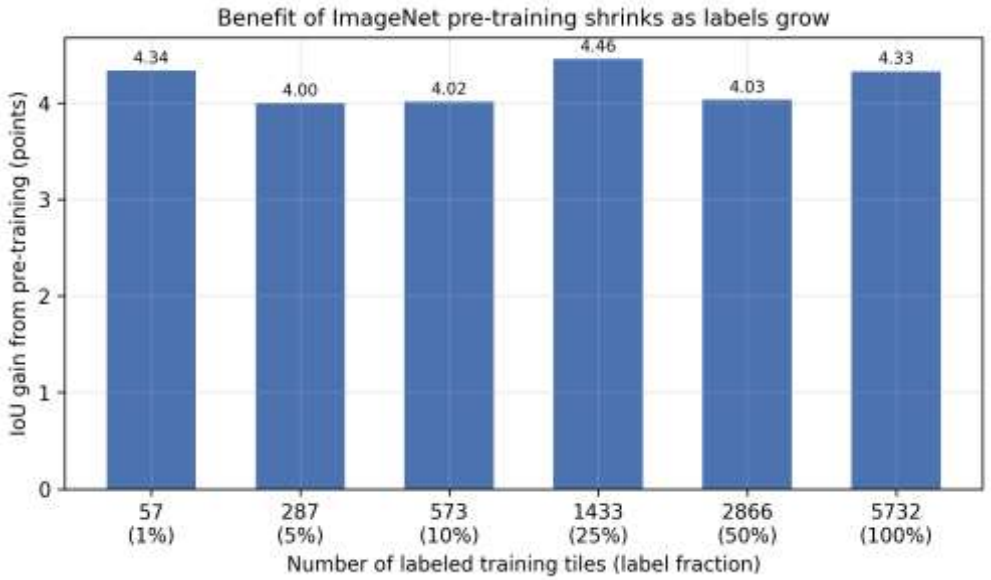


Fig 2. Pretraining gain is independent of the label budget. The dashed line denotes the mean and the shaded band ±1 standard deviation.

Across a hundredfold range from 57 to 5,732 tiles, the gain remains between 4.00 and 4.46 points, averaging 4.20 ± 0.20 points. Pretraining therefore provides an additive contribution maintained across the whole budget range, not a initial advantage that weakens as data accumulate.

The scope of this finding needs to be set carefully. The constancy of the gain was measured under fixed gradient-step and compute budgets; because the same protocol was applied at every label level, that constancy is directly observed. The practical meaning is nonetheless clear: compute budgets are limited under real application conditions, and within that limit pretraining provides an advantage of roughly four points independently of label quantity. The low-level edge and texture features transferred by ImageNet pretraining appear to provide a level of representation that the WHU dataset cannot yield on its own within the given budget.

C. Data Equivalence: A Ratio That Could Not Be Computed

One of the outputs anticipated at the outset was a data equivalence coefficient computed from the horizontal distance between the two curves: how many times more annotated tiles training from scratch requires to reach a given level of performance. The coefficient could not be computed, but the reason is not a methodological shortcoming; it is that the finding is stronger than expected.

An equivalence coefficient requires the two curves to overlap over a common performance range. The highest value reached by the from-scratch condition is 78.11%, while the lowest value of the pretrained

condition is 79.81%. The two curves do not overlap at any performance level; the gap between them is 1.70 points.

The practical meaning can be stated directly: under an equal compute budget, a pretrained model trained on only 57 annotated tiles outperforms a from-scratch model trained on 5,732 annotated tiles by 1.83 IoU points. No measurable upper bound for the equivalence coefficient is therefore available, only a lower bound: pretraining is equivalent to at least a hundredfold increase in labelled data. Across the budget range tested, training from scratch does not reach even the lowest performance of the pretrained model. A comparative view sharpens the result. Raising the number of labelled tiles from 57 to 5,732, that is, increasing the annotation workload a hundredfold, yields 2.50 IoU points in the pretrained condition. Using pretrained weights instead of training from scratch at a budget of 57 tiles yields 4.34 points at no additional annotation cost. At this scale, pretraining is worth more than a hundredfold increase in labels.

D. Log-Linear Fit and Its Limitations

A log-linear model of the form $IoU = a + b \cdot \log_{10}(n)$ was fitted to summarise the relationship between the number of labelled tiles and performance (Table 4). The slope coefficient b is interpreted as the IoU gain from a tenfold increase in the number of labelled tiles. The slopes of the two curves are practically identical (1.18 and 1.17). The 4.15-point difference between the intercepts is consistent with the mean gain reported earlier (4.20 points). This confirms quantitatively that the two curves are parallel on the log axis, and that pretraining changes the level rather than the slope.

Table 4. Log-linear fit coefficients.

Initialization	**Slope b (IoU / decade)**	**Intercept a**	**R^2**
ImageNet	1.18	78.20	0.85
From scratch	1.17	74.05	0.75

The explanatory power of the log-linear model is limited, however (R^2 = 0.846 and 0.748), because the relationship is not in fact log-linear but saturating. In the pretrained condition the slope is 2.36 points per decade over the 57–287 tile range and falls to 0.65 points per decade over the 287–5,732 tile range, a ratio of roughly 3.6. A single slope coefficient averages these two regimes, and so understates the early gain and overstates the late one.

E. Training Dynamics

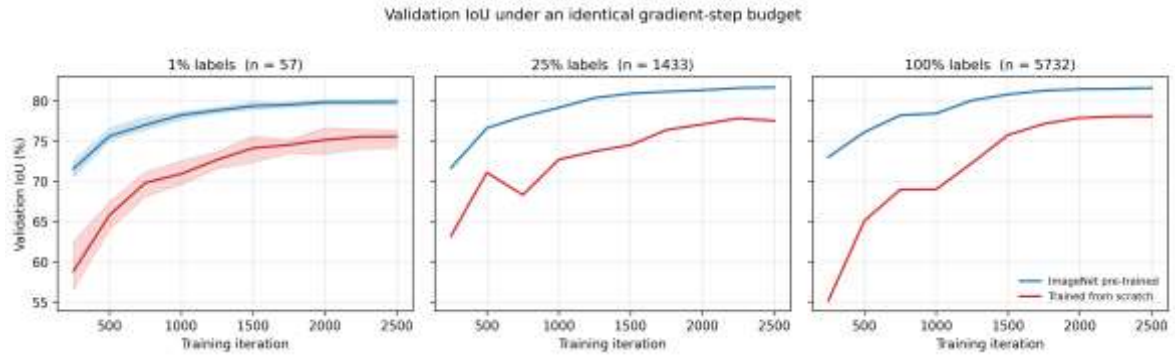


Fig 3. Validation IoU trajectories under an identical gradient-step budget. The shaded band indicates the minimum–maximum range across three seeds and is shown only for the 1% condition.

Figure 3 shows the initial advantage given by pretraining. At the first evaluation point of the 1% condition (iteration 250) the pretrained model reaches 71.5% validation IoU while the from-scratch model reaches only 58.8%. The advantage narrows over the course of training but has not closed by the end of the budget. In the fully labelled condition the from-scratch model begins to plateau only after iteration 2,000.

The convergence state at the end of the budget determines how the findings should be read, so it was examined across all runs. In sixteen of the twenty-four runs, the highest validation IoU was observed at the

final iteration (2,500); this was the case for 8 of 12 runs in the pretrained arm and 8 of 12 runs in the from-scratch arm. The mean IoU increase over the last 750 iterations is 0.38 points in the pretrained arm (range 0.20–0.71) and 0.95 points in the from-scratch arm (range 0.49–1.42).

These numbers point to a clear caveat: both arms reached the budget ceiling, that is, neither reached full convergence. The residual slope of the pretrained arm is smaller by roughly a factor of 2.5, but the budget remains a binding constraint for both arms. The present experiment therefore cannot separate how much of the 4.20-point difference between the arms comes from the quality of the solution reached and how much from the speed of convergence. The finding that the gain is independent of the label budget holds, since the same protocol was applied at every label level, but it should be read with the qualification "under an equal compute budget".

One fact is independent of this caveat: within the given budget both arms had saturated with respect to the number of labelled tiles. Moving from 287 to 5,732 tiles, that is, to twenty times more labelled data, yields only 0.84 points in the pretrained arm and 0.51 points in the from-scratch arm. The binding constraint in this regime is not the amount of labelled data; further investment in annotation does not pay off, whichever initialisation is used. The curve of the from-scratch model in the 25% condition shows a temporary drop between iterations 500 and 750. Fluctuations of this kind are expected behaviour in randomly initialised networks near the learning-rate peak that follows warm-up, and they did not affect final performance.

F. Seed Variability in the Low-Label Regime

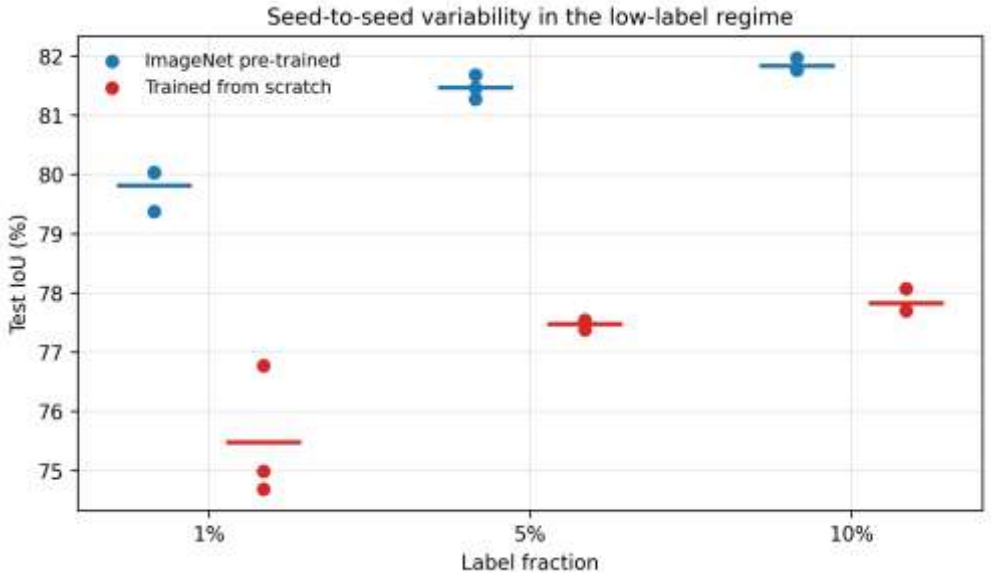


Fig 4. Individual run results for conditions repeated with three seeds. Horizontal lines denote the means.

Figure 4 and Table 5 set out the risk carried by seed selection in the low label regime. In the 1% condition the three runs of the from-scratch model gave 74.68%, 74.98% and 76.77%, a range of 2.09 points and a standard deviation of 1.13 points. The three runs of the pretrained model in the same condition fell between 79.37% and 80.04%, a range confined to 0.67 points. Pretraining reduces variability by roughly a factor of three.

Table 5. Individual run results for the 1% and 5% label-ratio conditions repeated with three seeds (test IoU, %).

Label ratio	**Initialization**	**Seed 42**	**Seed 43**	**Seed 44**	**Mean**	**Range**
1%	ImageNet	80.04	80.03	79.37	**79.81**	**0.67**
1%	From scratch	76.77	74.98	74.68	**75.48**	**2.09**
5%	ImageNet	81.68	81.27	81.46	**81.47**	**0.41**
5%	From scratch	77.37	77.51	77.54	**77.47**	**0.17**

This points to a second and less emphasised benefit of pretraining: it does not only raise mean performance, it also reduces sensitivity to which tiles were annotated. For a practitioner annotating on a limited budget, that means predictability of the outcome.

Variability falls quickly as the number of labels grows. At 287 tiles the standard deviation drops below 0.21 points in both conditions, and to 0.09 points in the from-scratch arm. The same trend continues at 573 tiles: the three runs of the from-scratch arm gave 77.70%, 77.70% and 78.07%, a range of 0.37 points. The variance-reducing effect of pretraining is no longer observed at this level, so it is specific to the lowest budgets, where subset selection is decisive. This observation supports the design choice of a single seed at fractions of 25% and above, though it should be remembered that variability was not measured for those conditions.

G. Precision–Recall Balance

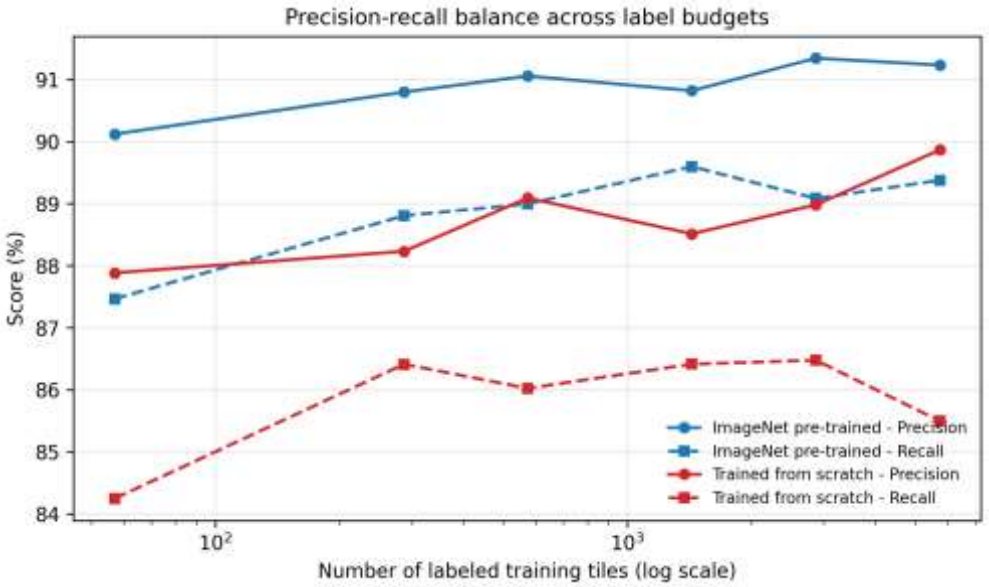


Figure 5. Precision and recall as a function of label budget. Solid lines denote precision and dashed lines recall.

Figure 5 shows that the difference between the two conditions lies in the character of the errors as well as in the level. Precision and recall run relatively close together in the pretrained model, with a difference of 1.85 points in the fully labelled condition (91.23% against 89.38%). In the from-scratch model the gap widens. In the fully labelled condition precision is 89.87% while recall stands at 85.52%, a difference of 4.35 points. High precision combined with low recall indicates that the model predicts building regions more narrowly than they are, that is, it tends towards under-segmentation. This shows up in particular as small buildings and building edges being missed.

This divergence is not uniform across the label budget. The precision-recall difference in the 1% condition is 2.65 points in the pretrained arm and 3.63 points in the from-scratch arm, while in the 5% condition the two arms converge (1.99 and 1.82 points) and the ordering reverses. Averaged over all twenty-four runs, the difference is 2.14 points in the pretrained arm (range 1.22–3.04) and 2.88 points in the from-scratch arm (range 0.53–5.74). The width of the ranges is as informative as the difference in means: the error balance of the from-scratch arm is far less stable from run to run.

The contribution of pretraining should therefore be read not only as an increase in accuracy but as a balancing of the error distribution. Where under-segmentation costs more than over-segmentation, as in compiling a building inventory, this difference matters in practice.

H. Qualitative Evaluation

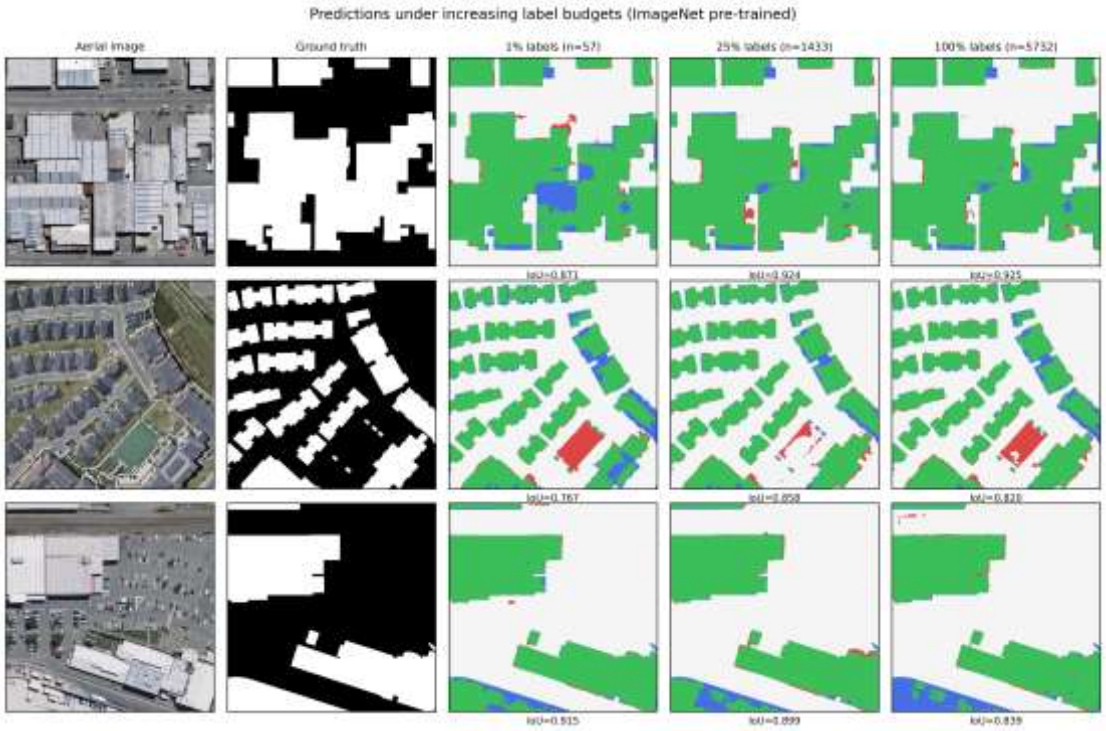


Fig 6. Predictions of the pretrained model under increasing label budgets. In the error maps, green indicates true positive, red false positive, and blue false negative pixels.

Figure 6 gives the visual counterpart of the quantitative findings. In the dense commercial area in the first row, the difference between the prediction obtained with 57 annotated tiles (IoU = 0.871) and the one obtained with 5,732 tiles (IoU = 0.925) consists mainly of fine corrections at building boundaries and the closing of a false-negative region in the central region. Building bodies are segmented correctly at all three budgets.

The second row shows a pattern worth noting: the large false-positive region (red) in the right central region remains almost unchanged at all three budgets. The structure in question is not annotated as a building in the reference mask, yet the model consistently classifies it as one. Errors of this kind cannot be removed by increasing the number of labelled tiles; their source is not model capacity or data quantity but the reference annotation itself, or the visual ambiguity of the structure.

The third row shows that per-tile performance does not rise monotonically with the number of labelled tiles: the IoU values are 0.915, 0.899 and 0.839 respectively. The wide false-negative band along the lower edge widened as the number of labelled tiles grew. Improvement in aggregate metrics is therefore not reflected uniformly in individual tiles, and for some individual tiles, additional data may not improve performance.

IV. DISCUSSION

Label saturation sets in early and the return on annotation investment falls quickly. The pretrained condition reaches 79.81% IoU with only 57 annotated tiles (1%), which corresponds to 97.0% of the performance of the fully labelled pretrained condition (82.31%). At 287 tiles the ratio rises to 99.0%. A fivefold increase in the number of labelled tiles yields 1.66 points, while the subsequent twentyfold increase yields only 0.84 points. The same asymmetry appears in the piecewise slopes of the log-linear fit: 2.36 points per decade over the 57 to 287 tile range, falling to 0.65 points per decade over the 287 to 5,732 range. For an institution with a limited budget this means that annotating a few hundred tiles carefully for training returns markedly more than investing in an expansion of the training set to thousands of tiles. The

qualitative assessment points to one source of the saturation: false-positive regions arising from structures left unannotated in the reference mask are not removed even when the number of labelled tiles is increased a hundredfold.

No evidence was found that the pretraining gain narrows with the label budget. Across a hundredfold range the gain stayed between 4.00 and 4.46 points, averaging 4.20 points, and the slopes of the log-linear fit are practically identical in the two arms (1.18 and 1.17). Because the two curves do not overlap at any performance level, no data equivalence coefficient could be computed: the pretrained model trained on 57 tiles outperformed the from-scratch model trained on 5,732 tiles by 1.83 points. This pattern departs from the established expectation in the transfer learning literature. Zoph et al. [15] reported that in COCO object detection the benefit of ImageNet pretraining disappears in the fully labelled condition, and that under strong data augmentation it lowers accuracy by 1.0 AP. Vats et al. [13] showed that the advantage of in-domain self-supervised representations over ImageNet pretraining closes as the label fraction increases. The most likely source of the discrepancy is the training regime. Zoph et al. trained all their models until convergence and reported that pretraining accelerates convergence by a factor of 1.3–8; by contrast, all conditions here were bounded by a binding step budget. The gain measured under a limited budget is therefore larger than the gain that would be measured in a converged regime.

The fixed budget choice affects the comparison differently for the two factors. Across label fractions, budget equality is neutral: every condition receives the same number of updates. Across initialisations, the comparison structurally favours the pretrained arm. In sixteen of the twenty-four runs, the highest validation performance was observed at the final iteration, and the increase over the last 750 iterations was measured as 0.38 points in the pretrained arm against 0.95 points in the from-scratch arm; the difference is still narrowing at the end of the budget. How much of the 4.20-point difference comes from the quality of the final solution and how much from the speed of convergence therefore cannot be separated, and the reported gain should be regarded as an upper bound. One observation is independent of this caveat: within the given budget both arms saturated with respect to the amount of labelled data; moving from 287 to 5,732 tiles yields 0.84 points in the pretrained arm and 0.51 points in the from-scratch arm. The main practical implication of the study rests on this observation.

The effect of pretraining is not limited to mean performance: in the 1% condition the standard deviation across seeds is 1.13 points for training from scratch against 0.38 points for the pretrained condition, and in the fully labelled condition the precision–recall difference is 4.35 and 1.85 points respectively. Both effects are specific to those conditions, however; they disappear at 287 tiles and above, and the ordering in variability even reverses. They are therefore not presented as a general property.

The limitations of the study fall under three headings. The first is the convergence constraint discussed above. The second is that the optimisation procedure is not exactly identical in the two arms: the discriminative learning rate applied to the encoder in the pretrained arm results in a 3.3-fold difference in encoder learning rate between the arms, and the base learning rate was not tuned separately for each. The third is that the label fraction was applied to the training set only, the validation set being kept complete in every condition; the reported label budgets therefore fall short of the total annotation cost, and the values at the lowest fractions should be interpreted as an optimistic upper bound. Fractions of 25% and above were run with a single seed, and the results are confined to one architecture and one dataset; because the split of the version used differs from that of the original distribution, the absolute values should not be compared directly with the literature.

V. CONCLUSION

This study examined quantitatively the relationship between the amount of labelled data and performance in building segmentation from high-resolution aerial orthoimagery, and the effect of ImageNet pretraining on that relationship. Architecture (FPN + ResNet-18), loss function, data augmentation and evaluation protocol were held constant across all conditions; the label fraction was varied over six levels between 1% and 100%, and each level was trained with two initialisations. An identical gradient-step budget (2,500

updates) was given to every condition, so that the effect of data quantity was measured under a protocol equalised for training duration and computational cost. Performance saturates early: 57 annotated tiles (1%) are sufficient to reach 79.81% IoU, 97.0% of the performance obtained with the full label budget (82.31%), and the twentyfold increase from 287 to 5,732 tiles adds only 0.84 points. Under an equal compute budget the gain from ImageNet pretraining remained between 4.00 and 4.46 points across a hundredfold range of label budgets, averaging 4.20 points; because the two curves do not overlap at any performance level, no data equivalence coefficient could be computed, and the pretrained model trained on 57 tiles outperformed the from-scratch model trained on 5,732 tiles by 1.83 points. The limitations identified in this study point directly to future work.

The first step is to train both arms to convergence with their own learning rate settings, establishing whether the difference between them is asymptotic or budget-dependent. Such high label efficiency brings the question of which tiles to annotate to the fore, and it is worth investigating whether active learning or diversity-based sampling strategies can reach higher performance at the same budget. Testing representations pretrained on remote sensing imagery rather than ImageNet, and reproducing the findings on datasets from different geographies, are the natural extensions of the study.

Acknowledgment
The authors declare that no specific funding was received for this research.

REFERENCES

[1] M. Buyukdemircioglu, R. Can, S. Kocaman, and M. Kada, "DEEP LEARNING BASED BUILDING FOOTPRINT EXTRACTION FROM VERY HIGH RESOLUTION TRUE ORTHOPHOTOS AND NDSM," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. V-2–2022, pp. 211–218, May 2022, doi: 10.5194/isprs-annals-V-2-2022-211-2022.

[2] E. Sumer and M. Turker, "An adaptive fuzzy-genetic algorithm approach for building detection using high-resolution satellite images," *Computers, Environment and Urban Systems*, vol. 39, pp. 48–62, May 2013, doi: 10.1016/j.compenvurbsys.2013.01.004.

[3] D. ARIKAN İSPİR and F. YILDIZ, "Using deep learning algorithms for built-up area extraction from high-resolution GÖKTÜRK-1 satellite imagery," *Earth Sci Inform*, vol. 18, no. 1, p. 21, Dec. 2024, doi: 10.1007/s12145-024-01518-w.

[4] S. Ji, S. Wei, and M. Lu, "Fully Convolutional Networks for Multisource Building Extraction From an Open Aerial and Satellite Imagery Data Set," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 1, pp. 574–586, Jan. 2019, doi: 10.1109/TGRS.2018.2858817.

[5] H. Liu *et al.*, "DE-Net: Deep Encoding Network for Building Extraction from High-Resolution Remote Sensing Imagery," *Remote Sensing*, vol. 11, no. 20, p. 2380, Jan. 2019, doi: 10.3390/rs11202380.

[6] H. Wang and F. Miao, "Building extraction from remote sensing images using deep residual U-Net," *European Journal of Remote Sensing*, vol. 55, no. 1, pp. 71–85, Dec. 2022, doi: 10.1080/22797254.2021.2018944.

[7] G. Yang, Q. Zhang, and G. Zhang, "EANet: Edge-Aware Network for the Extraction of Buildings from Aerial Images," *Remote Sensing*, vol. 12, no. 13, p. 2161, Jan. 2020, doi: 10.3390/rs12132161.

[8] J. Chang, Y. Cen, and G. Cen, "Asymmetric Network Combining CNN and Transformer for Building Extraction from Remote Sensing Images," *Sensors*, vol. 24, no. 19, p. 6198, Jan. 2024, doi: 10.3390/s24196198.

[9] D. Zheng *et al.*, "Weakly Supervised Building Extraction From High-Resolution Remote Sensing Images Based on Building-Aware Clustering and Activation Refinement Network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–15, 2024, doi: 10.1109/TGRS.2024.3438248.

[10] Y. Hua, D. Marcos, L. Mou, X. X. Zhu, and D. Tuia, "Semantic Segmentation of Remote Sensing Images With Sparse Annotations," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2021.3051053.

[11] L. Xia, X. Zhang, J. Zhang, H. Yang, and T. Chen, "Building Extraction from Very-High-Resolution Remote Sensing Images Using Semi-Supervised Semantic Edge Detection," *Remote Sensing*, vol. 13, no. 11, p. 2187, Jan. 2021, doi: 10.3390/rs13112187.

[12] B. Neupane, J. Aryal, and A. Rajabifard, "BUILDING FOOTPRINT SEGMENTATION USING TRANSFER LEARNING: A CASE STUDY OF THE CITY OF MELBOURNE," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. X-4-W3-2022, pp. 173–179, Oct. 2022, doi: 10.5194/isprs-annals-X-4-W3-2022-173-2022.

[13] A. Vats *et al.*, "Terrain-Informed Self-Supervised Learning: Enhancing Building Footprint Extraction From LiDAR Data With Limited Annotations," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 62, pp. 1–10, 2024, doi: 10.1109/TGRS.2024.3391391.

[14] P. Liu *et al.*, "A Systematic Study on Pretraining Strategies for Low-Label Remote Sensing Image Semantic Segmentation," *Sensors*, vol. 26, no. 4, p. 1385, Jan. 2026, doi: 10.3390/s26041385.

[15] B. Zoph *et al.*, “Rethinking Pre-training and Self-training,” Nov. 15, 2020, *arXiv*: arXiv:2006.06882. doi: 10.48550/arXiv.2006.06882.
[16] S. Kornblith, J. Shlens, and Q. V. Le, “Do Better ImageNet Models Transfer Better?,” Jun. 17, 2019, *arXiv*: arXiv:1805.08974. doi: 10.48550/arXiv.1805.08974.
[17] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, “Feature Pyramid Networks for Object Detection,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul. 2017, pp. 936–944. doi: 10.1109/CVPR.2017.106.
[18] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778. Accessed: Aug. 13, 2026. [Online]. Available: https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Türkiye'de Elektrik Enerjisi Göstergelerinin Holt-D Modeliyle Tahmini

Agah Başdeğirmen [1] ve Ali Şimşek [2*]

[1] *Yönetim ve Organizasyon Bölümü / Isparta Meslek Yüksekokulu, Isparta Uygulamalı Bilimler Üniversitesi, Türkiye*
[2] *Yönetim ve Organizasyon Bölümü / Isparta Meslek Yüksekokulu, Isparta Uygulamalı Bilimler Üniversitesi, Türkiye*

[*] *(alisimsek@isparta.edu.tr)*

Özet - Artan enerji talebi, elektrik sisteminin üretim, tüketim ve kapasite göstergelerine ilişkin güvenilir öngörülerin uzun dönemli enerji planlamasındaki önemini artırmaktadır. Bu çalışmanın amacı, Türkiye'de elektrik tüketimi, elektrik üretimi, kişi başına toplam elektrik tüketimi ve kurulu güç göstergelerini bütüncül olarak değerlendirmektir. Nicel araştırma yönteminin benimsendiği bu çalışmada, Türkiye İstatistik Kurumunun enerji istatistiklerinden alınan 2000-2024 dönemine ait yıllık ikincil veriler kullanılmıştır. Dört gösterge serisi Holt-D modeliyle analiz edilmiş, model performansı kare hatalar toplamı, ortalama mutlak hata, kök ortalama kare hata, ortalama mutlak yüzde hata ve ortalama mutlak ölçeklendirilmiş hata ölçütleriyle değerlendirilmiştir. Geçmiş dönem incelemesinde en yüksek yıllık ortalama bileşik büyüme %6,23 ile kurulu güçte belirlenmiştir. Modellerin ortalama mutlak yüzde hata değerleri %2,21 ile %3,76 arasında değişmiş, ortalama mutlak ölçeklendirilmiş hata değerlerinin tamamı 1'in altında kalmıştır. En güçlü tahmin performansı %2,21 ortalama mutlak yüzde hata ve 0,338 ortalama mutlak ölçeklendirilmiş hata ile kurulu güç modelinde elde edilmiştir. Tahminlere göre 2025-2030 döneminde toplam elektrik tüketiminin %13,9 artarak 362,2 milyon megavatsaate (MWh), elektrik üretiminin %12,3 artarak 405,7 milyon MWh, kişi başına tüketimin %11,7 artarak 4.126,1 kilovatsaate (kWh) ve kurulu gücün %20,6 artarak 146.683,2 megavata (MW) ulaşması beklenmektedir. Elde edilen bulgular, incelenen tüm göstergelerde artış beklendiğini ortaya koymaktadır. Ayrıca kurulu kapasitenin diğer göstergelere kıyasla daha hızlı genişleyeceği öngörülmektedir.

Anahtar Kelimeler - Elektrik Tüketimi, Elektrik Üretimi, Kişi Başına Toplam Elektrik Tüketimi, Kurulu Güç, Holt-D Modeli

I. GİRİŞ

Islatince ve Haydaroğlu (2015) ile Albostan vd. (2013), dünya genelinde artan tüketimin petrol başta olmak üzere fosil enerji kaynaklarının geleceğini tehdit ettiğini ve ekonomik ilerleme ile ulusal güvenliğin sürdürülebilmesi için gerekli olan kesintisiz ve güvenilir enerji arzının önemini artırdığını belirtmektedir. Artan enerji talebi ve fosil yakıtların çevresel etkileri karşısında Türkiye'nin enerji tasarrufuna ve yenilenebilir enerji kaynaklarına yönelmesi ve dışa bağımlılığını azaltmak amacıyla özellikle güneş enerjisi yatırımlarını ve teşvik mekanizmalarını güçlendirmesi gerekmektedir (İraz vd., 2014). Geleceğe yönelik talep tahminleri, ihtiyaç duyulacak ham madde miktarını, bulundurulması gereken bitmiş ürün stok düzeyini, sevk edilecek ürün sayısını, işe alınacak personel sayısını, inşa edilecek tesis sayısını ve hatta satın alınması gereken ofis malzemelerinin miktarını belirlemektedir (Chase, 2013). Bu çalışmada, elektrik enerjisi sisteminin kapasite, üretim ve tüketim boyutlarını bütüncül olarak değerlendirmek amacıyla elektrik tüketimi, elektrik üretimi, kişi başına toplam elektrik tüketimi ve kurulu güç değişkenleri tahmin edilmiştir.

Kurulu güç, bir jeneratörün, ana tahrik ünitesinin veya diğer elektrik üretim ekipmanlarının üretici tarafından belirlenen koşullarda sağlayabileceği en yüksek anma gücü olarak tanımlanmaktadır. Bu güç genellikle megavat (MW) cinsinden ifade edilmekte ve jeneratör üzerinde bulunan bilgi etiketinde belirtilmektedir (EIA, 2026a).

Brüt üretim, elektrik üretim birimlerince üretilen ve jeneratör çıkış terminalinde kilovatsaat (kWh) veya megavatsaat (MWh) cinsinden ölçülen toplam elektrik enerjisi miktarı olarak tanımlanmaktadır (EIA, 2026b). Net üretim, brüt üretim miktarından üretim tesislerindeki hizmetler veya yardımcı sistemler için tüketilen elektrik enerjisinin çıkarılmasıyla elde edilen elektrik enerjisi miktarı olarak tanımlanmaktadır. Pompaj depolamalı hidroelektrik santrallerinde suyun pompalanması amacıyla kullanılan elektrik enerjisi, tesis içi tüketim olarak değerlendirilerek brüt üretimden düşülmektedir (U.S. Energy Information Administration [EIA], 2026c). Kişi başına elektrik enerjisi tüketimi (kWh/kişi), elektrik santralleri ile birleşik ısı ve güç santrallerinin toplam üretiminden iletim, dağıtım ve dönüşüm kayıpları ile santrallerin kendi tüketimlerinin çıkarılmasıyla elde edilen elektrik enerjisi miktarının nüfusa bölünmesi sonucunda hesaplanan gösterge olarak tanımlanmaktadır (World Bank, 2026). Tablo 1'de, literatürde Holt-D modeli ve ilişkili üstel düzeltme yaklaşımlarının kullanıldığı Soysal ve Ömürgönülşen (2010), Yürekli vd. (2017), Tüzemen (2020), Parıldar vd. (2024), Söylemez vd. (2024), Aslan ve Bayraktutan (2025), Bolakar Tosun (2025) ve Ekin (2026) tarafından gerçekleştirilen araştırmalar özetlenmektedir.

Tablo 1. Holt-D Modeli ve İlişkili Tahmin Yöntemlerinin Kullanıldığı Çalışmalara İlişkin Literatür Özeti

Yıl	Yazar Soyadı	Çalışmanın Amacı ve Sonucu
2010	Soysal & Ömürgönülşen	Winter ve diğer zaman serisi yöntemleriyle turist sayısını tahmin etmektir..
		Winter yöntemiyle turist sayısının Haziran'a doğru arttığı tahmin edilmiştir
2017	Yürekli vd.	FAO Penman-Monteith ve Holt yaklaşımıyla referans bitki su tüketimini incelemektir.
		Referans bitki su tüketiminde istasyonlar arasında farklı eğilimler saptanmıştır.
2020	Tüzemen	ARIMA ve üstel düzeltme yöntemleriyle Cumhuriyet altını fiyatlarını tahmin etmektir.
		ARIMA tahminlerinde Cumhuriyet altını fiyatlarının sürekli artacağı öngörülmüştür.
2024	Parıldar vd.	Holt-Winter yöntemiyle yoğun bakım hemşire talebini tahmin etmektir.
		Tahmin edilen hemşire talebinin on iki hemşireyle karşılanabildiği belirlenmiştir.
2024	Söylemez vd.	Üstel düzeltme, Holt ve ARIMA yöntemleriyle tahıl üretimini tahmin etmektir.
		Buğday, arpa, mısır ve yulaf üretiminde artış öngörülmüştür.
2025	Aslan & Bayraktutan	Holt üstel düzeltme yöntemiyle Kocaeli turizm göstergelerini tahmin etmektir.
		Kocaeli'nin toplam turist sayısının 2040'a kadar artacağı öngörülmüştür.
2025	Bolakar Tosun	Holt-Winters yöntemiyle trafik kazaları ve araç sayısını tahmin etmektir.
		Trafik kazaları ve araç sayısının 2026'ya kadar artacağı öngörülmüştür.
2026	Ekin	Üstel düzeltme ve Holt-Winters yöntemleriyle arızalı parça sayısını tahmin etmektir.
		Arızalı parça sayıları Basit Üstel Düzgünleştirmeyle en başarılı tahmin edilmiştir.

Kaynak: Yazarlar tarafından oluşturulmuştur.

Bu çalışmada Holt-D modeline dayalı zaman serisi tahmini yöntemi kullanılmıştır. Holt-D modeli, düzensiz eğilimleri sönümlemek üzere geliştirilmiş üstel düzeltme temelli bir tahmin modelidir (Gardner & McKenzie, 1985). Veri toplama tekniği olarak ikincil veri kullanılmıştır. Bu teknik, önceden derlenmiş verilerin araştırma amacıyla sistematik biçimde kullanılmasını kapsamaktadır (Johnston, 2014). Veri analizi, dört gösterge serisine Holt-D modelinin uygulanmasını ve model başarısının kare hatalar toplamı (Sum of Squared Errors, SSE), ortalama mutlak hata (Mean Absolute Error, MAE), kök ortalama kare hata (Root Mean Squared Error, RMSE), ortalama mutlak yüzde hata (Mean Absolute Percentage Error, MAPE) ve ortalama mutlak ölçeklendirilmiş hata (Mean Absolute Scaled Error, MASE) ölçütleri üzerinden değerlendirilmesini içermektedir (Hyndman & Koehler, 2006).

Araştırmanın temel amacı, Türkiye elektrik enerjisi sisteminin kapasite, üretim ve tüketim boyutlarını bütüncül olarak değerlendirmek üzere elektrik tüketimi, elektrik üretimi, kişi başına toplam elektrik tüketimi ve kurulu güç göstergelerine ilişkin tahminler üretmektir. Veri seti 2000-2024 dönemini kapsamaktadır. Veriler, Türkiye İstatistik Kurumunun (TÜİK) Enerji İstatistiklerinden elde edilmiştir. Verilerin analizi ve tahminlerin üretilmesi Microsoft Excel 365 kullanılarak gerçekleştirilmiştir.

Araştırmanın önemi, Türkiye'deki elektrik göstergelerinin öngörülmesi yoluyla uzun dönemli enerji planlamasına temel bilgi sağlamasıdır.

Bu çalışma beş bölümden oluşmaktadır. Giriş bölümünde enerji arzı ve talep tahmininin önemi açıklanmakta, temel elektrik enerjisi göstergeleri tanımlanmakta, ilgili literatür özetlenmekte ve araştırmanın amacı, kapsamı, veri kaynağı ile yöntemi sunulmaktadır. İkinci bölümde Holt-D modelinin kuramsal yapısı, düzey, trend ve sönümleme parametreleri, tahmin denklemleri ve model performansının değerlendirilmesinde kullanılan hata ölçütleri açıklanmaktadır. Üçüncü bölümde Türkiye'nin 2000-2024 dönemine ait elektrik tüketimi, elektrik üretimi, kişi başına elektrik tüketimi ve kurulu güç verileri tanıtılmakta ve bu göstergelerin dönem içindeki değişimleri incelenmektedir. Dördüncü bölümde Holt-D modelinin uygulama sonuçları sunulmakta, model parametreleri ve tahmin performansı değerlendirilmekte ve enerji göstergelerine ilişkin 2025-2030 dönemi tahminleri yorumlanmaktadır. Beşinci ve son bölümde ise araştırmadan elde edilen bulgular genel olarak değerlendirilmektedir.

II. HOLT-D MODELİ

Gardner ve McKenzie (1985), doğrusal bir trendin üstel olarak düzeltilmesine yönelik standart Holt (1960) formülasyonunu esas almaktadır. Bu formülasyon Winters (1960) ve Holt (2004) çalışmalarına dayanmaktadır. Serinin yerel düzeyi Denklem (1), trend bileşeni Denklem (2) ve t tahmin başlangıç noktasından m dönem sonrası için hesaplanan tahmin değeri Denklem (3) ile ifade edilmektedir.

$$S_t = \alpha X_t + (1-\alpha)(S_{t-1} + T_{t-1}) \quad \textbf{(1)}$$

$$T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)T_{t-1} \quad \textbf{(2)}$$

$$\hat{X}_t(m) = S_t + mT_t \quad \textbf{(3)}$$

Burada S_t, serinin t dönemindeki yerel düzeyini, T_t ise trend bileşenini göstermektedir. $\hat{X}_t(m)$, t tahmin başlangıç noktasından itibaren m dönem sonrası için hesaplanan tahmin değeridir. α ve γ düzeltme parametreleri, model daha geniş bir aralıkta kararlı olmasına rağmen uygulamada genellikle 0 ile 1 arasında sınırlandırılmaktadır. Trend tahmini, φ ile gösterilen otoregresif sönümleme parametresi kullanılarak değiştirilebilir. Bu düzenleme sonucunda elde edilen Holt-D modelinde düzey bileşeni Denklem (4), trend bileşeni Denklem (5) ve m dönem sonrası tahmin değeri Denklem (6) ile hesaplanmaktadır.

$$S_t = \alpha X_t + (1-\alpha)(S_{t-1} + \phi T_{t-1}) \quad \textbf{(4)}$$

$$T_t = \gamma(S_t - S_{t-1}) + (1-\gamma)\phi T_{t-1} \quad \textbf{(5)}$$

$$\hat{X}_t(m) = S_t + \sum_{i=1}^{m} \phi^i T_t \quad \textbf{(6)}$$

Denklem (6)'daki toplam ifadesinin açık aritmetik yazımı Denklem (7)'de gösterilmiştir.

$$\hat{X}_t(m) = S_t + \phi T_t + \phi^2 T_t + \phi^3 T_t + \cdots + \phi^m T_t \quad \textbf{(7)}$$

Holt-D modeli, φ sönüm parametresinin değerine bağlı olarak trendin gelecekteki davranışına ilişkin dört farklı durumu kapsamaktadır. $\varphi = 0$ olduğunda tahminlerde trend bulunmaz ve model basit üstel düzeltme modeline eşdeğer hâle gelir. $0 < \varphi < 1$ olduğunda trend sönümlenir ve tahminler Denklem (8)'de gösterilen asimptotik düzeye yaklaşır.

$$\hat{X}_t(\infty) = S_t + \frac{\phi T_t}{1-\phi} \quad \textbf{(8)}$$

$\varphi = 1$ olduğunda Holt-D modeli standart Holt modeline eşdeğer olmakta ve trend doğrusal biçimde devam etmektedir. Son olarak $\varphi > 1$ olduğunda trend üstel bir yapı göstermektedir. Bu durum, otomatik tahmin modellerinde oldukça riskli bir seçenek olarak değerlendirilmektedir. Birçok tahmin sisteminde aynı üstel düzeltme modelinin bütün zaman serilerine uygulanması yaygındır. Her tahmin döneminde binlerce seriye ilişkin tahminlerin oluşturulması gereken büyük ölçekli stok sistemlerinde bu uygulama çoğu zaman kaçınılmazdır. Holt-D modelindeki her parametrenin 0 ile 1 aralığında belirlenmesi durumunda modelin ortalama tahmin doğruluğunun standart Holt modelinden daha yüksek olması beklenmektedir. Holt-D modelinin otomatik tahmin yöntemi olarak kullanılmasına ilişkin gerekçe, Lewandowski yaklaşımının gerekçesiyle benzerlik göstermektedir. Verilerde güçlü ve kalıcı bir trend

bulunması durumunda φ parametresinin 1'e yakın bir değer alması beklenmektedir. Bu durumda Holt-D modeliyle elde edilen tahminler standart Holt modelinin tahminlerine oldukça yakın olmaktadır. Verilerin aşırı derecede gürültülü olması veya trendin düzensiz bir yapı göstermesi durumunda ise φ < 1 seçilerek trendin zaman içinde sönümlenmesi sağlanmaktadır.

Denklem (4)-(6) ile verilen Holt-D modeli, işlemsel açıdan karmaşık görünmekle birlikte önemli ölçüde sadeleştirilebilir. Bir dönem ileriye yönelik tahmin hatası e_t, Denklem (9)'da tanımlanmaktadır.

$$e_t = X_t - \hat{X}_{t-1}(1) = X_t - S_{t-1} - \phi T_{t-1} \quad \textbf{(9)}$$

Bir dönem ileriye yönelik bu tahmin hatası kullanıldığında Denklem (4) ve Denklem (5), sırasıyla Denklem (10) ve Denklem (11) biçiminde yeniden yazılabilir.

$$S_t = \hat{X}_{t-1}(1) + h_1 e_t \quad \textbf{(10)}$$

$$T_t = \phi T_{t-1} + h_2 e_t \quad \textbf{(11)}$$

Denklem (10) ve Denklem (11)'de kullanılan h_1 ve h_2 katsayıları sırasıyla Denklem (12) ve Denklem (13)'te tanımlanmaktadır.

$$h_1 = \alpha \quad \textbf{(12)}$$

$$h_2 = \alpha\gamma \quad \textbf{(13)}$$

Modelin düzey ve trend bileşenlerinin birbirinden ayrılmasıyla daha ileri bir sadeleştirme gerçekleştirilebilir. İlk olarak A_t ve B_t bileşenleri sırasıyla Denklem (14) ve Denklem (15)'te tanımlanmaktadır.

$$A_t = S_t + \frac{\phi T_t}{1-\phi} \quad \textbf{(14)}$$

$$B_t = -\frac{\phi T_t}{1-\phi} \quad \textbf{(15)}$$

Burada A_t, tahminlerin uzun dönemde yaklaştığı asimptotik düzeyi göstermektedir. B_t bileşeninin işareti ise Denklem (11)'deki trend bileşeninin işaretinin tersidir. Bu bileşenler Denklem (16) ve Denklem (17) kullanılarak düzeltilmekte, m dönem sonrasına ilişkin tahmin ise Denklem (18) ile hesaplanmaktadır.

$$A_t = A_{t-1} + g_1 e_t \quad \textbf{(16)}$$

$$B_t = \phi B_{t-1} + g_2 e_t \quad \textbf{(17)}$$

$$\hat{X}_t(m) = A_t + \phi^m B_t \quad \textbf{(18)}$$

φ < 1 olduğunda Denklem (16)-(18) ile elde edilen tahminler, Denklem (6), Denklem (10) ve Denklem (11) ile elde edilen tahminlerle aynıdır. Bu eşdeğerliğin sağlanması için g_1 ve g_2 parametreleri sırasıyla Denklem (19) ve Denklem (20)'de tanımlanmaktadır.

$$g_1 = \alpha + \frac{\alpha\gamma\phi}{1-\phi} \quad \textbf{(19)}$$

$$g_2 = -\frac{\alpha\gamma\phi}{1-\phi} \quad \textbf{(20)}$$

Holt-D modelinin Denklem (16)-(18) ile verilen bu biçiminin temel avantajı, Denklem (6)'daki toplam işlemini ortadan kaldırmasıdır. Ayrıca bu gösterim, tahminlerin uzun dönemde yaklaşacağı asimptotik düzeyin doğrudan hesaplanmasını sağlamaktadır. Bununla birlikte φ = 1 olduğunda ortaya çıkan standart doğrusal Holt modeli için ayrı bir hesaplama düzenlemesinin yapılması gerekmektedir.

III. VERİ SETİ

Türkiye'de 2000-2024 yılları arasında elektrik tüketimi, elektrik üretimi, kişi başına toplam elektrik tüketimi ve kurulu güç göstergelerinin yıllar itibarıyla değişimi Tablo 2'de sunulmuştur.

Tablo 2. Türkiye'de Elektrik Tüketimi, Kişi Başına Elektrik Tüketimi, Elektrik Üretimi ve Kurulu Güç Göstergeleri

Yıllar	Elektrik Tüketimi (Mwh)	Elektrik Üretimi (Mwh)	Kişi Başına Toplam Elektrik Tüketimi (Kwh)	Kurulu Güç (Mw)
2000	97159804,0	124921600,0	1501,0	27264,1
2001	96365767,0	122724700,0	1469,0	28332,4
2002	102242104,0	129399500,0	1540,0	31845,8
2003	110988575,0	140580500,0	1652,0	35587,0
2004	119971414,0	150698300,0	1764,0	36824,0
2005	129002340,0	161956200,0	1873,0	38843,5
2006	141438534,0	176299800,0	2028,0	40564,8
2007	154852338,0	191558129,4	2194,0	40835,7
2008	161947528,0	198418000,0	2264,0	41817,2
2009	156894070,0	194812926,0	2162,0	44761,2
2010	172050628,0	211207700,0	2334,0	49524,1
2011	186099551,0	229395100,0	2490,0	52911,1
2012	194923349,0	239496800,0	2577,0	57059,4
2013	198045181,0	240153953,0	2583,0	64007,5
2014	207375078,0	251962817,0	2669,0	69519,8
2015	217312250,0	261783303,5	2760,0	73146,7
2016	231203746,0	274407749,0	2897,0	78497,4
2017	249022646,0	297277522,9	3082,0	85200,0
2018	258232177,0	304801885,0	3149,0	88550,8
2019	257273130,0	303897559,6	3094,0	91267,0
2020	262702129,0	306703092,4	3142,0	95890,6
2021	286691464,0	334723100,0	3386,0	99819,6
2022	286576318,0	328379300,0	3360,0	103809,3
2023	289371741,0	331148900,0	3390,0	110914,0
2024	308673931,0	354570200,0	3603,0	116265,1

Kaynak: Türkiye İstatistik Kurumu. (2026). Enerji istatistikleri. Merkezi Dağıtım Sistemi (MEDAS). 2 Temmuz 2026 tarihinde https://biruni.tuik.gov.tr/medas/?locale=tr

Tablo 2'deki veriler yıllık değişim oranları açısından değerlendirildiğinde, Türkiye'de elektrik tüketiminin 2000-2024 döneminde genel olarak arttığı, ancak artış hızının yıllara göre farklılaştığı görülmektedir. Elektrik tüketimi 2000'de 97,2 milyon MWh iken 2024'te 308,7 milyon MWh'ye ulaşarak yaklaşık %217,7 artmıştır. Dönemin yıllık ortalama bileşik büyüme oranı %4,93'tür. En yüksek artışlar 2010'da %9,66, 2006'da %9,64, 2007'de %9,48 ve 2021'de %9,13 olarak gerçekleşmiştir. Buna karşılık tüketim 2001, 2009, 2019 ve 2022'de azalmış, en yüksek düşüş %3,12 ile 2009'da görülmüştür. 2023'te %0,98 olan artışın 2024'te %6,67'ye yükselmesi dikkat çekmektedir. Elektrik üretimi, tüketimle genel olarak paralel bir seyir izlemiştir. Üretim 2000-2024 döneminde yaklaşık %183,83 artarken yıllık ortalama bileşik büyüme oranı %4,44 olmuştur. Üretim 2021'de %9,14 artmış, 2022'de %1,90 azalmış ve 2024'te %7,07 yükselmiştir. Kişi başına toplam elektrik tüketimi ise 1.501 kWh'den 3.603 kWh'ye çıkarak yaklaşık %140,04 artmış, yıllık ortalama bileşik büyüme oranı %3,72 olarak gerçekleşmiştir. Kişi başına tüketimin 2021'de %7,77 artması, 2022'de %0,77 azalması, 2023'te sınırlı toparlanması ve 2024'te %6,28 yükselmesi dalgalı bir seyre işaret etmektedir.

İncelenen göstergeler arasında en yüksek uzun dönemli büyüme kurulu güçte gerçekleşmiştir. Türkiye'nin kurulu gücü 2000'de 27.264,1 MW iken 2024'te 116.265,1 MW'ye ulaşarak yaklaşık %326,44 artmıştır. Yıllık ortalama bileşik büyüme oranı %6,23 ile elektrik tüketimi, elektrik üretimi ve kişi başına tüketimin üzerinde gerçekleşmiştir. Kurulu güçte 2002'de %12,40, 2010'da %10,64

ve 2013'te %12,18 oranında yüksek artışlar görülmüştür. Artış oranı 2023'te %6,84, 2024'te ise %4,82 olmuştur. Bu bulgular, elektrik üretim kapasitesinin tüketim ve fiilî üretimden daha hızlı büyüdüğünü göstermektedir.

IV. HOLT-D MODELİ UYGULAMA SONUÇLARI

Tablo 3'te, elektrik enerjisi göstergeleri için oluşturulan tahmin modellerine ait parametre değerleri ile SSE, MAE, RMSE, MAPE ve MASE tahmin performans ölçütlerine ilişkin bilgiler gösterilmektedir.

Tablo 3. Elektrik Enerjisi Göstergelerine İlişkin Model Parametreleri ve Tahmin Performans Ölçütleri

Göstergeler	α	γ	φ	SSE	MAE	RMSE	MAPE	MASE
Toplam Elektrik Tüketimi	1,00	0,23	0,98	1.481.836.940.046.030,0	6.572.340,8	7.857.684,5	%3,50	0,700
Elektrik Üretimi	0,84	0,27	0,98	2.400.968.301.507.450,0	8.562.774,7	10.002.017,1	%3,76	0,804
Kişi Başına Elektrik Tüketimi	1,00	0,23	0,98	237.950,7	83,0	99,6	%3,41	0,787
Kurulu Güç	1,00	0,55	0,98	64.605.520,6	1.252,4	1.640,7	%2,21	0,338

Kaynak: Yazarlar tarafından oluşturulmuştur.

Tablo 3'te yer alan performans ölçütleri birlikte değerlendirildiğinde, dört göstergeye ilişkin tahmin modellerinin genel olarak yüksek tahmin başarısı gösterdiği söylenebilir. **MAPE** açısından Lewis (1982) tarafından önerilen ve literatürde yaygın olarak kullanılan sınıflandırmaya göre, %10'un altındaki değerler çok yüksek tahmin doğruluğuna, %10-20 arasındaki değerler iyi, %20-50 arasındaki değerler kabul edilebilir, %50'nin üzerindeki değerler ise düşük tahmin doğruluğuna işaret etmektedir. Bu ölçüte göre tüm serilerin MAPE değerleri %10'un oldukça altında kalmaktadır. Dolayısıyla dört göstergenin de tahmin doğruluğu oldukça yüksektir. **MASE** ölçütünde temel referans değeri 1'dir. MASE'nin 1'den küçük olması kullanılan tahmin modelinin naif tahmin yöntemine göre daha başarılı olduğunu, 1'e eşit olması benzer performans gösterdiğini, 1'den büyük olması ise naif yönteme göre daha düşük tahmin performansı sergilediğini ifade etmektedir (Hyndman & Koehler, 2006). Toplam elektrik tüketimi için MASE 0,700, elektrik üretimi için 0,804, kişi başına elektrik tüketimi için 0,787 ve kurulu güç için 0,338 olarak belirlenmiştir. Bu sonuç, kullanılan tahmin modellerinin naif tahmin yöntemine göre daha başarılı bir tahmin performansı sergilediğini göstermektedir. **MAE**, tahmin edilen değerlerle gerçekleşen değerler arasındaki mutlak farkların ortalamasını göstermektedir. Bu nedenle MAE değerinin küçülmesi, tahmin hatalarının azalmasına ve tahmin performansının iyileşmesine işaret etmektedir (Willmott & Matsuura, 2005). Toplam elektrik tüketiminde MAE yaklaşık 6,57 milyon MWh, elektrik üretiminde 8,56 milyon MWh, kişi başına elektrik tüketiminde 83 kWh ve kurulu güçte yaklaşık 1.252 MW düzeyindedir. **RMSE** de benzer biçimde tahmin hatalarının büyüklüğünü ölçmekle birlikte hataların karelerine dayandığından büyük tahmin hatalarına MAE'ye göre daha fazla ağırlık vermektedir (Chai & Draxler, 2014). Toplam elektrik tüketiminin RMSE değeri yaklaşık 7,86 milyon MWh, elektrik üretiminin 10,00 milyon MWh, kişi başına elektrik tüketiminin 99,6 kWh ve kurulu gücün 1.640,7 MW düzeyindedir.

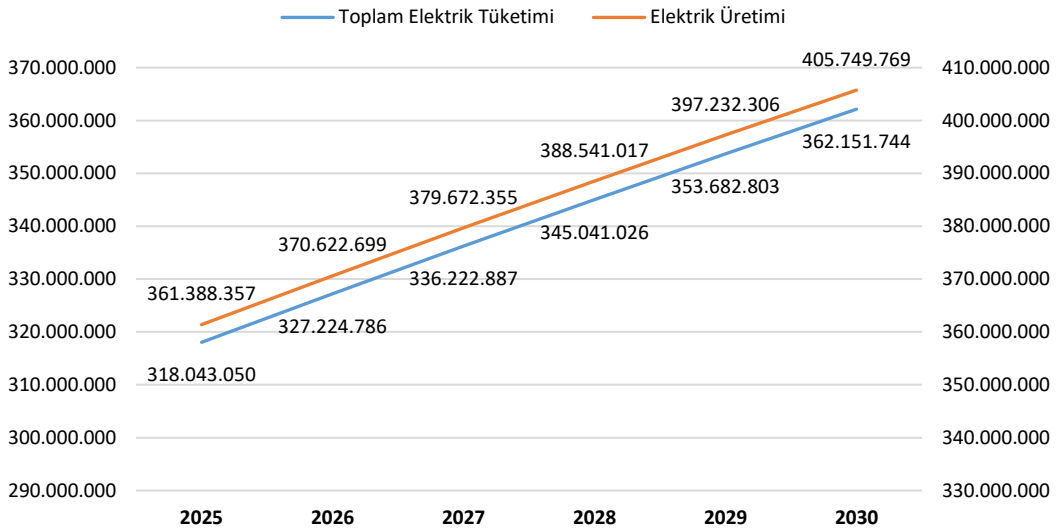
SSE, tahmin hatalarının kareleri toplamını göstermekte ve daha düşük değerler daha iyi uyuma işaret etmektedir. Ancak SSE doğrudan serinin büyüklüğünden ve gözlem sayısından etkilendiğinden elektrik tüketimi ile kişi başına tüketim gibi farklı ölçeklerdeki serilerin SSE değerlerini birbiriyle karşılaştırmak doğru değildir. Bu nedenle SSE daha çok aynı seri için farklı modeller arasından en uygun modelin seçilmesinde kullanılan bir ölçüttür.

Model parametreleri de sonuçların yorumlanmasında önemlidir. Düzey yumuşatma parametresi olan α'nın 1'e yakın değerler alması, modelin serideki son gözlemlere yüksek ağırlık verdiğini göstermektedir. Toplam elektrik tüketimi, kişi başına elektrik tüketimi ve kurulu güç serilerinde α değerinin 1,00 olması, bu serilerde güncel gözlemlerin tahmin düzeyinin belirlenmesinde oldukça etkili olduğunu göstermektedir. Elektrik üretiminde α değerinin 0,84 olması da yeni gözlemlere yüksek ağırlık verildiğini, ancak geçmiş bilginin diğer serilere kıyasla bir miktar daha fazla korunduğunu göstermektedir.

Tablodaki γ parametresi, trend bileşeninin güncellenme hızını ifade etmektedir. Toplam elektrik tüketimi ve kişi başına elektrik tüketiminde γ yaklaşık 0,23, elektrik üretiminde 0,27, kurulu güçte ise

0,55 düzeyindedir. Kurulu güçteki daha yüksek γ değeri, trend bileşeninin yeni değişimlere diğer serilere göre daha hızlı uyum sağladığını göstermektedir. Son olarak φ = 0,98 değeri, modelde trendin sönümlendirilerek geleceğe taşındığını ve sönümlendirmenin oldukça sınırlı olduğunu göstermektedir. Değerin 1'e çok yakın olması, trend etkisinin tahmin ufku boyunca büyük ölçüde korunduğuna işaret etmektedir.

Genel olarak değerlendirildiğinde, MAPE değerlerinin %2,21-%3,76 aralığında ve MASE değerlerinin tamamının 1'in altında olması, kullanılan modellerin hem mutlak yüzde hata bakımından yüksek doğruluğa sahip olduğunu hem de naif tahmin yöntemine göre daha başarılı sonuçlar verdiğini ortaya koymaktadır. Performans göstergeleri içinde özellikle kurulu güç modeli, %2,21 MAPE ve 0,338 MASE değerleriyle en güçlü tahmin performansını göstermektedir. Elektrik üretimi modeli ise dört seri içinde en yüksek MAPE ve MASE değerlerine sahip olmakla birlikte bu değerler yine de modelin tahmin performansının oldukça başarılı olduğunu göstermektedir.



Şekil 1. Türkiye'de Toplam Elektrik Tüketimi ve Elektrik Üretiminin 2025-2030 Dönemi Tahmin Değerleri

Kaynak: Yazarlar tarafından oluşturulmuştur.

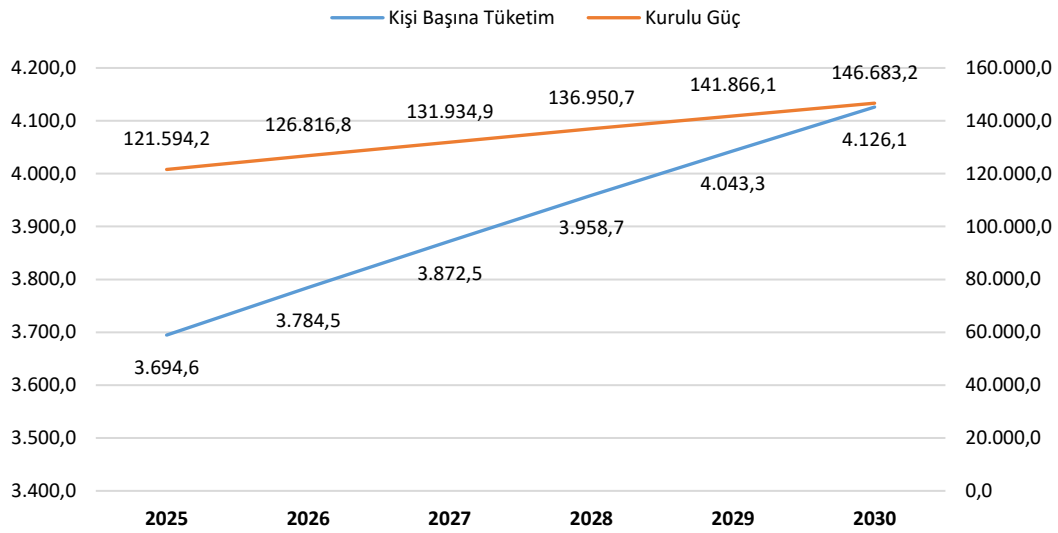
Şekil 1 incelendiğinde, 2025-2030 döneminde hem toplam elektrik tüketiminin hem de elektrik üretiminin düzenli ve kesintisiz bir artış eğilimi göstereceği tahmin edilmektedir. Toplam elektrik tüketiminin 2025 yılında yaklaşık 318,0 milyon MWh düzeyinden 2030 yılında 362,2 milyon MWh düzeyine yükselmesi beklenmektedir. Buna göre tahmin dönemi boyunca toplam elektrik tüketimindeki artış yaklaşık 44,2 milyon MWh, diğer bir ifadeyle yaklaşık %13,9 düzeyindedir. Elektrik üretiminin ise aynı dönemde yaklaşık 361,4 milyon MWh'den 405,7 milyon MWh'ye yükselmesi ve böylece yaklaşık 44,3 milyon MWh artması öngörülmektedir. Bu değişim, üretimde yaklaşık %12,3'lük bir artışa karşılık gelmektedir.

Yıllar itibarıyla değerlendirildiğinde, her iki serinin de oldukça istikrarlı bir artış eğilimi izlediği görülmektedir. Toplam elektrik tüketimi 2026 yılında 327,2 milyon MWh, 2027 yılında 336,2 milyon MWh, 2028 yılında 345,0 milyon MWh ve 2029 yılında 353,7 milyon MWh düzeyine ulaşmaktadır. Elektrik üretimi ise aynı yıllarda sırasıyla 370,6 milyon MWh, 379,7 milyon MWh, 388,5 milyon MWh ve 397,2 milyon MWh olarak tahmin edilmektedir. Bu görünüm, modelin tahmin ufku boyunca ani sıçramalar veya keskin düşüşler yerine istikrarlı ve doğrusal eğilime yakın bir büyüme yapısı öngördüğünü göstermektedir.

Elektrik üretiminin incelenen tüm yıllarda toplam elektrik tüketiminin üzerinde kalacağı da dikkat çekmektedir. Üretim ile tüketim arasındaki farkın 2025 yılında yaklaşık 43,4 milyon MWh, 2030 yılında ise yaklaşık 43,5 milyon MWh düzeyinde gerçekleşmesi beklenmektedir. Dolayısıyla tahmin sonuçları, iki göstergenin birbirine yakın mutlak artışlarla büyümesine karşın üretimin tüketimin üzerinde

seyretmeye devam edeceğine işaret etmektedir. Bununla birlikte bu farkın doğrudan bir "arz fazlası" olarak yorumlanması uygun değildir. Böyle bir değerlendirme için elektrik ihracatı ve ithalatı, iletim ve dağıtım kayıpları, santrallerin iç tüketimleri ve diğer sistem bileşenlerinin ayrıca dikkate alınması gerekmektedir.

Genel olarak Şekil 1, Türkiye'de elektrik talebi ile üretimin 2025-2030 döneminde birlikte artmaya devam edeceğini ve üretim düzeyinin tahmin dönemi boyunca tüketimin üzerinde kalacağını ortaya koymaktadır. Ayrıca tüketimin dönemsel toplam artış oranının üretimden bir miktar yüksek olması, elektrik talebindeki göreli büyümenin üretime kıyasla daha hızlı gerçekleşebileceğini göstermektedir. Bu bulgu, uzun dönemli enerji planlamasında tüketim artışının yakından izlenmesinin önemini ortaya koymaktadır.



Şekil 2. Türkiye'de Kişi Başına Elektrik Tüketimi ve Kurulu Gücün 2025-2030 Dönemi Tahmin Değerleri

Kaynak: Yazarlar tarafından oluşturulmuştur.

Şekil 2 incelendiğinde, 2025-2030 döneminde kişi başına elektrik tüketimi ile kurulu gücün birlikte ve kesintisiz biçimde artacağı, ancak her iki göstergede de yıllık artış hızının zaman içinde sınırlı ölçüde yavaşlayacağı öngörülmektedir.

Kişi başına elektrik tüketiminin 2025 yılında 3.694,6 kWh düzeyinden 2030 yılında 4.126,1 kWh düzeyine yükselmesi beklenmektedir. Böylece tahmin dönemi boyunca kişi başına tüketimde yaklaşık 431,5 kWh, diğer bir ifadeyle %11,7 oranında artış öngörülmektedir. Yıllık değerler 2026'da 3.784,5 kWh, 2027'de 3.872,5 kWh, 2028'de 3.958,7 kWh ve 2029'da 4.043,3 kWh olarak tahmin edilmiştir. Bununla birlikte yıllık mutlak artışın 2025-2026 dönemindeki yaklaşık 89,9 kWh'den 2029-2030 döneminde 82,8 kWh'ye gerilemesi, tüketim düzeyi artmaya devam ederken büyüme hızının kademeli olarak yavaşladığını göstermektedir. Oransal olarak yıllık artışın da yaklaşık %2,4 düzeyinden %2,0 civarına gerilediği görülmektedir.

Kurulu güç tahminleri de benzer biçimde sürekli bir yükselişe işaret etmektedir. Kurulu gücün 2025 yılında 121.594,2 MW düzeyinden 2030 yılında 146.683,2 MW düzeyine ulaşacağı tahmin edilmektedir. Bu durum, beş yıllık tahmin döneminde yaklaşık 25.089 MW'lık kapasite artışına ve yaklaşık %20,6'lık toplam büyümeye karşılık gelmektedir. Kurulu güç 2026 yılında 126.816,8 MW, 2027 yılında 131.934,9 MW, 2028 yılında 136.950,7 MW ve 2029 yılında 141.866,1 MW olarak öngörülmektedir. Bununla birlikte yıllık kapasite artışının yaklaşık 5.223 MW'dan 4.817 MW'a doğru azalması, kurulu güçteki büyümenin devam etmesine rağmen artış hızının giderek daha ılımlı bir seyir izleyeceğine işaret etmektedir.

İki gösterge karşılaştırıldığında, tahmin dönemi boyunca kurulu gücün oransal büyümesinin kişi başına elektrik tüketimindeki büyümeden daha yüksek olduğu görülmektedir. Kişi başına tüketimin 2025-2030 döneminde yaklaşık %11,7 artması beklenirken kurulu güçteki artış yaklaşık %20,6'dır. Bu bulgu, model

sonuçları çerçevesinde elektrik sistemindeki kurulu kapasitenin kişi başına tüketimde beklenen artıştan daha hızlı genişleyeceğini göstermektedir. Ancak bu durum tek başına kapasite fazlası veya arz güvenliği açısından yeterli bir gösterge değildir. Böyle bir değerlendirme için kapasite kullanım oranları, üretim teknolojilerinin bileşimi, yenilenebilir enerji kaynaklarının payı, puant talep ve elektrik üretim miktarlarının da birlikte incelenmesi gerekmektedir.

Genel olarak Şekil 2, 2025-2030 döneminde Türkiye'de hem kişi başına elektrik tüketiminin hem de kurulu gücün istikrarlı biçimde yükselmeye devam edeceğini, buna karşın yıllık büyüme hızlarında kademeli bir yavaşlama yaşanacağını ortaya koymaktadır. Özellikle kurulu gücün daha yüksek oranda büyümesinin öngörülmesi, elektrik üretim kapasitesindeki genişlemenin tahmin dönemi boyunca devam edeceğine ilişkin önemli bir bulgu olarak değerlendirilebilir.

V. Sonuç ve Tartışma

Bu çalışma, Türkiye elektrik enerjisi sisteminin kapasite, üretim ve tüketim boyutlarını bütüncül biçimde değerlendirmek amacıyla elektrik tüketimi, elektrik üretimi, kişi başına toplam elektrik tüketimi ve kurulu güç göstergelerini incelemiştir. Türkiye İstatistik Kurumunun enerji istatistiklerinden elde edilen 2000-2024 dönemine ait yıllık veriler, sönümlü eğilimleri dikkate alan Holt-D modeliyle analiz edilmiş ve 2025-2030 dönemi için tahminler üretilmiştir. Model performansı kare hatalar toplamı, ortalama mutlak hata, kök ortalama kare hata, ortalama mutlak yüzde hata ve ortalama mutlak ölçeklendirilmiş hata ölçütleriyle değerlendirilmiştir. Böylece geçmiş dönemdeki değişimlerin ortaya konulması, tahmin modellerinin başarısının incelenmesi ve gelecekteki enerji göstergelerinin öngörülmesi aynı analiz çerçevesinde bir araya getirilmiştir.

Geçmiş dönem bulguları, incelenen dört göstergenin de 2000-2024 döneminde uzun vadede büyüdüğünü, ancak büyümenin göstergeler arasında farklı hızlarda gerçekleştiğini göstermektedir. Elektrik tüketimi 97,2 milyon MWh'den 308,7 milyon MWh'ye yükselmiştir. Dönem genelindeki artış yaklaşık %217,7, yıllık ortalama bileşik büyüme oranı ise %4,93 olmuştur. Elektrik üretimi aynı dönemde yaklaşık %183,83 artmış ve yıllık ortalama %4,44 büyümüştür. Kişi başına toplam elektrik tüketimi 1.501 kWh'den 3.603 kWh'ye çıkarak yaklaşık %140,04 artmış, yıllık ortalama büyüme oranı %3,72 olarak gerçekleşmiştir. En yüksek uzun dönemli büyüme ise kurulu güçte belirlenmiştir. Kurulu güç 27.264,1 MW'den 116.265,1 MW'ye ulaşmış, toplam artış yaklaşık %326,44 ve yıllık ortalama büyüme %6,23 olmuştur. Bu görünüm, üretim kapasitesinin elektrik tüketimi ve fiilî üretimden daha hızlı genişlediğini ortaya koymaktadır.

Holt-D modellerine ilişkin performans ölçütleri, dört enerji göstergesi için de yüksek tahmin doğruluğuna işaret etmektedir. Kare hatalar toplamı, ortalama mutlak hata ve kök ortalama kare hata ölçütlerinin serilerin ölçü birimlerine bağlı olması, bu değerlerin farklı ölçeklerdeki göstergeler arasında doğrudan karşılaştırılmamasını gerektirmektedir. Bu nedenle göreli performansın yorumlanmasında ölçekten bağımsız ölçütler daha açıklayıcıdır. Ortalama mutlak yüzde hata toplam elektrik tüketiminde %3,50, elektrik üretiminde %3,76, kişi başına elektrik tüketiminde %3,41 ve kurulu güçte %2,21 olarak hesaplanmıştır. Bütün değerlerin %10'un altında bulunması, modellerin mutlak yüzde hata bakımından başarılı sonuçlar verdiğini göstermektedir. Ortalama mutlak ölçeklendirilmiş hata değerleri ise sırasıyla 0,700, 0,804, 0,787 ve 0,338'dir. Bu değerlerin tamamının 1'in altında olması, Holt-D modellerinin karşılaştırmada kullanılan naif tahmin yaklaşımından daha düşük hata ürettiğini ortaya koymaktadır. Kurulu güç modeli %2,21 ortalama mutlak yüzde hata ve 0,338 ortalama mutlak ölçeklendirilmiş hata ile en güçlü performansı göstermiştir. Elektrik üretimi modeli, dört seri arasında ölçekten bağımsız her iki hata ölçütünün de en yüksek değerlerine sahip olsa da ulaşılan düzeyler bu modelin de başarılı olduğunu göstermektedir.

Model parametreleri, serilerin yeni gözlemlere ve eğilim değişikliklerine verdikleri tepkilerin farklılaştığını göstermektedir. Toplam elektrik tüketimi, kişi başına elektrik tüketimi ve kurulu güç modellerinde düzey yumuşatma parametresinin 1,00 olması, güncel gözlemlerin tahmin düzeyi üzerinde belirleyici olduğunu ortaya koymaktadır. Elektrik üretiminde bu parametrenin 0,84 olması, son gözlemlere yine yüksek ağırlık verildiğini, ancak geçmiş bilginin diğer serilere göre daha fazla korunduğunu göstermektedir. Trend güncelleme parametresi toplam ve kişi başına elektrik tüketiminde 0,23, elektrik üretiminde 0,27 ve kurulu güçte 0,55'tir. Kurulu güçteki daha yüksek değer, bu serinin

trend bileşeninin yeni değişimlere daha hızlı uyum sağladığını göstermektedir. Tüm modellerde sönümleme parametresinin 0,98 olması ise eğilimin geleceğe taşındığını, ancak sınırlı ölçüde sönümlendirildiğini ortaya koymaktadır.

2025-2030 tahminleri, toplam elektrik tüketimi ile elektrik üretiminin düzenli ve kesintisiz biçimde artacağını göstermektedir. Toplam elektrik tüketiminin 2025 yılındaki yaklaşık 318,0 milyon MWh düzeyinden 2030 yılında 362,2 milyon MWh'ye yükselmesi beklenmektedir. Bu değişim, 44,2 milyon MWh'lik artışa ve yaklaşık %13,9'luk büyüme oranına karşılık gelmektedir. Elektrik üretiminin ise 361,4 milyon MWh'den 405,7 milyon MWh'ye yükselmesi, böylece 44,3 milyon MWh ve yaklaşık %12,3 oranında artması beklenmektedir. Üretimin bütün tahmin yıllarında tüketimin üzerinde kalması ve iki gösterge arasındaki farkın 2025 yılında yaklaşık 43,4 milyon MWh, 2030 yılında ise 43,5 milyon MWh olması beklenmektedir. Bununla birlikte tüketimin toplam artış oranının üretimden daha yüksek olması, talepteki göreli büyümenin üretime kıyasla daha hızlı gerçekleşebileceğine işaret etmektedir.

Kişi başına elektrik tüketimi ve kurulu güç tahminleri de artış eğiliminin devam edeceğini, ancak yıllık artışların zamanla sınırlı ölçüde yavaşlayacağını göstermektedir. Kişi başına tüketimin 2025 yılındaki 3.694,6 kWh düzeyinden 2030 yılında 4.126,1 kWh'ye yükselerek 431,5 kWh ve yaklaşık %11,7 oranında artması beklenmektedir. Yıllık mutlak artışın yaklaşık 89,9 kWh'den 82,8 kWh'ye gerilemesi, tüketim yükselirken büyüme hızının kademeli biçimde azaldığını ortaya koymaktadır. Kurulu gücün ise 121.594,2 MW'den 146.683,2 MW'ye çıkarak yaklaşık 25.089 MW'lık kapasite artışıyla %20,6 oranında büyüyeceği tahmin edilmektedir. Yıllık kapasite artışının yaklaşık 5.223 MW'den 4.817 MW'ye gerilemesi, kapasite genişlemesinin süreceğini, ancak artış hızının daha ılımlı bir görünüm kazanacağını göstermektedir. Kurulu gücün kişi başına tüketimden daha hızlı büyümesi, geçmiş dönemde belirlenen kapasite ağırlıklı genişleme eğiliminin tahmin döneminde de devam edeceğine işaret etmektedir.

Bulguların yorumlanmasında enerji sisteminin diğer bileşenlerinin dikkate alınması gerekmektedir. Elektrik üretiminin tüketimin üzerinde tahmin edilmesi, tek başına arz fazlası bulunduğu anlamına gelmemektedir. Elektrik ihracatı ve ithalatı, iletim ve dağıtım kayıpları, santrallerin iç tüketimleri ve diğer sistem bileşenleri değerlendirilmeden bu farka ilişkin daha kapsamlı bir sonuca ulaşılamaz. Benzer biçimde kurulu gücün tüketimden daha hızlı büyümesi, tek başına kapasite fazlasının veya arz güvenliğinin göstergesi değildir. Kapasite kullanım oranları, üretim teknolojilerinin bileşimi, yenilenebilir enerji kaynaklarının payı, puant talep, elektrik üretim miktarları ve üretim kaynaklarının yapısı birlikte incelenmelidir. Bu kapsam sınırları içinde Holt-D modeliyle elde edilen tahminler, Türkiye'de elektrik tüketimi, üretimi ve kapasitesindeki olası gelişimin birlikte izlenmesine olanak sağlamakta ve uzun dönemli enerji planlaması için temel bilgi sunmaktadır.

KAYNAKLAR

A. Albostan, Y. Çekiç, ve L. Eren, "Rüzgar Enerjisinin Türkiye'nin Enerji Arz Güvenliğine Etkisi", Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, c. 24, sy 4, Mar. 2013.

Aslan, Y., & Bayraktutan, Y. (2025). Sanayi şehirlerinin turizm şehrine dönüşümü: Kocaeli için simülasyon. Econder Uluslararası Akademik Dergi, 9(1), 38–64.

Bolakar Tosun, H. (2025). Increasing number of vehicles and traffic accidents in Turkey: Current situation and future predictions. Journal of Innovations in Civil Engineering and Technology (JICIVILTECH), 7(1), 1–9.

Chase, C. W., Jr. (2013). Demand-driven forecasting: A structured approach to forecasting (2. bs.). John Wiley & Sons.

Ekin, E. (2026). Quantitative demand forecasting of spare parts in the aviation industry: A comparative analysis. Journal of Aviation Research, 8(1), 1–32.

Gardner, E. S., Jr., & McKenzie, E. (1985). Forecasting trends in time series. Management Science, 31(10), 1237–1246.

H. Islatince ve C. Haydaroğlu, "Türk Imalat Sanayiinde Energi Verimliliği Ve Yoğunluğunun Analizi", Dumlupınar Üniversitesi Sosyal Bilimler Dergisi, sy 24, Haz. 2015.

Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. International Journal of Forecasting, 20, 5–10.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. International Journal of Forecasting, 22(4), 679-688.

Johnston, M. P. (2014). Secondary data analysis: A method of which the time has come. Qualitative and Quantitative Methods in Libraries, 3(3), 619-626.

Parildar, O., Akyurek, C. E., & Akay, D. (2024). A decision support system for skill-based nurse scheduling in an intensive care unit. Gazi University Journal of Science, 37(1), 284–308.

R. İraz, İ. Altınışık, ve H. S. Peker, "Güneş Enerjisi Yatırımlarına Yönelik Teşvikler Ve Türkiye'deki Durum", Selçuk Üniversitesi Sosyal Bilimler Meslek Yüksekokulu Dergisi, c. 13, sy 1-2, ss. 69–78, May. 2014.

Soylemez, I., Nalici, M. E., & Ünlü, R. (2024). Türkiye'de tahıl üretiminin tahminlemesi: Karşılaştırmalı analiz. Erciyes Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 40(2), 288–303.

Soysal, M., & Ömürgönülşen, M. (2010). Türk turizm sektöründe talep tahmini üzerine bir uygulama. Anatolia: Turizm Araştırmaları Dergisi, 21(1), 128–136.

Türkiye İstatistik Kurumu. (2026). Enerji istatistikleri. Merkezi Dağıtım Sistemi (MEDAS). 2 Temmuz 2026 tarihinde https://biruni.tuik.gov.tr/medas/?locale=tr

Tüzemen, A. (2020). Cumhuriyet altını fiyatlarının ARIMA yöntemi kullanılarak ileri tahmini. Karadeniz Teknik Üniversitesi Sosyal Bilimler Enstitüsü Sosyal Bilimler Dergisi, (20), 361–381.

U.S. Energy Information Administration. (2026a). *Generator nameplate capacity* [kurulu güç]. Erişim tarihi: 2 Temmuz 2026, https://www.eia.gov/tools/glossary/index.php

U.S. Energy Information Administration. (2026b). *Gross generation* [Brüt üretim]. Erişim tarihi: 2 Temmuz 2026, https://www.eia.gov/tools/glossary/index.php

U.S. Energy Information Administration. (2026c). *Net generation* [Net üretim]. Erişim tarihi: 2 Temmuz 2026, https://www.eia.gov/tools/glossary/index.php

Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. Management Science, 6(3), 324–342.

World Bank. (2026). *Electric power consumption (kWh per capita)* [Kişi başına elektrik enerjisi tüketimi (kWh/kişi)]. World Development Indicators. Erişim tarihi: 2 Temmuz 2026, https://databank.worldbank.org/metadataglossary/world-development-indicators/series/EG.USE.ELEC.KH.PC

Yürekli, K., Ünlükara, A., & Cömert, M. M. (2017). Holt yaklaşımı ile referans bitki su tüketiminin mevsimsel değişiminin saptanması. Gaziosmanpaşa Bilimsel Araştırma Dergisi, 6(Özel Sayı [BSM-2017]), 75–81.
Lewis, C. D. (1982). Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting. Butterworth Scientific.
Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. International Journal of Forecasting, 22(4), 679-688.
Willmott, C. J., & Matsuura, K. (2005). Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Climate Research, 30(1), 79–82.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Yapay Zekâ Kavramının Tarihsel ve Kavramsal Dönüşümü

Ali Şimşek[1*] ve Agah Başdeğirmen[2]

[1] *Yönetim ve Organizasyon Bölümü / Isparta Meslek Yüksekokulu, Isparta Uygulamalı Bilimler Üniversitesi, Türkiye*
[2] *Yönetim ve Organizasyon Bölümü / Isparta Meslek Yüksekokulu, Isparta Uygulamalı Bilimler Üniversitesi, Türkiye*

[*] *(alisimsek@isparta.edu.tr)*

Özet - Yapay zekâ kavramı, makinelerin yalnızca hesaplama yapabilen araçlar olarak görülmesinden öğrenebilen, karar verebilen, içerik üretebilen ve çevreleriyle etkileşim kurabilen sistemler olarak değerlendirilmesine uzanan önemli bir dönüşüm geçirmiştir. Bu çalışmanın amacı, söz konusu dönüşümü zekâ, düşünme, anlama, bilgi temsili, öğrenme ve özerklik kavramları üzerinden tarihsel ve kavramsal olarak incelemektir. Nitel araştırma yaklaşımının benimsendiği çalışmada doküman incelemesi yöntemi kullanılmıştır. Yapay zekânın kuramsal ve yöntemsel gelişimini temsil eden birincil akademik çalışmalar, hakemli makaleler, akademik kitaplar, konferans bildirileri ve güncel incelemeler ölçüt örnekleme ve amaçlı örnekleme yoluyla belirlenmiştir. İncelenen kaynaklar dokuz tarihsel dönem altında sınıflandırılmış, kronolojik ve tematik bir yaklaşımla çözümlenmiştir. Bulgular, erken dönemde nöral etkinlik ve biçimsel mantık temelinde açıklanan zekânın zamanla sembolik problem çözme, uzman bilgisi, olasılıksal çıkarım, veriden öğrenme ve derin sinir ağları çerçevesinde yeniden tanımlandığını göstermektedir. Transformer mimarileriyle birlikte ön eğitim, ölçeklendirme ve farklı görevlere uyarlanabilme özellikleri ön plana çıkmış, üretken ve çok modlu modeller ise yapay zekâyı doğal dil, görüntü ve diğer veri türlerini birlikte işleyebilen genel amaçlı sistemlere dönüştürmüştür. Bununla birlikte Turing'in davranış temelli ölçütleri ile Searle'ün anlamaya ilişkin eleştirileri, gözlemlenebilir başarı ile bilişsel anlama arasındaki ayrımın güncelliğini koruduğunu göstermektedir. Sonuç olarak yapay zekâ kavramının dönüşümü, teknik ilerleme kadar insan zekâsının niteliğine ilişkin felsefi ve toplumsal tartışmalarla da biçimlenmektedir.

Anahtar Kelimeler - Bilgi Temsili, Doküman İncelemesi, Makine Zekâsı, Tarihsel Gelişim, Yapay Zekâ

I. GİRİŞ

Yapay zekâ (İng. artificial intelligence; İngilizce kıs. [AI], Türkçe kıs. [YZ]), insan bilişinin makineler aracılığıyla modellenmesini amaçlayan disiplinler arası bir alandır. YZ'nin tarihsel gelişimi, sinir ağlarından çağdaş öğrenme sistemlerine uzanan bir dönüşüm sürecini kapsamaktadır. Bu süreç, düşünme ve anlama kavramlarına ilişkin felsefi tartışmaların derinleşmesine katkıda bulunmuştur. Günümüzde bu alandaki kuramsal birikim, farklı uygulamalar aracılığıyla sürekli olarak genişlemektedir. YZ, klasik mantık programlama, makine öğrenmesi ve veri madenciliği gibi farklı araştırma alanlarını kapsamaktadır [1].

YZ alanındaki ilk çalışma olarak kabul edilen McCulloch ve Pitts'in [2] araştırmasında sinir sistemi, birbirine bağlı nöronlardan oluşan bir ağ olarak ele alınmıştır. Nöronların belirli bir eşik değerini aşan uyarılar karşısında "ya hep ya hiç" ilkesine göre etkinleştiği ve her nöral etkinliğin bir önerme olarak temsil edilebileceği varsayılmıştır. Bu varsayım doğrultusunda nöral ilişkiler, iki değerli önerme mantığı temelinde matematiksel olarak incelenmiştir. Hebb [3], bir hücrenin başka bir hücrenin yinelenen veya sürekli biçimde etkinleşmesine katkıda bulunması durumunda hücreler arasındaki bağlantının etkinliğinin artabileceğini ileri süren nörofizyolojik bir öğrenme postulatı geliştirmiştir. Bu hipotez, daha sonra dijital bilgisayar ortamında gerçekleştirilen sinir ağı benzetimleri aracılığıyla sınanmıştır [4]. Marvin Minsky ve Dean Edmonds, ilk sinir ağı bilgisayarlarından biri olarak kabul edilen SNARC'ı (Stochastic Neural

Analog Reinforcement Calculator, Stokastik Sinirsel Analog Pekiştirme Hesaplayıcısı) 1951 yılında geliştirmişlerdir [5]. McCarthy vd. tarafından 1955 yılında sunulan öneri doğrultusunda 1956 yılında gerçekleştirilen Dartmouth Yaz Araştırma Projesi, YZ alanının kuruluşunda belirleyici bir dönüm noktası olarak kabul edilmektedir [6]. YZ, karmaşık ve değişen ortamlarda öğrenebilen, tepki verebilen ve karar alabilen bilgisayar sistemlerinin geliştirilmesini amaçlayan yeni bir disiplin olarak ortaya çıkmıştır [7].

Alan Turing [8], "makineler düşünebilir mi?" sorusunda yer alan "makine" ve "düşünmek" kavramlarını doğrudan tanımlamanın yol açabileceği güçlükler üzerinde durmuş ve problemi daha açık biçimde tanımlanan taklit oyunu üzerinden ele almıştır. Oyunda bir erkek, bir kadın ve bu iki katılımcıdan ayrı bir odada bulunan bir sorgulayıcı yer almaktadır. Yazılı veya daktiloyla hazırlanmış yanıtlar aracılığıyla yürütülen iletişim sürecinde sorgulayıcı, katılımcıların kimliklerini belirlemeye çalışmaktadır. Erkeğin yerini bir makinenin alması durumunda sorgulayıcının yanılma düzeyi, makinenin performansının değerlendirilmesinde bir ölçüt olarak kullanılmaktadır. Bu düzenek, fiziksel özellikleri değerlendirme kapsamı dışında bırakarak insan ile makinenin gözlemlenebilir davranışlarının doğrudan karşılaştırılmasına olanak sağlamaktadır.

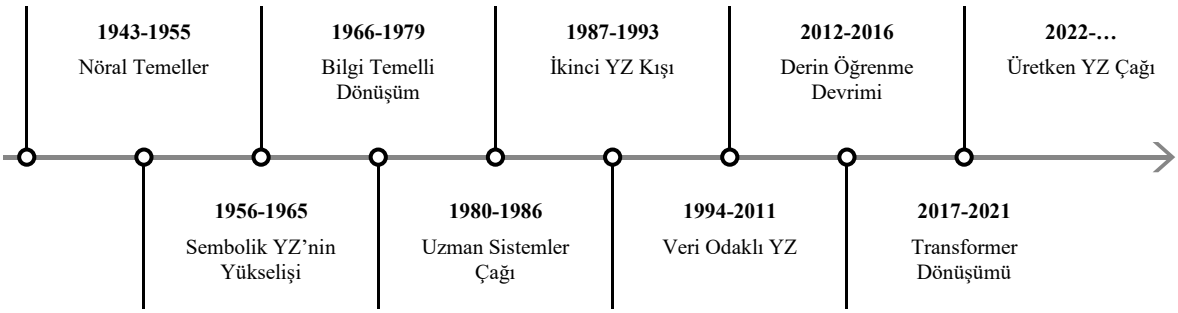
Searle [9], güçlü YZ yaklaşımına göre uygun biçimde programlanmış bir bilgisayarın gerçekten bir zihne ve bilişsel durumlara sahip olduğu yönündeki iddiayı, Çince sembollere dayalı bir düşünce deneyi aracılığıyla incelemiştir. Deneyde Çince bilmeyen bir kişi, kapalı bir odada kendisine verilen Çince sembolleri İngilizce olarak sunulan kurallar doğrultusunda yalnızca biçimsel olarak işlemekte ve ana dili Çince olan kişilerin yanıtlarından ayırt edilemeyen çıktılar üretmektedir. Buna karşın söz konusu kişi Çinceyi anlamamaktadır. Searle, doğru girdiye, çıktıya ve programa sahip olmanın anlamayı güvence altına almadığını ve Turing testini geçmenin anlamanın varlığını kanıtlamadığını savunmaktadır. Zayıf YZ'yi ise bilgisayarların zihnin incelenmesine yönelik hipotezlerin oluşturulmasına ve sınanmasına olanak tanıyan güçlü araştırma araçları olarak kullanılmasıyla ilişkilendirmektedir.

Rich'e [10] göre YZ, bilgisayarların hâlihazırda insanlar tarafından daha başarılı biçimde gerçekleştirilen işleri nasıl yapabileceğini inceleyen bir alandır. Kaplan'a [11] göre YZ, sınırlı verilerden hareketle uygun genellemeleri zamanında oluşturabilme yeteneğine dayanmaktadır. Kaplan ve Haenlein'e [12] göre YZ, bir sistemin dış verileri doğru biçimde yorumlayabilme, bu verilerden öğrenebilme ve edindiği bilgileri esnek uyum yoluyla belirli amaçlara ulaşmak ve görevleri gerçekleştirmek için kullanabilme yeteneğidir. Poole ve Mackworth'a [13] göre YZ, içinde bulunduğu çevreyi algılayan, amaçları doğrultusunda eylemde bulunan ve karar verme süreçleri hesaplama yoluyla açıklanabilen etmenlerin tasarlanmasını, geliştirilmesini ve incelenmesini konu alan bir bilim ve mühendislik alanıdır.

Günümüzde YZ, yaşamı ve çalışmayı dönüştüren, hızla gelişen bir alandır. Sanal asistanlar, sürücüsüz araçlar, robotlar, tıbbi görüntüleme ve tanı araçları ile büyük verideki örüntüleri belirleyen makine öğrenmesi başlıca uygulamalardır [14]. Bu araştırma, nitel araştırma yaklaşımı kapsamında doküman incelemesine dayalı araştırma yöntemi tercih edilmiştir. Araştırmanın veri kaynaklarını, yapay zekâ kavramının kuramsal ve yöntemsel gelişimini temsil eden birincil akademik çalışmalar, hakemli makaleler, akademik kitaplar, konferans bildirileri ve güncel inceleme çalışmaları oluşturmuştur. Yapay zekânın tanımlanmasına katkıda bulunan, alanın önemli bir kuramsal veya teknolojik dönüşümünü temsil eden ve yayın bilgileri doğrulanabilen çalışmalar araştırmaya dâhil edilmiştir. İncelenen kaynaklar öncelikle tarihsel dönemlere göre sınıflandırılmış, ardından zekâ anlayışı, bilgi temsili, öğrenme biçimi, insan-makine ilişkisi, başarı ölçütleri, yöntemsel sınırlılıklar ve etik sonuçlar bakımından tematik olarak çözümlenmiştir. Elde edilen bulgular kronolojik ve karşılaştırmalı bir yapı içerisinde yorumlanmıştır.

II. YAPAY ZEKÂNIN TARİHİ

Bu bölümde YZ'nin tarihsel gelişimi, 1943'ten günümüze uzanan dokuz dönem çerçevesinde nöral ve kuramsal temellerin oluşumundan sembolik yaklaşımın kurumsallaşmasına, bilgi temelli sistemlerden istatistiksel ve derin öğrenme yöntemlerine, Transformer mimarilerinden üretken ve çok modlu modellere uzanan bir süreç olarak kronolojik biçimde incelenmektedir. Bu dönemlendirme aracılığıyla alanın kuramsal yönelimleri, yöntemsel dönüşümleri, başarıları ve sınırlılıkları ile teknolojik gelişmelerin günümüzdeki etik, toplumsal ve bilimsel tartışmalara nasıl zemin hazırladığı bütüncül biçimde ortaya konulmaktadır. YZ'nin tarihsel gelişimi Şekil 1'de gösterilmektedir.



Şekil 1. Yapay Zekânın Tarihsel Gelişimi

Hesaplamalı Zekânın Nöral ve Kuramsal Temelleri (1943-1955): Bu dönem, biyolojik sinir sistemlerinin mantıksal ve matematiksel modellerle açıklanmaya çalışıldığı ve insan zekâsına özgü olduğu kabul edilen davranışların makinelere aktarılabileceği düşüncesinin geliştiği bir evreyi kapsamaktadır [2], [3], [6], [8], [15] - [17]. Dönemin başlangıcında nöral etkinlikler önerme mantığı temelinde modellenmiş ve böylece sinir ağları ile hesaplama arasında biçimsel bir ilişki kurulmuştur [2]. Aynı dönemde geri bildirim, amaç ve öngörü kavramları aracılığıyla canlıların ve makinelerin davranışları ortak bir çerçevede incelenmiştir [15]. Bu kuramsal yönelim, sinaptik bağlantıların deneyim yoluyla güçlenebileceğini öne süren öğrenme postulatıyla gelişmiştir [3]. Ardından satranç üzerinden programlanabilir arama ve karar verme süreçleri ele alınmış [16] ve makine zekâsı taklit oyunu aracılığıyla işlemsel bir değerlendirme problemine dönüştürülmüştür [8]. Ortaya çıkan bu düşünsel birikim, zihnin ve beynin ilk modern hesaplamalı kuramı olarak değerlendirilmiştir [17]. Söz konusu birikim, öğrenme ve zekânın makineler tarafından benzetilebileceğini öne süren 1955 tarihli Dartmouth önerisiyle ortak bir araştırma gündemine dönüşmüş ve YZ'nin kurumsallaşmasına zemin hazırlamıştır [6].

Yapay Zekânın Kurumsallaşması ve Sembolik Yaklaşımın Yükselişi (1956-1965): Bu dönem, YZ'nin bağımsız bir araştırma alanı olarak kurumsallaştığı, sembolik bilgi işleme ve sezgisel problem çözme yaklaşımlarının geliştiği, öğrenme ve örüntü tanımaya yönelik hesaplamalı modellerin oluşturulduğu süreci kapsamaktadır [18] - [23]. Sezgisel yöntemlerle sembolik mantık teoremlerini kanıtlayabilen mantık kuramı makinesi, karmaşık bilgi işlemenin biçimsel olarak tanımlanabileceğini göstermiştir [18]. Problem çözme davranışı, sembolleştirilmiş bilgiyi belleklerde saklayıp işlemleri kurallarla birleştiren programlarla açıklanmış, mantık kuramcısı programının işleyişiyle karşılaştırılmıştır [19]. Rastlantısal bağlantılardan oluşan algılayıcı modelinin pekiştirmeyle örüntü tanıma, çağrışımsal öğrenme ve genelleme kapasitesi olasılıksal açıdan incelenmiştir [20]. Dama oyununda ezberleme ve genellemeye dayalı iki yöntem sınanmış, programın deneyimle geliştiricisinden daha iyi oynayabildiği gösterilmiştir [21]. Liste işleme (**Lis**t **P**rocessing, LISP), sembolik ifadeleri listeler ve özyinelemeli fonksiyonlarla işleyerek çıkarım ve otomatik teorem ispatı çerçevesi oluşturmuştur [22]. Arama, örüntü tanıma, öğrenme, planlama ve tümevarım araştırmaları sistematikleştirilmiş, sezgisel yöntemlerle güçlü problem çözme sistemlerinin geliştirilebileceği ileri sürülmüştür [23].

Sınırlılıkların Görünürleşmesi ve Bilgi Temelli Dönüşüm (1966-1979): Bu dönem, erken dönem YZ modellerinin temsil ve ölçeklenebilirlik bakımından taşıdığı sınırlılıkların belirginleşmesini, başlangıçtaki iyimser beklentilerin sorgulanmasını ve araştırmaların genel amaçlı sistemler geliştirme hedefinden belirli problem alanlarına özgü bilginin yapılandırılmasına yönelmesini kapsamaktadır [24] - [30]. Algılayıcı sistemlerin hesaplama gücü matematiksel olarak incelenmiş, belirli mimarilerin bazı örüntüleri temsil etme ve ayırt etmedeki temel sınırlılıkları gösterilmiştir [30]. YZ araştırmaları ileri otomasyon, merkezî sinir sistemi modellemesi ve bütünleştirme biçiminde sınıflandırılmış, kombinatoryal patlama ile ölçeklenebilirliğin ilerlemeyi engellediği savunulmuştur [26]. Probleme özgü sezgisel bilgiyi grafik aramaya dâhil eden A^* algoritması geliştirilmiş, belirli koşullarda en düşük maliyetli yolu bulduğu gösterilmiştir [24]. Sezgisel Dendritik Algoritma programının kimyagerlerin bilgisiyle kütle spektrumundan ve ampirik formülden molekül yapılarını üretmesi, elemesi ve olabilirliklerine göre sıralaması incelenmiştir [25].

Bilgisayar bilimi ampirik bir alan olarak değerlendirilmiş, fiziksel sembol sistemlerinin genel akıllı eylem için gerekli ve yeterli olduğu savunulmuştur [27]. Enfeksiyon tanısı ve antimikrobiyal tedavi seçiminde MYCIN'in üretim kuralları ve kesinlik faktörlerine dayalı akıl yürütme ve açıklama mekanizmaları incelenmiştir [28]. DENDRAL ve MYCIN temelinde yüksek performansın genel çıkarımdan çok alan uzmanlığına dayandığı savunulmuş ve bilgi mühendisliğinin ilkeleri ortaya konmuştur [29].

Uzman Sistemlerin Ticarileşmesi ve Yükselişi (1980-1986): Bu dönem, uzman sistemlerin araştırma prototipi aşamasından düzenli endüstriyel kullanıma geçişini, kural tabanlı çıkarım ve bilgi mühendisliği araçlarının olgunlaşmasını ve nöral ağ araştırmalarının yeniden güç kazanmasını kapsamaktadır [31] - [37]. Rete algoritması, kural tabanlı uzman sistemlerde örüntüleri ağ yapısında derleyip eşleşmeleri saklayarak çalışma belleğindeki işlem maliyetini azaltmak amacıyla geliştirilmiştir [31]. Basit nöronların etkileşimiyle içerik adreslenebilir çağrışımsal bellek, örüntü tamamlama ve hata düzeltme sağlayan geri beslemeli sinir ağı modeli geliştirilmiştir [32]. **D**igital **E**quipment **C**orporation (DEC) tarafından R1 kural tabanlı sisteminin müşteri siparişlerini VAX-11/780 bilgisayarı yapılandırmalarına dönüştürmesi, değişiklikleri belirlemesi ve montaj şemaları üretmesi incelenmiştir [33]. Sembol düzeyinin üzerinde bilginin ortam, rasyonelliğin davranış yasası sayıldığı bilgisayar sistemlerine ilişkin bağımsız düzeyi tanımlayan bilgi düzeyi hipotezi geliştirilmiştir [34]. Yaklaşık on yıllık MYCIN deneyiminde bilginin temsili, edinimi, belirsiz akıl yürütme, açıklama, performans değerlendirmesi ve EMYCIN ile genelleştirme incelenmiştir [35]. Uzman sistemlerin gücü çıkarımdan çok alan uzmanlığına bağlanmış, bilgi mühendisliğinin temsil, kullanım ve edinim boyutlarındaki endüstriyel uygulama potansiyeli değerlendirilmiştir [36]. Geri yayılım yöntemi, çıktı hatasını geriye yayarak ağırlıkları ayarlamak ve gizli birimlerin göreve uygun temsiller öğrenmesini sağlamak için tanıtılmıştır [37].

İkinci Yapay Zekâ Kışı ve Yöntemsel Yeniden Yapılanma (1987-1993): Bu dönem, uzman sistemlere yönelik ticari beklenti ve yatırımların gerilemesine karşın YZ araştırmalarının zamansal öğrenme, bağlantıcı örüntü tanıma, olasılıksal ardışık modelleme, davranış temelli robotik ve modelden bağımsız pekiştirmeli öğrenme yaklaşımları doğrultusunda yeniden yapılanmasını kapsamaktadır [38] - [43]. Öğrenme sinyalini tahmin hatası yerine ardışık tahminlerin farkından üreten zamansal fark yöntemleri tanıtılmış, belirli durumlardaki yakınsama ve optimalliği gösterilmiştir [38]. Yerel alıcı alanlar ve ağırlık paylaşımı eklenen geri yayılımlı ağ, normalleştirilmiş el yazısı posta kodu rakamlarını sınıflandırmak üzere geliştirilmiştir [39]. Gizli markov modelleri, gözlem olasılığını hesaplama, en olası durum dizisi ve parametre kestirimi problemleriyle incelenmiş, konuşma tanıma uygulamaları açıklanmıştır [40]. Gizli durumları bağlam birimleriyle geri besleyen basit yinelemeli ağın zamansal örüntüler ile sözdizimsel ve anlamsal sözcük kategorilerini öğrenebildiği gösterilmiştir [41]. Merkezî dünya modeli ve sembolik temsiller yerine algı-eylem etkileşimiyle çevreye doğrudan bağlanan paralel davranış katmanlarına dayalı robotik yaklaşım önerilmiştir [42]. Çevrenin geçiş modeli önceden bilinmeden q-öğrenmenin, tüm durum-eylem çiftlerinin tekrarlı örneklenmesi koşuluyla optimal eylem değerlerine 1 olasılıkla yakınsadığı kanıtlanmıştır [43].

Olasılıksal, İstatistiksel ve Veri Odaklı Yapay Zekâ (1994-2011): Bu dönem, YZ'nin belirsizliği olasılıksal olarak modelleyen, genelleme başarısını istatistiksel ölçütlerle değerlendiren ve giderek büyüyen veri kümelerinden otomatik biçimde öğrenen yöntemlere yönelmesini kapsamaktadır [5], [44] - [53]. Destek vektör ağları, yüksek boyutlu özellik uzayında en geniş marjlı karar yüzeyi oluşturarak doğrusal ayrılamayan verilerde genelleme gücü göstermiştir [44]. YZ, çevresini algılayıp amaçlı eyleyen rasyonel etmenler çerçevesinde bütünleştirilmiş, akıl yürütme, karar, öğrenme, doğal dil, görme ve robotik incelenmiştir [5]. Yinelemeli ağlardaki sönme-büyüme sorununu sabit akış ve öğrenilebilir kapılarla azaltan uzun kısa süreli bellek mimarisinin uzun bağımlılıkları öğrenebildiği gösterilmiştir [45]. Evrişimli ağların pikselden el yazısı tanımadaki üstünlüğü gösterilmiş, LeNet-5 ve graf dönüştürücülerle bütüncül eğitilen ticari belge tanıma sistemi açıklanmıştır [46]. Rastgele örneklem ve özniteliklerle kurulan ağaçları birleştiren ormanda hata ağaç sayısıyla sınıra yakınsamış, başarım ağaçların gücü ve korelasyonuna bağlanmıştır [47]. İstatistiksel uygulamalar veri ve algoritmik modelleme kültürlerine ayrılmış, makine öğrenmesinden yararlanılması ve modellerin bağımsız test tahmin doğruluğuyla değerlendirilmesi savunulmuştur [48]. Kısıtlı Boltzmann makineleriyle ön eğitilen ve geri yayılımla ayarlanan derin otokodlayıcıların verileri temel bileşenler analizinden daha etkili sıkıştırdığı gösterilmiştir

[49]. WordNet sözcüksel veri tabanının 5.247 eş anlam kümesiyle 3,2 milyon etiketli görüntüyü içeren ImageNet görüntü veri kümesi tanıtılmış, ölçek ve doğruluğun görsel tanımadaki önemi gösterilmiştir [50]. Web ölçeğindeki metin ve kullanıcı verilerinde basit, ölçeklenebilir modellerin karmaşık modelleri aşmasıyla veri, temsil ve denetimsiz öğrenmenin önemi vurgulanmıştır [51]. Koşullu bağımsızlıklarla yüksek boyutlu dağılımları sıkıştıran olasılıksal grafik modelleri, Bayes ve Markov ağlarında temsil, çıkarım ve veriden öğrenmeyle incelenmiştir [52]. DeepQA mimarisinin erişim, dil, temsil, çıkarım ve öğrenmeyi güven puanlamasıyla birleştirmesiyle Watson'ın açık alan sorularında insan uzman düzeyine ulaştığı gösterilmiştir [53].

Derin Öğrenme Devrimi (2012-2016): Bu dönem, büyük ölçekli veri kümeleri ve artan hesaplama gücü sayesinde çok katmanlı sinir ağlarının uçtan uca eğitilebilmesini, hiyerarşik temsillerin veriden otomatik olarak öğrenilmesini ve derin öğrenme yöntemleriyle farklı uygulama alanlarında önemli başarılar elde edilmesini kapsamaktadır [54] - [61]. Gizli Markov modellerinde akustik durum olasılığı kestiriminde Gauss karışımları yerine derin sinir ağlarının çeşitli konuşma tanıma görevlerindeki üstünlüğü gösterilmiştir [54]. Üretici ve ayırt edici ağların karşıt amaçlı en küçük-en büyük optimizasyonu yoluyla birlikte eğitildiği üretici çekişmeli ağlar çerçevesi geliştirilmiştir [55]. Otomatik kodlamalı varyasyonel Bayes, yeniden parametreleme dönüşümü kullanılarak stokastik alt sınır optimizasyonu sağlayarak sürekli gizil modellerde ölçeklenebilir çıkarım ve otokodlama sunmuştur [56]. Değişken uzunluklu girdiyi kodlayan ve çıktıyı üreten iki uzun kısa süreli bellek (**L**ong **S**hort-**T**erm **M**emory, LSTM) diziden diziye mimarinin İngilizce-Fransızca çeviride tümce temsillerini doğrudan öğrenebildiği gösterilmiştir [57]. Yumuşak hizalama, sabit bağlam darboğazını azaltarak kod çözücünün hedef üretiminde kaynak bölümlerine odaklanmasını ve hizalamayla çeviriyi birlikte öğrenmesini sağlamıştır [58]. Derin öğrenmede ham veriden soyut hiyerarşik temsiller öğrenme incelenmiş, evrişimli ve yinelemeli ağların görüntü, konuşma ve metindeki başarıları sentezlenmiştir [59]. Kimlik kısayollarıyla hedef yerine artık dönüşüm öğrenen 152 katmanlı derin ağlar etkili eğitilmiş ve ImageNet sınıflandırmasında üstün başarı sağlamıştır [60]. AlphaGo sistemi, insan oyunlarından denetimli, kendi oyunlarından pekiştirmeli öğrenmeyi politika-değer ağlarıyla Monte Carlo aramasında birleştirerek profesyonel go oyuncusunu yenmiştir [61].

Transformer Mimarileri ve Temel Modeller (2017-2021): Bu dönem, öz-dikkat temelli Transformer mimarisinin geliştirilmesini, büyük ölçekli ön eğitimin genel amaçlı temel modellerin oluşturulmasına zemin hazırlamasını, bu yaklaşımın görsel ve bilimsel alanlara yayılmasını ve modellerin ölçeklendirilmesinin çevresel, ekonomik ve toplumsal sonuçlarının tartışılmasını kapsamaktadır [62] - [72]. 1,2 milyon ImageNet görüntüsünde GPU, doğrultulmuş doğrusal birimler ve seyreltmeyle eğitilen derin evrişimli ağların sınıflandırma başarısını belirgin artırdığı gösterilmiştir [62]. Yinelemesiz ve evrişimsiz, ölçeklendirilmiş nokta çarpımı ile çok başlı öz-dikkate dayanan Transformer'ın çeviride hız, paralellik ve başarı sağladığı gösterilmiştir [63]. Etiketsiz metinlerden maskeli dil ve cümle tahminiyle çift yönlü Transformer temsilleri öğrenen BERT, on bir doğal dil görevine uyarlanmıştır [64]. Büyük dil modellerinin enerji tüketimi, karbon salımı ve finansal maliyetleri nicelleştirilmiş, hesaplamaya eşitsiz erişim vurgulanarak şeffaf araştırma uygulamaları önerilmiştir [65]. 175 milyar parametreli otoregresif GPT-3'ün doğal dil açıklamalarıyla görevleri parametre güncellemeden sıfır, tek ve az örnekli koşullarda gerçekleştirebildiği gösterilmiştir [66]. Veriye aşamalı Gauss gürültüsü ekleyen ileri süreci öğrenilmiş ters Markov zinciriyle çeviren gürültü giderici difüzyon modelleri nitelikli görüntüler üretmiştir [67]. Dil modeli kaybının ölçekleme yasaları belirlenmiş, sabit hesaplama bütçesinde büyük modellerin sınırlı veriyle yakınsama öncesi eğitimi daha verimli bulunmuştur [68]. Büyüyen dil modelleri çevresel-finansal maliyetler, dokümantasyon yetersizliği, önyargı ve biçim-anlam karışıklığı nedeniyle eleştirilmiş, veri seçimi, belgeleme ve paydaş değerlendirmesi önerilmiştir [69]. Görüntüleri sözcük belirteci benzeri sabit yamalara dönüştüren Görsel Transformer'ın büyük ölçekli ön eğitimin ardından evrişimli ağlarla rekabet edebildiği gösterilmiştir [70]. AlphaFold modeli, çoklu hizalamalar ve amino asit çiftlerini dikkat ağında işleyerek proteinlerin üç boyutlu yapılarını çoğunlukla deneysel doğruluğa yakın kestirmiştir [71]. 400 milyon görüntü-metin çiftiyle karşılaştırmalı eğitilen karşıt dil-görüntü ön eğitimi (**C**ontrastive **L**anguage-**I**mage **P**re-training, CLIP) doğal dil denetimiyle göreve özgü eğitim gerektirmeden sıfır örnekli aktarım sağladığı gösterilmiştir [72].

Üretken ve Çok Modlu Yapay Zekânın Yaygınlaşması (2022-Günümüz): Bu dönem, insan geri bildirimine dayalı hizalama ve düşünce zinciri istemleri aracılığıyla yetenekleri geliştirilen büyük dil modellerinin yaygınlaşmasını, farklı veri türlerini birlikte işleyebilen çok modlu temel modellerin ortaya çıkmasını ve üretken YZ'nin eğitim, araştırma, iş yaşamı ve toplum üzerindeki etkilerinin tartışılmasını kapsamaktadır [73] - [80]. İnsan gösterimleriyle ince ayar ve geri bildirimli pekiştirmeyle 1,3 milyar parametreli InstructGPT modelinin 175 milyar parametreli GPT-3'ü tercihlerde aştığı gösterilmiştir [73]. İstemlerdeki düşünce zinciri örneklerinin yeterince büyük dil modellerinde aritmetik, sağduyu ve sembolik akıl yürütme görevlerindeki performansı belirgin artırdığı gösterilmiştir [74]. Dil modellerindeki 21 risk altı başlıkta sınıflandırılmıştır: ayrımcılık, bilgi tehlikeleri, yanlış bilgilendirme, kötüye kullanım, insan-bilgisayar zararları ve çevresel-sosyoekonomik etkiler [75]. ChatGPT'nin fırsatları ve etik-hukuki-toplumsal güçlükleri 43 uzmanın görüşleriyle incelenmiş, bilgi-etik, dijital dönüşüm, eğitim ve bilimsel araştırma eksenlerinde gündem önerilmiştir [76]. Dil modellerinin üretim, kişiselleştirme ve etkileşim olanakları değerlendirilmiş, risklere karşı eleştirel doğrulama, YZ okuryazarlığı ve insan denetimli pedagoji önerilmiştir [77]. Çok modlu temel modellerde uzman modellerden birleşik görsellere, çok modlu dil modellerine ve araç kullanan genel asistanlara geçiş gösterilmiştir [78]. Üretken YZ, veri örüntülerinden anlamlı içerik oluşturan sosyoteknik yapı olarak tanımlanmış, model-sistem-uygulama düzeyleri ayrılarak bilgi sistemleri araştırma gündemi geliştirilmiştir [79]. Üretken YZ'nin yaklaşık yetmiş yılı kural, model, derin ve temel model aşamalarına ayrılmış, uygulamalar, sorunlar ve gelecek yönleri değerlendirilmiştir [80].

Sonuç olarak YZ, insan zekâsına özgü algılama, öğrenme, akıl yürütme ve karar verme süreçlerinin hesaplamalı sistemlerle modellenmesini amaçlayan disiplinler arası bir araştırma alanıdır [5]. Bu alanın tarihsel gelişimi yalnızca teknik ilerlemelerle değil, zekânın tanımı ve temsil biçimlerine ilişkin kuramsal dönüşümlerle de şekillenmiştir [81]. Sembolik yaklaşımlardan istatistiksel öğrenmeye ve derin sinir ağlarına uzanan süreç, bilimsel beklentiler ve hesaplama kapasitesindeki artışla birlikte evrimleşmiştir [59], [82]. Uzman sistemlerden derin öğrenmeye geçiş, veri temelli genelleme ve ölçeklenebilirlik sorunlarını öne çıkarmış ve bu dönüşüm alanın yöntemsel sınırlarını yeniden tanımlamıştır [29], [83]. Günümüzde Transformer tabanlı modeller çok modlu sistemlerle birlikte geniş uygulama alanları bulurken etik, şeffaflık ve hesap verebilirlik tartışmaları da giderek daha kritik hâle gelmiştir [63], [69].

III. SONUÇ VE TARTIŞMA

İlk dönem, nöral ve mantıksal temellerin birlikte geliştiğini göstermektedir. Nöral etkinliklerin mantıksal modellenmesi, sinaptik bağlantıların deneyimle güçlenmesi ve geri bildirim, öğrenmenin hesaplamalı temsilini desteklemiştir. Satrançta arama ve karar vermenin programlanması, makine zekâsını davranış üzerinden değerlendirilen araştırma problemine dönüştürmüştür. Dartmouth girişimiyle ortak gündem kazanan alan, sembolik bilgi işleme ve sezgisel problem çözmeyle bağımsızlaşmıştır. Ancak davranışsal başarı ile gerçek anlama arasındaki ayrım, Turing'in taklit oyunu ve Searle'ün düşünce deneyi ekseninde güncelliğini korumaktadır.

Sembolik dönemde zekâ, sembol işleme, aramayla problem çözme ve uzmanlığı biçimselleştirme üzerinden tanımlanmıştır. Logic Theorist, LISP ve erken sistemler, problem çözme ile örüntü tanımanın hesaplanabilirliğini göstermiştir. Ancak temsil, kombinatoryal patlama ve ölçeklenebilirlik sorunları genel amaçlı beklentileri sınırlayarak araştırmaları alana özgü uzman sistemlere yöneltmiştir. DENDRAL ve MYCIN, performansın hem çıkarıma hem yapılandırılmış uzmanlık bilgisine bağlı olduğunu göstermiştir. Ticarileşme bilgi mühendisliğini güçlendirmiş, geri yayılım nöral ağ araştırmalarını canlandırmıştır. İkinci yapay zekâ kışı araştırmaları sonlandırmayarak zamansal ve pekiştirmeli öğrenme, bağlantıcı örüntü tanıma ve davranış temelli robotik doğrultusunda yeniden yapılanmaya yol açmıştır.

1994 sonrasındaki olasılıksal, istatistiksel ve veri odaklı yaklaşım, yapay zekâ kavramını dönüştürmüştür. Olasılıksal belirsizlik temsili, bağımsız verilerde genellemenin ölçülmesi ve otomatik öğrenme, açık kurallardan veriye dayalı örüntülere geçişi hızlandırmıştır. Destek vektör yöntemleri, olasılıksal grafik modeller, rastgele ormanlar ve yinelemeli ağlar güçlü araçlar sağlamıştır. Veri ile hesaplama gücündeki artış, derin ağların çok katmanlı temsilleri ham veriden öğrenmesini mümkün kılmıştır. Konuşma tanıma, görüntü sınıflandırma, makine çevirisi ve oyunlardaki gelişmeler, zekâyı temsil öğrenme, genelleme ve uyarlanma kapasitesiyle ilişkilendirmiştir.

Transformer mimarileri ve temel modeller veri odaklı dönüşümü ölçeklendirmiştir. Öz-dikkat uzun dizileri paralel işlemeyi, büyük ölçekli ön eğitim ise modelleri sınırlı düzenlemelerle farklı problemlere uyarlamayı kolaylaştırmıştır. Dil, görüntü ve bilimsel verilerle geliştirilen temel modeller, belirli görevlere yönelik sistemlerden çok görevli genel amaçlı yapılara yönelimi göstermiştir. Üretken modeller ve insan geri bildirimine dayalı hizalama, yapay zekâyı yeni içerik oluşturabilen etkileşimli sistemlere dönüştürmüştür. Çok modlu modeller, farklı veri türlerini birlikte işleyerek algılama, üretme, uyarlanma ve araç kullanmayı birleştirmiştir.

Genel sentez, yöntemlerin birbirini ortadan kaldıran doğrusal bir ilerleme izlemediğini göstermektedir. Sembolik çıkarım, uzmanlık bilgisi, olasılıksal modelleme, bağlantıcı öğrenme ve insan geri bildirimi farklı ağırlıklarla sürmüştür. Dönüşüm, zekânın temsil birimleri, başarı ölçütleri ve çevreyle ilişkisi üzerinden gerçekleşmiştir. Dolayısıyla yapay zekâ tarihi, teknik başarıların kronolojisinden çok bilgi, öğrenme, anlam ve özerkliğin yeniden tanımlandığı çok katmanlı bir süreçtir. Alanın sınırları, hesaplama ve verinin yanı sıra insan zekâsına ilişkin beklentiler, uygulama gereksinimleri ve toplumsal etkilerle genişlemiştir.

Günümüzde teknik gelişmeler, etik ve toplumsal değerlendirmeleri yapay zekânın ayrılmaz parçasına dönüştürmüştür. Büyük modellerin enerji tüketimi, maliyeti, kaynak eşitsizliği, veri önyargıları, yanlış bilgi, kötüye kullanım, şeffaflık ve hesap verebilirlik sorunları, başarının performansla sınırlanamayacağını göstermektedir. Eğitim ve bilimde içerik üretimi, kişiselleştirme ve etkileşim önem kazanırken bilgi doğrulama, yapay zekâ okuryazarlığı ve insan denetimi gereklidir. Çalışma, ölçüt ve amaçlı örneklemeyle seçilen yayınlar ile dokuz dönemli tarihsel sınıflandırmayla sınırlıdır. Gelecekte alternatif dönemlendirmelerin, disiplinlere göre tanımların ve üretken sistemlerin anlam, özerklik ve sorumluluğa etkilerinin incelenmesi katkı sağlayabilir. Sonuç olarak yapay zekâ, insan bilişini taklit etmeye yönelik sınırlı bir girişimden öğrenen, üreten ve farklı verilerle etkileşen sistemleri kapsayan sosyoteknik bir alana dönüşmüştür. Bu dönüşüm, teknolojik kapasite ile etik sorumluluğun birlikte ele alınmasını gerektirmektedir.

KAYNAKLAR

[1] C. C. Aggarwal, Artificial Intelligence: A Textbook. Cham, İsviçre: Springer, 2021.

[2] W. S. McCulloch ve W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” Bull. Math. Biophys., c. 5, sayı 4, ss. 115-133, Aralık 1943.

[3] D. O. Hebb, The Organization of Behavior: A Neuropsychological Theory. New York, NY, ABD: John Wiley & Sons, 1949.

[4] N. Rochester, J. H. Holland, L. H. Haibt ve W. L. Duda, “Tests on a cell assembly theory of the action of the brain, using a large digital computer,” IRE Trans. Inf. Theory, c. IT-2, sayı 3, ss. 80-93, Eylül 1956.

[5] S. J. Russell ve P. Norvig, Artificial Intelligence: A Modern Approach, seri Prentice Hall Series in Artificial Intelligence. Englewood Cliffs, NJ, ABD: Prentice Hall, 1995.

[6] J. McCarthy, M. L. Minsky, N. Rochester ve C. E. Shannon, “A proposal for the Dartmouth summer research project on artificial intelligence: August 31, 1955,” AI Mag., c. 27, sayı 4, ss. 12-14, Kış 2006.

[7] R. E. Neapolitan ve X. Jiang, Artificial Intelligence: With an Introduction to Machine Learning, 2. bs. Boca Raton, FL, ABD: CRC Press, 2018.

[8] A. M. Turing, “Computing machinery and intelligence,” Mind, c. 59, sayı 236, ss. 433-460, Ekim 1950.

[9] J. R. Searle, “Minds, brains, and programs,” Behav. Brain Sci., c. 3, sayı 3, ss. 417-424, Eylül 1980.

[10] E. Rich, K. Knight ve S. B. Nair, Artificial Intelligence, 3. bs. New Delhi, Hindistan: Tata McGraw-Hill, 2010.

[11] J. Kaplan, Artificial Intelligence: What Everyone Needs to Know. New York, NY, ABD: Oxford Univ. Press, 2016.

[12] A. Kaplan ve M. Haenlein, “Siri, Siri, in my hand: Who’s the fairest in the land? On the interpretations, illustrations, and implications of artificial intelligence,” Bus. Horiz., c. 62, sayı 1, ss. 15-25, Ocak-Şubat 2019.

[13] D. L. Poole ve A. K. Mackworth, Artificial Intelligence: Foundations of Computational Agents, 3. bs. Cambridge, Birleşik Krallık: Cambridge Univ. Press, 2023.

[14] A. Banafa, Introduction to Artificial Intelligence (AI), seri River Publishers Series in Rapids. Gistrup, Danimarka: River Publishers, 2024.

[15] A. Rosenblueth, N. Wiener ve J. Bigelow, “Behavior, purpose and teleology,” Philos. Sci., c. 10, sayı 1, ss. 18-24, Ocak 1943.

[16] C. E. Shannon, “Programming a computer for playing chess,” Philos. Mag., seri 7, c. 41, sayı 314, ss. 256-275, Mart 1950.

[17] G. Piccinini, “The first computational theory of mind and brain: A close look at McCulloch and Pitts’s ‘logical calculus of ideas immanent in nervous activity’,” Synthese, c. 141, sayı 2, ss. 175-215, Ağustos 2004.

[18] A. Newell ve H. A. Simon, “The logic theory machine: A complex information processing system,” RAND Corp., Santa Monica, CA, ABD, RAND Belgesi P-868, 12 Temmuz 1956.
[19] A. Newell, J. C. Shaw ve H. A. Simon, “Elements of a theory of human problem solving,” Psychol. Rev., c. 65, sayı 3, ss. 151-166, Mayıs 1958.
[20] F. Rosenblatt, “The perceptron: A probabilistic model for information storage and organization in the brain,” Psychol. Rev., c. 65, sayı 6, ss. 386-408, Kasım 1958.
[21] A. L. Samuel, “Some studies in machine learning using the game of checkers,” IBM J. Res. Develop., c. 3, sayı 3, ss. 210-229, Temmuz 1959.
[22] J. McCarthy, “Recursive functions of symbolic expressions and their computation by machine, Part I,” Commun. ACM, c. 3, sayı 4, ss. 184-195, Nisan 1960.
[23] M. Minsky, “Steps toward artificial intelligence,” Proc. IRE, c. 49, sayı 1, ss. 8-30, Ocak 1961.
[24] P. E. Hart, N. J. Nilsson ve B. Raphael, “A formal basis for the heuristic determination of minimum cost paths,” IEEE Trans. Syst. Sci. Cybern., c. SSC-4, sayı 2, ss. 100-107, Temmuz 1968.
[25] B. G. Buchanan, G. L. Sutherland ve E. A. Feigenbaum, “Heuristic DENDRAL: A program for generating explanatory hypotheses in organic chemistry,” Machine Intelligence 4 içinde, B. Meltzer ve D. Michie, Ed. Edinburgh, Birleşik Krallık: Edinburgh Univ. Press, 1969, ss. 209-254.
[26] J. Lighthill, “Artificial intelligence: A general survey,” Artificial Intelligence: A Paper Symposium içinde, Science Research Council, 1973, ss. 1-21. Çevrim içi erişim: https://www.chilton-computing.org.uk/inf/literature/reports/lighthill_report/p001.htm (Erişim tarihi: 2 Temmuz 2026).
[27] A. Newell ve H. A. Simon, “Computer science as empirical inquiry: Symbols and search,” Commun. ACM, c. 19, sayı 3, ss. 113-126, Mart 1976.
[28] E. H. Shortliffe, Computer-Based Medical Consultations: MYCIN, seri Artificial Intelligence. New York, NY, ABD: American Elsevier, 1976, c. 2.
[29] E. A. Feigenbaum, “The art of artificial intelligence: I. Themes and case studies of knowledge engineering,” Proc. 5th Int. Joint Conf. Artif. Intell. (IJCAI-77) içinde, Cambridge, MA, ABD, 22-25 Ağustos 1977, ss. 1014-1029.
[30] M. L. Minsky ve S. A. Papert, Perceptrons: An Introduction to Computational Geometry, genişletilmiş bs. Cambridge, MA, ABD: MIT Press, 1988.
[31] C. L. Forgy, “Rete: A fast algorithm for the many pattern/many object pattern match problem,” Artif. Intell., c. 19, sayı 1, ss. 17-37, Eylül 1982.
[32] J. J. Hopfield, “Neural networks and physical systems with emergent collective computational abilities,” Proc. Natl. Acad. Sci. USA, c. 79, sayı 8, ss. 2554-2558, Nisan 1982.
[33] J. McDermott, “R1: A rule-based configurer of computer systems,” Artif. Intell., c. 19, sayı 1, ss. 39-88, Eylül 1982.
[34] A. Newell, “The knowledge level,” Artif. Intell., c. 18, sayı 1, ss. 87-127, Ocak 1982.
[35] B. G. Buchanan ve E. H. Shortliffe, Ed., Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project, seri The Addison-Wesley Series in Artificial Intelligence. Reading, MA, ABD: Addison-Wesley, 1984.
[36] E. A. Feigenbaum, “Knowledge engineering: The applied side of artificial intelligence,” Ann. N.Y. Acad. Sci., c. 426, sayı 1, ss. 91-107, Kasım 1984.
[37] D. E. Rumelhart, G. E. Hinton ve R. J. Williams, “Learning representations by back-propagating errors,” Nature, c. 323, sayı 6088, ss. 533-536, Ekim 1986.
[38] R. S. Sutton, “Learning to predict by the methods of temporal differences,” Mach. Learn., c. 3, sayı 1, ss. 9-44, Ağustos 1988.
[39] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard ve L. D. Jackel, “Backpropagation applied to handwritten zip code recognition,” Neural Comput., c. 1, sayı 4, ss. 541-551, Aralık 1989.
[40] L. R. Rabiner, “A tutorial on hidden Markov models and selected applications in speech recognition,” Proc. IEEE, c. 77, sayı 2, ss. 257-286, Şubat 1989.
[41] J. L. Elman, “Finding structure in time,” Cogn. Sci., c. 14, sayı 2, ss. 179-211, Nisan-Haziran 1990.
[42] R. A. Brooks, “Intelligence without representation,” Artif. Intell., c. 47, sayı 1-3, ss. 139-159, Ocak 1991.
[43] C. J. C. H. Watkins ve P. Dayan, “Q-learning,” Mach. Learn., c. 8, sayı 3-4, ss. 279-292, Mayıs 1992.
[44] C. Cortes ve V. Vapnik, “Support-vector networks,” Mach. Learn., c. 20, sayı 3, ss. 273-297, Eylül 1995.
[45] S. Hochreiter ve J. Schmidhuber, “Long short-term memory,” Neural Comput., c. 9, sayı 8, ss. 1735-1780, Kasım 1997.
[46] Y. LeCun, L. Bottou, Y. Bengio ve P. Haffner, “Gradient-based learning applied to document recognition,” Proc. IEEE, c. 86, sayı 11, ss. 2278-2324, Kasım 1998.
[47] L. Breiman, “Random forests,” Mach. Learn., c. 45, sayı 1, ss. 5-32, Ekim 2001.
[48] L. Breiman, “Statistical modeling: The two cultures,” Statist. Sci., c. 16, sayı 3, ss. 199-231, Ağustos 2001.
[49] G. E. Hinton ve R. R. Salakhutdinov, “Reducing the dimensionality of data with neural networks,” Science, c. 313, sayı 5786, ss. 504-507, Temmuz 2006.
[50] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li ve L. Fei-Fei, “ImageNet: A large-scale hierarchical image database,” Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR) içinde, Miami, FL, ABD, 2009, ss. 248-255.
[51] A. Halevy, P. Norvig ve F. Pereira, “The unreasonable effectiveness of data,” IEEE Intell. Syst., c. 24, sayı 2, ss. 8-12, Mart-Nisan 2009.
[52] D. Koller ve N. Friedman, Probabilistic Graphical Models: Principles and Techniques, seri Adaptive Computation and Machine Learning. Cambridge, MA, ABD: MIT Press, 2009.

[53] D. Ferrucci vd., "Building Watson: An overview of the DeepQA project," AI Mag., c. 31, sayı 3, ss. 59-79, Güz 2010.
[54] G. Hinton vd., "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," IEEE Signal Process. Mag., c. 29, sayı 6, ss. 82-97, Kasım 2012.
[55] I. J. Goodfellow vd., "Generative adversarial nets," Proc. Adv. Neural Inf. Process. Syst. 27 (NIPS) içinde, 2014, ss. 2672-2680.
[56] D. P. Kingma ve M. Welling, "Auto-encoding variational Bayes," Proc. Int. Conf. Learn. Represent. (ICLR) içinde, 2014.
[57] I. Sutskever, O. Vinyals ve Q. V. Le, "Sequence to sequence learning with neural networks," Proc. Adv. Neural Inf. Process. Syst. 27 (NIPS) içinde, 2014, ss. 3104-3112.
[58] D. Bahdanau, K. Cho ve Y. Bengio, "Neural machine translation by jointly learning to align and translate," Proc. Int. Conf. Learn. Represent. (ICLR) içinde, San Diego, CA, ABD, 2015.
[59] Y. LeCun, Y. Bengio ve G. Hinton, "Deep learning," Nature, c. 521, sayı 7553, ss. 436-444, Mayıs 2015.
[60] K. He, X. Zhang, S. Ren ve J. Sun, "Deep residual learning for image recognition," Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR) içinde, Las Vegas, NV, ABD, 2016, ss. 770-778.
[61] D. Silver vd., "Mastering the game of Go with deep neural networks and tree search," Nature, c. 529, sayı 7587, ss. 484-489, Ocak 2016.
[62] A. Krizhevsky, I. Sutskever ve G. E. Hinton, "ImageNet classification with deep convolutional neural networks," Commun. ACM, c. 60, sayı 6, ss. 84-90, Haziran 2017.
[63] A. Vaswani vd., "Attention is all you need," Proc. Adv. Neural Inf. Process. Syst. 30 (NIPS) içinde, 2017, ss. 5998-6008.
[64] J. Devlin, M.-W. Chang, K. Lee ve K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," Proc. 2019 Conf. North Amer. Chapter Assoc. Comput. Linguistics: Human Lang. Technol. (NAACL-HLT) içinde, c. 1, Minneapolis, MN, ABD, 2019, ss. 4171-4186.
[65] E. Strubell, A. Ganesh ve A. McCallum, "Energy and policy considerations for deep learning in NLP," Proc. 57th Annu. Meeting Assoc. Comput. Linguistics (ACL) içinde, Floransa, İtalya, 2019, ss. 3645-3650.
[66] T. B. Brown vd., "Language models are few-shot learners," Proc. Adv. Neural Inf. Process. Syst. (NeurIPS) içinde, c. 33, 2020, ss. 1877-1901.
[67] J. Ho, A. Jain ve P. Abbeel, "Denoising diffusion probabilistic models," Proc. Adv. Neural Inf. Process. Syst. (NeurIPS) içinde, c. 33, 2020, ss. 6840-6851.
[68] J. Kaplan vd., "Scaling laws for neural language models," arXiv:2001.08361, Ocak 2020.
[69] E. M. Bender, T. Gebru, A. McMillan-Major ve S. Shmitchell, "On the dangers of stochastic parrots: Can language models be too big?," Proc. 2021 ACM Conf. Fairness, Accountability, and Transparency (FAccT) içinde, 2021, ss. 610-623.
[70] A. Dosovitskiy vd., "An image is worth 16 × 16 words: Transformers for image recognition at scale," Proc. Int. Conf. Learn. Represent. (ICLR) içinde, 2021.
[71] J. Jumper vd., "Highly accurate protein structure prediction with AlphaFold," Nature, c. 596, sayı 7873, ss. 583-589, Ağustos 2021.
[72] A. Radford vd., "Learning transferable visual models from natural language supervision," Proc. 38th Int. Conf. Mach. Learn. (ICML) içinde, seri Proc. Mach. Learn. Res., c. 139, 2021, ss. 8748-8763.
[73] L. Ouyang vd., "Training language models to follow instructions with human feedback," Proc. Adv. Neural Inf. Process. Syst. (NeurIPS) içinde, c. 35, 2022, ss. 27730-27744.
[74] J. Wei vd., "Chain-of-thought prompting elicits reasoning in large language models," Proc. Adv. Neural Inf. Process. Syst. (NeurIPS) içinde, c. 35, 2022, ss. 24824-24837.
[75] L. Weidinger vd., "Taxonomy of risks posed by language models," Proc. 2022 ACM Conf. Fairness, Accountability, and Transparency (FAccT) içinde, Seul, Kore Cumhuriyeti, 2022, ss. 214-229.
[76] Y. K. Dwivedi vd., "Opinion paper: 'So what if ChatGPT wrote it?' Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy," Int. J. Inf. Manage., c. 71, Makale no. 102642, Ağustos 2023.
[77] E. Kasneci vd., "ChatGPT for good? On opportunities and challenges of large language models for education," Learn. Individ. Differ., c. 103, Makale no. 102274, Nisan 2023.
[78] C. Li vd., "Multimodal foundation models: From specialists to general-purpose assistants," Found. Trends Comput. Graph. Vis., c. 16, sayı 1-2, ss. 1-214, Mayıs 2024.
[79] S. Feuerriegel, J. Hartmann, C. Janiesch ve P. Zschech, "Generative AI," Bus. Inf. Syst. Eng., c. 66, sayı 1, ss. 111-126, Şubat 2024.
[80] R. He, J. Cao ve T. Tan, "Generative artificial intelligence: A historical perspective," Natl. Sci. Rev., c. 12, sayı 5, Makale no. nwaf050, Mayıs 2025.
[81] N. J. Nilsson, The Quest for Artificial Intelligence: A History of Ideas and Achievements. Cambridge, Birleşik Krallık: Cambridge Univ. Press, 2010.
[82] T. M. Mitchell, Machine Learning. New York, NY, ABD: McGraw-Hill, 1997.
[83] I. Goodfellow, Y. Bengio ve A. Courville, Deep Learning, seri Adaptive Computation and Machine Learning. Cambridge, MA, ABD: MIT Press, 2016.

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

Decoupling the Effects of Water to Cement Ratio and Firing Temperature on the Mechanical Performance of Vibration Cast Calcium Aluminate Cement Bodies

Ali KAFALI* and Safa POLAT[2]

[1]*Department of Mechatronics Engineering, Karabuk University, Turkiye*
[2]*Department of Mechanical Engineering, Karabuk University, Turkiye*

**(alikafali@karabuk.edu.tr) Email of the corresponding author*

Abstract - Mixing water and firing temperature both govern the mechanical performance of calcium aluminate cement (CAC) bodies, but they are rarely separated because in practice they move together. This study isolates them experimentally. Pure SECAR 80 cement pastes at water-to-cement (w/c) ratios of 0.30 and 0.40 were consolidated by vibration casting into Ø15 × 15 mm cylinders and fired at 900 °C or 1200 °C; cold crushing strength, green and fired bulk density, mass loss and calculated apparent porosity were determined. At 900 °C, raising the w/c ratio from 0.30 to 0.40 reduced the mean crushing strength from 34.2 ± 7.0 MPa to 14.6 ± 3.3 MPa (−57.3 %) and increased apparent porosity by 10.6 percentage points. A green-state mass balance showed the 0.30 paste consolidated essentially air-free, whereas the 0.40 batch retained 13.8 vol.% entrained air, indicating that excess water acts through two mechanisms: capillary porosity left by evaporated free water, and impaired vibratory compaction of an over-fluid, segregation-prone paste. At constant w/c = 0.40, raising the firing temperature from 900 °C to 1200 °C increased the mean strength by 84.5 % while apparent porosity fell by only 2.2 percentage points; this disproportionate gain is attributed to the onset of sintering and the development of a ceramic bond rather than to densification. Minimizing mixing water remains the dominant lever on strength.

Keywords – Calcium aluminate cement, SECAR 80, vibration casting, sintering temperature, apparent porosity.

I. INTRODUCTION

Calcium aluminate cements (CAC) are the principal hydraulic binders of the modern refractory industry. Unlike Portland cement, whose hydration chemistry is dominated by calcium silicates, CAC derive their setting behaviour from calcium monoaluminate ($CaO{\cdot}Al_2O_3$, CA) and calcium dialuminate ($CaO{\cdot}2Al_2O_3$, CA_2), which hydrate rapidly to metastable hexagonal hydrates and subsequently convert to the cubic hydrogarnet C_3AH_6 together with gibbsite [1, 2]. This chemistry gives CAC-bonded bodies a useful combination of rapid strength development at ambient temperature and, after firing, a refractoriness that ordinary hydraulic binders cannot approach.

The service life of a CAC-bonded component, however, is set less by the chemistry of the binder than by the porosity of the body produced around it. Two processing decisions dominate that porosity. The first is the quantity of water added at mixing. Water more than the stoichiometric hydration demand does not participate in the reaction; it occupies volume in the fresh paste, and on drying and firing it leaves behind a connected capillary pore network [3]. The second is the temperature of the subsequent heat treatment, which determines whether the load-bearing skeleton of the fired body is a dehydrated but essentially unsintered assembly of collapsed hydrate pseudomorphs, or a genuinely sintered ceramic in which neck growth has re-established continuity across former particle boundaries.

These two variables are rarely separated cleanly in the published literature, because in industrial practice they move together: a formulation requiring more water is usually also the one fired more aggressively to compensate. The consequence is that the relative magnitude of the two effects, and in

particular the extent to which a high firing temperature can rescue a body cast too wet, remains poorly quantified for pure cement systems.

A further complication is that mixing water influences the fired body through two distinct routes, not one. The first is the direct route just described. The second is indirect and concerns the forming step. Vibration casting with vibratory compaction relies on the paste possessing a yield stress low enough to flow and de-air under vibration, but high enough to resist segregation and bleeding. A paste that is too fluid will bleed, and the vibration intended to expel entrapped air will instead drive segregation and stabilise it. If this occurs, the fired body carries not only capillary porosity but also a population of coarse entrained air voids, which are far more damaging per unit volume because they act as pre-existing critical flaws rather than as a uniform reduction of load-bearing area [4].

The porosity dependence of the strength of brittle porous solids is conventionally described by the empirical exponential relation independently proposed by Ryshkewitch [5], in which strength decays exponentially with the pore volume fraction [6]. Provided the pore architecture is comparable between specimens, this relation offers a practical means of normalising strength data for porosity differences, and therefore of testing whether a strength change between two conditions can be explained by porosity alone or whether an additional bonding mechanism must be invoked. This normalisation is the central analytical tool of the present work.

The aim of this study is therefore to separate, within a single pure-cement system and a single forming route, the strength contribution of the water-to-cement ratio from that of the firing temperature. SECAR 80, a high-alumina CAC containing approximately 40 wt.% free α-Al_2O_3 in addition to CA and CA_2 clinker phases [1, 7], was chosen as a model binder because it is free of the silicate and ferrite phases that complicate the interpretation of thermal behaviour in lower-grade cements. A three-group design was adopted in which one pair of groups differs only in w/c ratio and a second pair differs only in firing temperature, so that each effect is isolated by a single comparison.

II. Materials and Method

A. Raw Material

The sole solid raw material was SECAR 80 calcium aluminate cement (Imerys Aluminates, France). No aggregate, filler, deflocculant, dispersant or set-controlling admixture was added, so that the observed behaviour could be attributed unambiguously to the cement matrix itself. Nominal characteristics are summarised in Table 1. The presence of a substantial fraction of free α-alumina distinguishes SECAR 80 from lower grades such as SECAR 51 and 71 and raises the true density of the powder above that of the clinker phases alone [1].

Table 1. Nominal characteristics of the Secar 80 cement

Characteristic	Nominal value
Designation	SECAR 80 (Imerys Aluminates)
Total Al_2O_3 content	≈ 80 wt.%
Free α-Al_2O_3 addition	≈ 40 wt.%
Principal crystalline phases	CA, CA_2, α-Al_2O_3
Blaine specific surface area	≈ 10 600 cm² g⁻¹
True density (calculated)	3.25 g cm⁻³

The true density in Table 1 was calculated as a volume-weighted mean of 40 wt.% α-Al_2O_3 (3.99 g cm⁻³) and 60 wt.% CA/CA_2 clinker (≈ 2.95 g cm⁻³). This value is used throughout the green-state analysis; it was not measured pycnometrically and its uncertainty is discussed in Section IV.E.

B. Mix Design and Specimen Preparation

Two water-to-cement ratios were investigated, w/c = 0.30 and w/c = 0.40 by mass, corresponding to 30 and 40 parts of deionised water per 100 parts of cement. These values bracket the range over which a pure SECAR 80 paste passes from a stiff, barely castable consistency to a freely flowing one, and therefore represent the practical decision a producer faces when selecting a placing consistency.

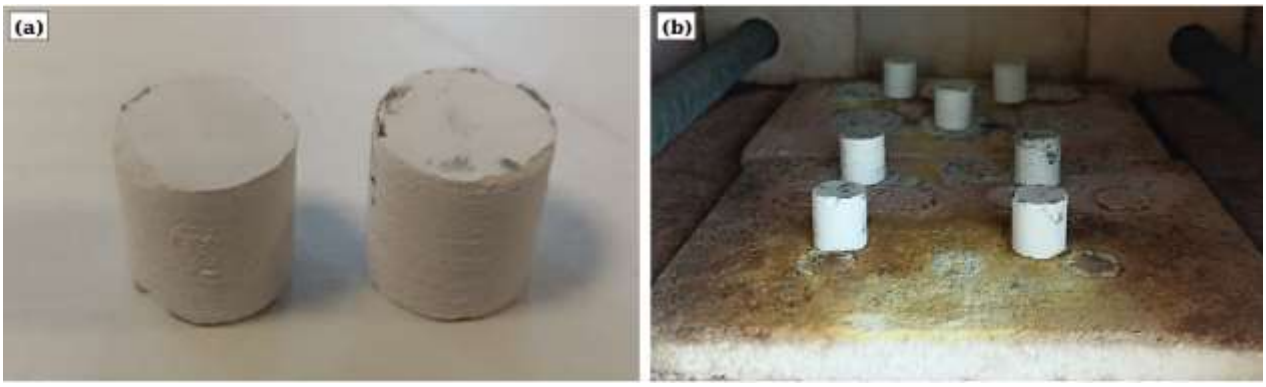


Fig. 1 (a) Representative as-prepared SECAR 80 specimens (Ø15 × 15 mm) before compression testing. (b) Specimen batch arranged on the kiln shelf of the electric chamber furnace prior to firing.

Cement and water were mixed mechanically until a homogeneous paste was obtained. The paste was cast into cylindrical moulds of 15 mm internal diameter and consolidated by vibration casting combined with vibratory compaction, the vibration being continued until no further evolution of surface bubbles was observed. Specimens were demoulded after setting and cured under ambient laboratory conditions before firing. The final geometry was a right cylinder of nominal diameter d_0 = 15 mm and height h = 15 mm, giving a geometric volume of 2.6507 cm³ and a loaded cross-sectional area of 176.71 mm².

Fig. 1(a) shows two representative specimens in the as-prepared condition. The cylindrical geometry, the flat end faces reproduced from the mould and the uniform pale colour characteristic of a hydrated high-alumina cement body are visible; these end faces are the surfaces subsequently loaded in compression.

C. Heat Treatment and Experimental Design

Specimens were fired in an electric chamber furnace at either 900 °C or 1200 °C. The two temperatures lie on opposite sides of the reported hydraulic-to-ceramic bond transition of high-alumina CAC systems, generally placed in the vicinity of 1000 °C [8, 9]: at 900 °C the hydrates are fully decomposed but solid-state sintering has not meaningfully begun, whereas 1200 °C lies within the domain in which sintering of CAC-derived phases is reported to initiate [10]. Fig. 1(b) shows a batch arranged on the kiln shelf before firing; specimens were spaced to allow free circulation of the furnace atmosphere around each cylinder.

The resulting design is given in Table 2. It is arranged so that each process variable is isolated by a single pairwise comparison. Groups S30-900 and S40-900 share a firing temperature and differ only in w/c ratio; groups S40-900 and S40-1200 share a w/c ratio and differ only in firing temperature.

Table 2. Experimental design and specimen matrix

Group	w/c	Water (g per 100 g cement)	Firing temperature
S30-900	0.30	30	900 °C
S40-900	0.40	40	900 °C
S40-1200	0.40	40	1200 °C

D. Characterisation

Specimens were weighed before and after firing to an accuracy of 0.01 g. Green (as-cast) and fired bulk densities were obtained by the geometric method, dividing specimen mass by the nominal geometric volume. Mass loss on firing was calculated from the difference between pre- and post-firing masses.

Apparent porosity was calculated from the fired bulk density and an assumed true density of the fired phase assemblage, taken as 3.10 g cm^{-3} for the 900 °C bodies (a mixture of $C_{12}A_7$, CA, CA_2 and residual α-Al_2O_3) and 3.15 g cm^{-3} for the 1200 °C body, where further conversion toward CA_2 and the onset of denser aluminate phases is expected:

$$P = (1 - \rho_{\mathbf{bulk}} / \rho_{\mathbf{true}}) \times 100 \quad (1)$$

This is a calculated total porosity based on a geometric bulk density and an assumed true density, not an Archimedes apparent porosity measured by liquid displacement. It therefore includes closed porosity and is sensitive to the assumed true density; the consequences are addressed in Section IV.E.

To separate the two mechanisms by which mixing water acts, the theoretical air-free density of each fresh paste was computed from a volumetric mass balance:

$$\rho_{\mathbf{paste,theor}} = (1 + w) / (1/\rho_{\mathbf{cement}} + w/\rho_{\mathbf{water}}) \quad (2)$$

where the water-to-cement ratio is by mass. The entrained air content of the consolidated green body follows as the shortfall of the measured green density below this theoretical value, and measures how effectively vibratory compaction removed air independently of how much water was present.

Cold crushing strength was determined by uniaxial compression of the cylinders. Group means, standard deviations and coefficients of variation were computed from individual specimen values. Finally, the Ryshkewitch relation [5] was used to normalise strength for porosity:

$$\sigma = \sigma_0 \exp(-bP) \quad (3)$$

where σ_0 is the extrapolated strength at zero porosity, P the pore volume fraction and b an empirical constant. The constants were fitted to the two groups fired at 900 °C, which share a thermal history and therefore a bonding state. The relation so obtained defines a 900 °C isotherm against which the 1200 °C body can be compared: any strength in excess of the isotherm prediction cannot be attributed to porosity and must arise from a change in the nature of the bond itself.

III. RESULTS

A. Cold Crushing Strength

Individual and aggregated crushing strengths are given in Table 3 and displayed in Fig. 2. The three groups are clearly separated, and the ranking is not monotonic in either process variable taken alone.

Table 3. Individual and aggregated cold crushing strength results

Group	Specimen	CCS (MPa)	Group mean ± SD (MPa)	CV (%)
S30-900	1	26.9	**34.2 ± 7.0**	**20.5**
	2	40.9		
	3	34.9		
S40-900	4	14.0	**14.6 ± 3.3**	**22.5**
	5	18.2		
	6	11.7		
S40-1200	7	22.2	**27.0 ± 4.9**	**18.1**
	8	32.1		
	9	23.5		
	10	30.2		

The reference group S30-900 developed the highest mean strength at 34.2 MPa. Increasing the mixing water to w/c = 0.40 while holding the firing temperature constant reduced the mean strength to 14.6 MPa, a reduction of 57.3 %. Firing the wetter formulation at 1200 °C instead of 900 °C recovered the mean strength to 27.0 MPa, an increase of 84.5 % over its 900 °C counterpart but still 21 % below the drier reference.

Scatter was moderate and comparable across groups, with coefficients of variation between 18.1 % and 22.5 %. This is unremarkable for the compressive failure of a highly porous brittle solid, in which failure is initiated by the largest flaw sampled by each specimen, and does not indicate a systematic difference in defect population between the groups.

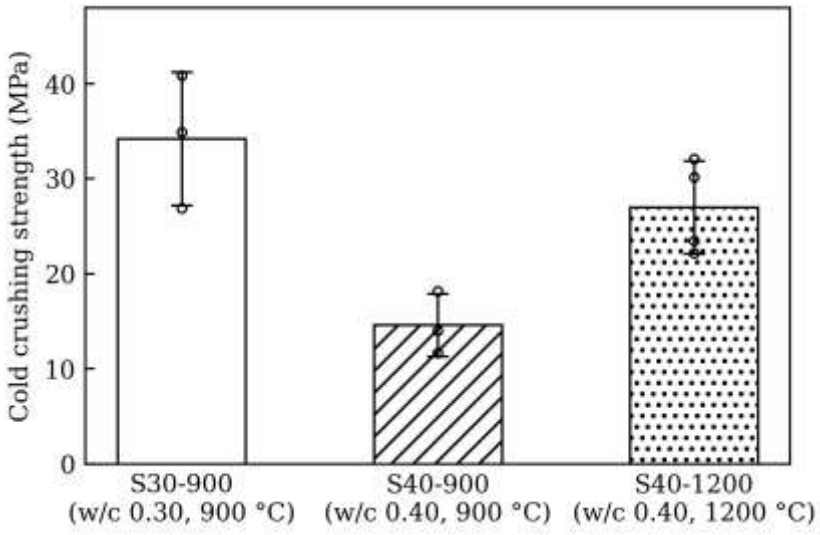


Fig. 2 Cold crushing strength of the three specimen groups. Bars show group means, error bars one standard deviation, open circles individual specimens.

The mode of failure was consistent across the specimens examined and is illustrated in Fig. 3, which shows a specimen on the lower platen immediately after testing. Failure did not occur by the propagation of a single dominant crack: no through-going shear plane, axial splitting crack or intact conical remnant of the kind associated with dense brittle cylinders was observed. Instead, the body disintegrated diffusely, producing a corona of fine powder and small angular fragments distributed radially over the platen around a partially retained core.

This comminution-dominated behaviour is what would be expected of a body in which roughly half of the volume is pore space. In such a structure the load is carried by a sparse skeleton of thin ligaments between pores; once the peak load is reached, these ligaments buckle and crush progressively over a large volume rather than allowing a single crack to run through a continuous solid. The observation is qualitative and is not by itself proof of the mechanism, but it is consistent with the high apparent porosities reported in Section III.B and with the steep porosity sensitivity of strength established in Section IV.C.



Fig. 3 A specimen on the lower platen immediately after compression testing, showing diffuse comminution into fine powder and angular fragments around a partially retained core rather than failure on a single fracture plane.

B. Density, Mass Loss and Porosity

Fired bulk density followed the same ranking as strength. The drier reference body reached 1.660 g cm^{-3} against 1.332 g cm^{-3} for the wetter body fired at the same temperature, a difference of 19.8 %. Calculated apparent porosity rose correspondingly from 46.5 % to 57.0 %.

Physical characteristics are collected in Table 4 and illustrated in Fig. 4.

Table 4. Green and fired physical properties

Group	Green ρ (g cm⁻³)	Theor. air-free ρ (g cm⁻³)	Entrained air (vol.%)	Fired ρ (g cm⁻³)	Mass loss (%)	Apparent porosity (%)
S30-900	2.150	2.139	≈ 0	1.660	22.8	46.5
S40-900	1.705	1.978	**13.8**	1.332	21.9	57.0
S40-1200	1.943	1.978	1.8	1.422	26.8	54.8

Mass loss on firing was similar for the two bodies fired at 900 °C, at 22.8 % and 21.9 %, consistent with essentially complete dehydration of the hydrate assemblage in both cases. The 1200 °C body lost appreciably more mass, 26.8 %, indicating that decomposition continued beyond the point reached at 900 °C. Despite this larger mass loss, the 1200 °C body was denser than its 900 °C counterpart, 1.422 against 1.332 g cm⁻³; volumetric shrinkage on firing therefore outpaced mass loss, which is the signature of incipient sintering rather than of simple dehydration.

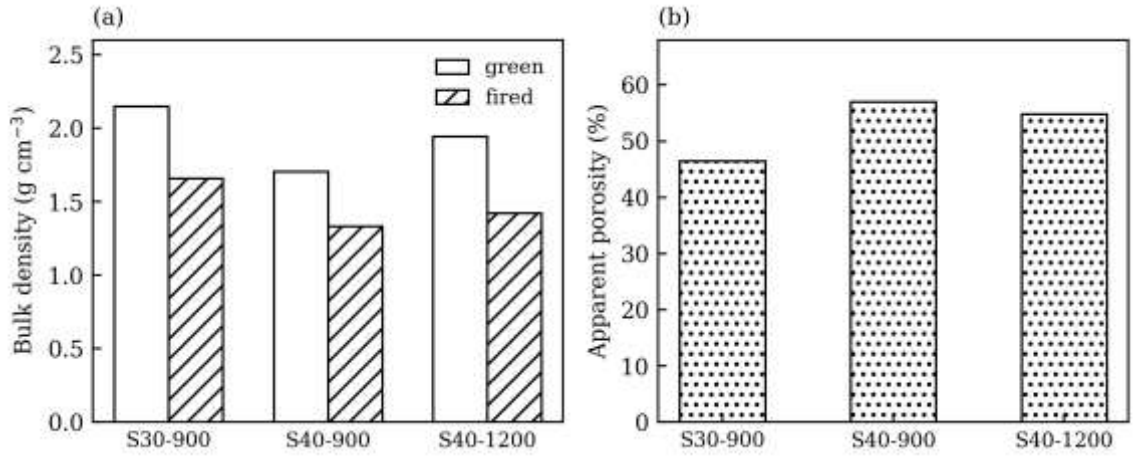


Fig. 4 (a) Green and fired bulk density, and (b) calculated apparent porosity of the three specimen groups.

C. *Green-State Mass Balance*

The green-state analysis, shown in Fig. 5, reveals a feature that the fired properties alone would conceal. At w/c = 0.30 the measured green density of 2.150 g cm⁻³ is indistinguishable from the theoretical air-free value of 2.139 g cm⁻³, implying that vibratory compaction expelled essentially all entrapped air and that the paste consolidated to its maximum attainable packing.

The inferred air content for this group is in fact marginally negative, −0.5 vol.%, meaning that the measured green density slightly exceeded the theoretical air-free value. A negative air content has no physical meaning; the result lies within the uncertainty of the assumed cement true density of 3.25 g cm⁻³ and implies that the true value is a little higher than assumed. It is reported as approximately zero, and the same offset would apply equally to the two w/c = 0.40 batches, so the comparison between groups is unaffected.

At w/c = 0.40 the two batches behaved very differently from one another. The batch fired at 1200 °C had a green density of 1.943 g cm⁻³ against a theoretical 1.978 g cm⁻³, corresponding to only 1.8 vol.% entrained air and therefore to successful de-airing. The batch fired at 900 °C, however, reached only 1.705 g cm⁻³, a shortfall of 13.8 vol.%. Since both batches were nominally identical in composition and were processed by the same route, this difference reflects batch-to-batch variability in the effectiveness of the vibration step at the higher water content, rather than a systematic consequence of the water content itself.

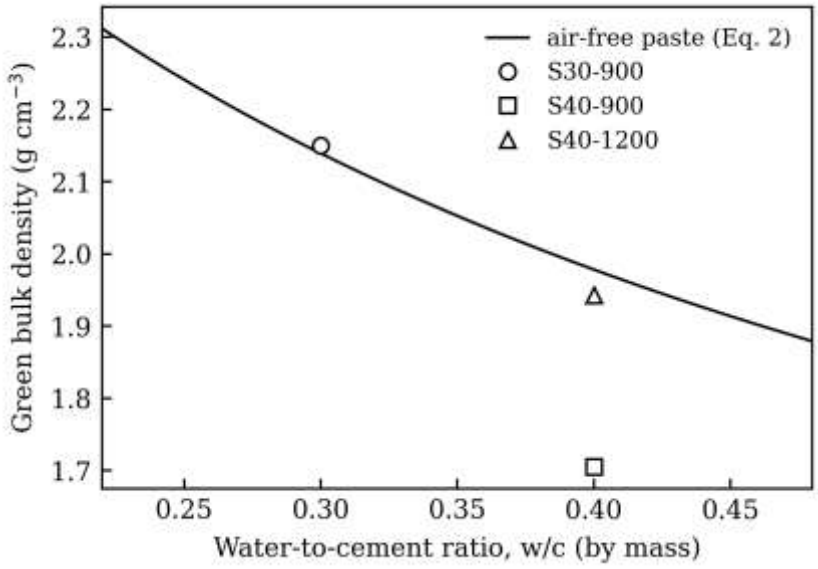


Fig. 5 Measured green bulk density of the three batches compared with the theoretical air-free paste density as a function of w/c ratio.

IV. Discussion

The comparison of S30-900 with S40-900 isolates the effect of mixing water at a fixed thermal history. A ten-point increase in w/c ratio, from 0.30 to 0.40, cost 57.3 % of the crushing strength. This is a very large penalty for what a producer would regard as a modest adjustment made purely to improve placing consistency, and it is the central practical finding of this work.

The magnitude becomes understandable once the two superimposed mechanisms are separated. The first is the direct capillary mechanism. Water more than the hydration demand occupies volume in the fresh paste and is expelled during drying and firing, leaving a connected capillary network [3]. This mechanism alone accounts for the difference between the theoretical air-free densities of the two pastes, 2.139 against 1.978 g cm^{-3}, and is unavoidable at a given w/c ratio.

The second mechanism is indirect and acts through the forming step. The green-state balance in Section III.C shows that the 0.30 paste consolidated air-free, whereas the wetter paste in one batch retained 13.8 vol.% entrained air despite receiving the same vibration treatment. This is consistent with the paste having become too fluid to be stabilised by vibration: bleeding and segregation permit air to be re-entrained or to remain trapped as coarse voids rather than being expelled. The distinction matters mechanically, because coarse entrained voids are not equivalent to an equal volume of fine capillary porosity. Fine, uniformly distributed capillary pores reduce the load-bearing cross-section, whereas coarse voids act as discrete stress concentrators and as pre-existing critical flaws; clustered or coarse pore populations are known to depress ceramic strength more severely than the same volume fraction of dispersed porosity [4].

Taken together, these observations indicate that the practical importance of controlling the water-to-solid ratio in ceramic body preparation is understated if it is framed only as a matter of final porosity. Excess water compromises the forming step itself, and the resulting defect population is qualitatively worse than the porosity figure alone would suggest.

The comparison of S40-900 with S40-1200 isolates the effect of firing temperature. Raising the temperature from 900 °C to 1200 °C increased the mean strength by 84.5 %, from 14.6 to 27.0 MPa, while the calculated apparent porosity fell by only 2.2 percentage points, from 57.0 % to 54.8 %. The disproportion between these two figures is the key observation: a strength change of this size cannot be a densification effect.

This behaviour is consistent with the established thermal sequence of CAC systems. Hydrate decomposition is essentially complete well below 900 °C, so a body fired at 900 °C consists of dehydrated hydrate pseudomorphs and residual clinker and alumina particles held together by little more than mechanical interlocking. High-alumina CAC are reported to pass through a strength minimum in this

intermediate range, with the recovery of strength and the transition from hydraulic to ceramic bonding occurring in the vicinity of 1000 °C [8, 9, 11, 12]. At 1200 °C the body lies above this transition and within the domain in which solid-state sintering of CAC-derived phases initiates [10], so that diffusional neck growth restores continuity across former particle boundaries.

It is worth stating what this study does not claim. Calcium hexaluminate (CA_6) formation, often invoked to explain strength development in fired CAC systems, is generally reported to require temperatures of approximately 1300–1400 °C, with only small amounts appearing at 1300 °C [10, 13–15]. At 1200 °C the strength gain observed here is therefore more plausibly attributed to sintering neck formation and to continued conversion within the $CA–CA_2–Al_2O_3$ assemblage than to CA_6 crystallisation. No phase analysis was performed, and this interpretation remains to be confirmed by X-ray diffraction.

The larger mass loss recorded for the 1200 °C body, 26.8 % against 21.9 %, supports the view that transformation continued beyond the state reached at 900 °C. That the body nevertheless became denser, despite losing more mass, requires that volumetric shrinkage exceeded the mass loss, which is the expected signature of sintering.

The two effects can be separated quantitatively by using porosity as the normalising variable. Fitting Eq. 3 to the two groups fired at 900 °C, which share a bonding state, gives $b = 8.03$ and $\sigma_0 = 1425$ MPa. The value of b lies within the range commonly reported for porous ceramics [1] and confirms the steep exponential sensitivity of strength to pore volume in this system.

Applying this 900 °C isotherm to the porosity of the 1200 °C body, P = 54.8 %, predicts a strength of 17.4 MPa. The measured value of 27.0 MPa exceeds this by 55 %, as shown in Fig. 6. Because normalisation has already absorbed the porosity difference between the two 40 % bodies, including that arising from their differing green densities, the excess cannot be explained by densification and must be assigned to the change in the character of the bond.

This treatment also addresses a potential objection to the experimental design. The two w/c = 0.40 batches did not have identical green densities: the batch fired at 1200 °C was better consolidated, at 1.943 against 1.705 g cm^{-3}. Comparing their fired strengths directly would therefore confound the temperature effect with a compaction difference. The Ryshkewitch normalisation removes this confound, because the superior consolidation of the 1200 °C batch is expressed in its lower fired porosity and is fully credited to the porosity term of Eq. 3 before the residual is attributed to bonding. The 55 % excess is what remains after that correction.

Two consequences follow for the preparation of CAC-bonded ceramic bodies. First, the water-to-solid ratio should be treated as the primary design variable and minimised to the lowest value consistent with adequate flow, since its influence on strength is both large and, once the body is cast, irreversible. Where a lower water content would compromise castability, the appropriate response is to restore flow by other means, for example through deflocculant addition or particle size distribution optimisation, rather than by adding water.

Second, firing above the hydraulic-to-ceramic bond transition offers a substantial but incomplete remedy. Raising the firing temperature to 1200 °C recovered a body cast at w/c = 0.40 to 27.0 MPa, which is 79 % of the strength achieved by a body cast at w/c = 0.30 and fired at only 900 °C. The additional thermal energy therefore did not fully compensate for the excess water, and the drier body would be expected to benefit at least as much from the same treatment. Firing temperature and water content are not interchangeable levers.

This is a preliminary study, and its conclusions should be read with the following constraints in mind. The experimental matrix is incomplete: the fourth cell of the design, w/c = 0.30 fired at 1200 °C, was not produced, which prevents a formal test of whether the water and temperature effects interact. Closing this cell is the single most valuable extension of the work.

Group sizes are small, at three to four specimens, and are insufficient for a Weibull analysis of the strength distribution or for statistical hypothesis testing at conventional confidence levels. The reported differences are large relative to the scatter but have not been tested formally.

The specimen geometry, a cylinder of 15 mm diameter with a height-to-diameter ratio of unity, is smaller than the geometries prescribed by standard refractory crushing tests and has an aspect ratio at which platen friction constrains lateral expansion. Absolute strength values are therefore likely to be

higher than would be measured on standard specimens and should be regarded as comparative rather than as design values.

Porosity was calculated from geometric bulk density and an assumed true density rather than measured by liquid displacement. The values reported are total porosities and are sensitive to the assumed true density; a change of ±0.05 g cm^{-3} in that assumption shifts the calculated porosity by roughly one percentage point.

No phase or microstructural analysis was performed. The attribution of the 1200 °C strength excess to sintering neck formation rests on the thermal behaviour reported in the literature and on the density–mass-loss relationship, not on direct evidence. X-ray diffraction to follow the CA–CA_2–CA_6 sequence and scanning electron microscopy to observe neck formation and pore morphology are the natural next steps. Finally, the present work characterises the cement matrix in isolation; extension to silicon carbide–calcium aluminate cement composites, in which the SiC aggregate would alter both the water demand of the mix and the sintering behaviour of the matrix, is planned as a continuation of this investigation.

V. Conclusion

Pure SECAR 80 calcium aluminate cement bodies were prepared at two water-to-cement ratios by vibration casting with vibratory compaction and fired at two temperatures, in a design permitting the two process variables to be isolated. Raising the water-to-cement ratio from 0.30 to 0.40 at a constant firing temperature of 900 °C reduced the mean cold crushing strength by 57.3 %, from 34.2 ± 7.0 MPa to 14.6 ± 3.3 MPa and increased the calculated apparent porosity from 46.5 % to 57.0 %.

A green-state mass balance showed that the w/c = 0.30 paste consolidated essentially air-free, whereas one w/c = 0.40 batch retained 13.8 vol.% entrained air under identical vibration. Excess mixing water therefore damages strength through two mechanisms: the capillary porosity left by evaporated free water, and the impaired vibratory de-airing of an over-fluid, segregation-prone paste.

At constant w/c = 0.40, raising the firing temperature from 900 °C to 1200 °C increased the mean strength by 84.5 %, from 14.6 to 27.0 MPa, while apparent porosity fell by only 2.2 percentage points; the strength gain is therefore not a densification effect. Porosity alone accounts for only 17.4 MPa of the strength of the 1200 °C body, whereas 27.0 MPa was measured; the 55 % excess indicates a shift from a dehydrated hydraulic bond to a sintered ceramic bond.

Firing above the bond transition partially but not fully compensates for an excessive water content: the wet body fired at 1200 °C reached only 79 % of the strength of the dry body fired at 900 °C. Minimising mixing water remains the dominant lever on the mechanical performance of CAC-bonded ceramic bodies.

References

1. Engbert, A., Plank, J.: Impact of sand and filler materials on the hydration behavior of calcium aluminate cement. Journal of the American Ceramic Society. 104, 1067–1075 (2021). https://doi.org/10.1111/jace.17505

2. Mishra, S., Chandra, K.S., Sarkar, D., Su, S.N., Chakraborty, J.N., Agarwalla, A.: Effect of Calcium Aluminate phases in High Alumina Cements and Low Cement Castable. Journal of Alloys and Compounds Communications. 6, 100052 (2025). https://doi.org/10.1016/j.jacomc.2024.100052

3. Gołaszewska, M., Gołaszewski, J., Chmiela, B.: Effect of Water to Cement Ratio on Properties of Calcium Sulfoaluminate Cement Mortars. Materials. 17, 2806 (2024). https://doi.org/10.3390/ma17122806

4. Dorey, R.A., Yeomans, J.A., Smith, P.A.: Effect of pore clustering on the mechanical properties of ceramics. J. Eur. Ceram. Soc. 22, 403–409 (2002). https://doi.org/10.1016/S0955-2219(01)00303-X

5. RYSHKEWITCH, E.: Compression Strength of Porous Sintered Alumina and Zirconia. Journal of the American Ceramic Society. 36, 65–68 (1953). https://doi.org/10.1111/j.1151-2916.1953.tb12837.x

6. Nielsen, L.F.: Strength and Stiffness of Porous Materials. Journal of the American Ceramic Society. 73, 2684–2689 (1990). https://doi.org/10.1111/j.1151-2916.1990.tb06746.x

7. SECAR® The original binder for high temperature applications, Imerys Aluminates, product range technical literature. Available: https://www.imerys.com/product-ranges/secar

8. Gao, S., Zhang, P., Li, N., Zhang, J., Luan, J., Ye, G., Liao, G.: Effect of CAC content on the strength of castables at temperatures between 300 and 1000 °C. Ceram. Int. 46, 14957–14963 (2020). https://doi.org/10.1016/j.ceramint.2020.03.024

9. Abolhasani, A., Samali, B., Aslani, F.: Physicochemical, Mineralogical, and Mechanical Properties of Calcium Aluminate Cement Concrete Exposed to Elevated Temperatures. Materials. 14, 3855 (2021). https://doi.org/10.3390/ma14143855

10. Tomba Martinez, A.G., Luz, A.P., Braulio, M.A.L., Sako, E.Y., Pandolfelli, V.C.: Revisiting CA6 formation in cement-bonded alumina-spinel refractory castables. J. Eur. Ceram. Soc. 37, 5023–5034 (2017). https://doi.org/10.1016/j.jeurceramsoc.2017.07.003

11. Lee, J.-M., Yang, H.J., Jang, I.-Y., Kim, S.K., Jung, D.-H.: Comparison of calcium aluminate cements on hydration and strength development at different initial curing regimes. Case Studies in Construction Materials. 17, e01596 (2022). https://doi.org/10.1016/j.cscm.2022.e01596

12. Kafali, A., Karaca, M.M., Polat, S.: Performance enhancement of cement matrix with B_4C, SiC, and hBN additives under thermal shock, mechanical stress, and wear conditions. J. Compos. Mater. 60, 1433–1447 (2026). https://doi.org/10.1177/00219983251384013

13. Chen, J., Chen, H., Yan, M., Cao, Z., Mi, W.: Formation mechanism of calcium hexaluminate. International Journal of Minerals, Metallurgy, and Materials. 23, 1225–1230 (2016). https://doi.org/10.1007/s12613-016-1342-9

14. Lu, H., Li, Y., Liu, D., Wen, J., Sun, D., Jin, E., Chen, J., Ding, D., Xiao, G.: Preparing porous calcium hexaluminate (CA6) ceramics with calcium aluminate cement (CAC) as the gelling agent and sole calcia source. J. Alloys Compd. 1010, 177625 (2025). https://doi.org/10.1016/j.jallcom.2024.177625

15. Xia, Y., Ding, C., Luo, W., Yang, H., Yuan, W.: Optimization of Properties of Calcium Hexaluminate-Based Insulating Castables with Calcium Aluminate Cement. Materials. 18, 2354 (2025). https://doi.org/10.3390/ma18102354

6th International Conference on Modern and Advanced Research

August 16-17, 2026: Konya, Turkey

© 2026 Published by All Sciences Academy

https://as-proceeding.com/index.php/icmar/home

KÜLTÜREL MİRAS ALANLARINDA ZİYARETÇİ DENEYİMLERİ

Ozan Çatır

Usak University

ozan.catir@usak.edu.tr

Özet - Kültürel miras alanları, kentlerin geçmişini günümüze taşımanın yanı sıra ziyaretçilerin mekânla kültürel, görsel ve duygusal ilişkiler kurmasına olanak sağlayan önemli turistik çekiciliklerdir. Bu araştırmada, İstanbul'daki Galata Kulesi ile Bakü'deki Maiden (Kız) Kulesine ilişkin ziyaretçi deneyimlerinin karşılaştırılması ve ortaya çıkan deneyim örüntülerinin yenileyici turizm çerçevesinde değerlendirilmesi amaçlanmıştır. Araştırmanın verileri, ziyaretçilerin TripAdvisor platformunda paylaştıkları çevrimiçi yorumlardan elde edilmiştir. Yorumların incelenmesinde kelime frekansı, konu modelleme ve duygu analizi gibi metin madenciliği tekniklerinden yararlanılmıştır. Bulgular, kültürel miras alanlarına ilişkin deneyimlerin tarihsel ve kültürel özelliklerin yanı sıra mekânsal algılar, destinasyonla kurulan bağ ve duygusal değerlendirmeler üzerinden şekillendiğini göstermektedir. Galata Kulesine ilişkin anlatılarda manzara, panoramik görünüm ve mekânsal deneyim belirginleşirken, Maiden Kulesinde tarihsel özellikler, kültürel kimlik ve kentsel simgesellik daha fazla öne çıkmıştır. Her iki kültürel miras alanına ilişkin değerlendirmelerde olumlu ifadelerin ağırlıkta olduğu görülmüştür. Elde edilen bulgular, kültürel miras alanlarının ziyaretçiler açısından yalnızca görülmesi gereken turistik çekicilikler olarak değil, kentle farklı biçimlerde ilişki kurulan deneyim alanları olarak algılandığını ortaya koymaktadır. Araştırma, bu deneyimlerin yenileyici turizm bağlamında ele alınmasına yönelik ziyaretçi temelli bir değerlendirme sunmaktadır.

Anahtar Kelimeler: Kültürel miras, ziyaretçi deneyimi, yenileyici turizm, metin madenciliği, çevrimiçi yorumlar

1. GİRİŞ

Ülkelerin kalkınması, gelirlerin arttırılması, istihdam ve kültürel etkileşim açısından turizm sektörü önemli bir yere sahiptir. Fakat turizm faaliyetlerinin artmasıyla birlikte destinasyonlarda çevresel bozulmayı, sosyal eşitsizlikleri ve kültürel aşınmayı etkilemektedir. Son yıllarda yapılan çalışmaları, sürdürülebilir turizm yaklaşımının sorgulanması gerektiğini vurgulamaktadır (Bellato vd., 2023; Amorim vd., 2025). Turizm sistemlerinin yenileyici bir bakış açısıyla tekrar geliştirilmesi gerektiği vurgulanmıştır (Bellato ve Pollock, 2025). Literatürdeki bu tartışmalar, turizmin bütüncül bir yaklaşımla çevresel ve toplumsal sistem üzerindeki etkilerinin değerlendirilmesi gerektiğini göstermektedir.

Yenileyici turizm yaklaşımı hem yerel halk ve toplumun kalkınmasını güçlendirmeyi amaçlayan hem de çevresel sistemlerin zarar görmesini en aza indiren bir turizm yaklaşımıdır. Turizm destinasyonlarını karmaşık bir sistem olarak ele alan bu yaklaşım, turizmin destinasyona sağlıklı bir şekilde etkilemesi gerektiğini savunmaktadır (Bellato vd., 2022; Bellato vd., 2023). Yenileyici turizm ekonomik getirisinin yanı sıra, yerel topluluklara, turizm işletmelerine, doğaya ve ziyaretçi deneyimine etki eden çok yönlü etkileşimlerin sonucudur (Iddawala ve Lee, 2025). Bu turizm yaklaşımı çevresel bütünlüğü destekleyen, kültürel değerleri güçlendiren ve toplum refahını geliştiren bir turizm bakış açısının geliştirilmesi gerektiğini vurgulamaktadır (Bellato ve Pollock, 2025).

Yenileyici turizm yaklaşımı birçok turizm çeşidi ile ilişkili olarak literatürde incelenmiştir. En fazla ilişkilendirilen turizm çeşitleri doğa temelli turizm, macera turizmi ve kırsal turizm çeşitleridir. Bu tür turizm çeşitleri uygulanırken, yerel kalkınmaya, topluluk katılımına ve çevresel korumaya dikkat edilmesi gerektiği çalışmalarda vurgulanmıştır (Schmidt Rojas, Sand ve Gross, 2025; Zurbarán-Arias ve Piñeros, 2025; Amorim vd., 2025). Kentsel turizm alanlarında da yenileyici turizm yaklaşımı uygulanabilir. Şehirlerde bulunan tarihi yapılar ve kültürel miras alanları ziyaretçilerin destinasyonla kurduğu deneyimsel ilişkiyi güçlendiren unsurlardır (Li, vd., 2025; Amorim vd., 2025).

Yenileyici turizm yaklaşımı ile ilgili mevcut çalışmalar daha çok kavramsal tartışmalar ve teorik çerçeveler üzerinde yoğunlaşmaktadır. Yeni gelişmekte olan bu kavramın destinasyon yönetimi ve turizm planlaması uygulamalarında nasıl kullanılacağı henüz yeterince incelenmemiştir (Bellato ve Pollock, 2025; Iddawala ve Lee, 2025). Özellikle ziyaretçilerin destinasyon algıları ve deneyimleri üzerinde sınırlı sayıda ampirik çalışmanın olması literatürde önemli bir araştırma boşluğu olduğunu göstermektedir. Ziyaretçi deneyimlerinin analizi turizm destinasyonlarının nasıl algılandığını anlamak ve yenileyici turizm yaklaşımlarının ziyaretçi perspektifinden nasıl değerlendirilebileceğini ortaya koymak açısından önemli bir araştırma alanı sunmaktadır.

Son yıllarda dijital platformların yaygınlaşmasıyla birlikte turistlerin seyahat deneyimlerini çevrimiçi ortamlarda paylaşmaları turizm araştırmaları için önemli veri kaynakları oluşturmuştur. Özellikle TripAdvisor gibi çevrimiçi seyahat platformlarında yer alan kullanıcı yorumları turistlerin destinasyon deneyimlerini ve algılarını doğrudan yansıtan metinsel veriler sunmaktadır. Bu tür verilerin metin madenciliği yöntemleri kullanılarak analiz edilmesi ziyaretçi deneyimlerinde öne çıkan temaların ve algıların sistematik biçimde ortaya çıkarılmasına olanak sağlamaktadır.

Bu çalışma, İstanbul'da bulunan Galata Kulesi ile Bakü'de bulunan Maiden (Kız) Kulesi ziyaretçi yorumlarını metin madenciliği yöntemleri kullanarak analiz etmeyi amaçlamaktadır. Araştırmada TripAdvisor platformunda yer alan ziyaretçi yorumları kelime frekansı analizi, konu modelleme ve duygu analizi teknikleri kullanılarak incelenmiş ve ziyaretçi deneyimlerinde öne çıkan temalar belirlenmiştir. Çalışmanın temel amacı kültürel miras alanlarının ziyaretçi deneyimlerinin yenileyici turizm perspektifi doğrultusunda nasıl değerlendirilebileceğini ortaya koymaktır.

Yenileyici turizm literatüründe ziyaretçi deneyimlerinin ve turist algılarının ampirik olarak incelenmesine yönelik çalışmaların sınırlı olması, özellikle kültürel miras alanları bağlamında yeni araştırmaların yapılmasını gerekli kılmaktadır. Bu araştırma, çevrimiçi ziyaretçi yorumlarını metin madenciliği yöntemleri kullanarak analiz etmeyi ve elde edilen bulguları yenileyici turizm perspektifi doğrultusunda değerlendirmeyi amaçlamaktadır.

Bu doğrultuda araştırma aşağıdaki sorulara yanıt aramaktadır:

AS1: Galata Kulesi ve Maiden (Kız) Kulesi ziyaretçi yorumlarında öne çıkan temel deneyim temaları nelerdir?

AS2: Ziyaretçi yorumlarından elde edilen temalar yenileyici turizm boyutları (kültürel miras algısı, mekânsal deneyim, yerel bağ kurma ve dönüştürücü deneyim) ile nasıl ilişkilendirilebilir?

AS3: Galata Kulesi ve Maiden (Kız) Kulesi ziyaretçi deneyimleri metin madenciliği bulguları açısından hangi yönlerden benzerlik ve farklılık göstermektedir?

Bu çalışmanın literatüre üç temel katkı sağlaması beklenmektedir. İlk olarak çalışma, yenileyici turizm literatüründe sınırlı sayıda bulunan ampirik araştırmalara katkı sağlayarak ziyaretçi deneyimlerinin yenileyici turizm perspektifi ile nasıl değerlendirilebileceğini ortaya koymaktadır. İkinci olarak çalışma, kültürel miras alanlarının ziyaretçi deneyimlerindeki rolünü inceleyerek kentsel turizm bağlamında yenileyici turizm tartışmalarına katkı sunmaktadır. Üçüncü olarak ise çalışma, turizm araştırmalarında metin madenciliği yöntemlerinin kullanımına yönelik metodolojik bir örnek sunarak çevrimiçi ziyaretçi yorumlarının analizine dayalı yeni araştırma yaklaşımlarına katkı sağlamaktadır

Bu çalışma dört bölümden oluşmaktadır. İlk bölümde çalışmanın teorik arka planı ele alınarak yenileyici turizm yaklaşımı ve turizm literatüründeki yeri tartışılmıştır. İkinci bölümde araştırma yöntemi açıklanmış, veri toplama süreci ve metin madenciliği teknikleri ayrıntılı olarak sunulmuştur. Üçüncü bölümde Galata Kulesi ve Maiden (Kız) Kulesi ziyaretçi yorumlarına ilişkin metin madenciliği analizlerinden elde edilen bulgular karşılaştırmalı olarak değerlendirilmiştir. Son bölümde ise araştırma bulguları yenileyici turizm perspektifi doğrultusunda tartışılmış ve çalışmanın teorik ve pratik katkıları ele alınmıştır.

2. Literatür Taraması

Sürdürülebilir turizm ile ilgili yaklaşımların sorgulanmaya başlamasıyla yeni bir paradigma olarak yenileyici turizm yaklaşımı ortaya çıkmıştır. Bu yaklaşım destinasyonun toplumsal kültürel ve ekolojik olarak güçlendirmeyi amaçlayan bir yaklaşımdır (Bellato vd., 2023; Dredge, 2022; Cave ve Dredge, 2021; Sheldon, 2022; Yeoman ve Sheldon, 2024). Alan yazınında yenileyici turizm bir dönüşüm aracı olarak konumlandırılmıştır. Turizm sistemleri bu yaklaşımda sosyo-ekolojik sistemler olarak değerlendirilmekte ve turizmle ilgili faaliyetlerin doğa ile uyumlaştırılması gerektiği vurgulanmıştır (Bellato vd., 2022; Duarte vd., 2024; Manthiou vd., 2025; Forte ve Mariotti, 2024). Yenileyici turizm yaklaşımı kültürel süreklilik toplumsal kalkınma ve ekolojik dayanıklılık boyutlarında ele alınan bir yaklaşımdır (Bellato, 2025; Dredge, 2022; Cave ve Dredge, 2021).

Geleneksel turizm yaklaşımları turizmi ekonomik kalkınmayı sağlayan bir unsur olarak görmektedir. Yenileyici turizm ise turizm faaliyetlerinin topluluklar, çevresel sistemler ve ekonomik süreçlerin etkileşim içinde geliştiği bir sistemi ortaya çıkarmaktadır (Bellato vd., 2022; Cave ve Dredge, 2022; Hall vd., 2025; Meneghello, 2023). Yenileyici turizm ekonomik getiri sağlarken, toplumsal ve çevresel süreçleri de desteklemektedir (Sheldon, 2022; Yeoman ve Sheldon, 2024). Yenileyici turizm yaklaşımında paydaşların katılımı ve yerel yönetişim turizm planlama süreçlerinde gereklidir. Turizm gelirlerinin adil dağıtılması ve sosyal dayanıklılık için paydaşların katılımı desteklenmektedir (Scheyvens ve Kaire Gataulu, 2025; Becken ve Kaur, 2021; Boluk ve Panse, 2022; Zurbarán-Arias ve Piñeros, 2025). Turizm sektöründe güç ilişkileri yerel katılımı bazı durumlarda engelleyebilir (Soma vd., 2025; Lemarié ve Bellato, 2025). Bu nedenle turizm planlamasında sosyal adalet, kapsayıcı yönetişim ve bilgi sistemleri planlamasında dikkat edilmesi gereken konulardır (Mandić ve Rastegar, 2025; Higgins-Desbiolles, 2024).

Yenileyici turizm yaklaşımı ile kültürel miras destinasyonları ilişkisi üzerine çalışmalar yapıldığı görülmektedir. Önceki çalışmalar doğa temelli ve kırsal turizm konularına odaklansa da, son yıllarda yapılan çalışmalar kentsel turizm ve kültürel miras alanları üzerine odaklanmıştır (Lemarié ve Bellato, 2025; Li vd., 2025; Levy ve Shani, 2025). Kentsel kültürel miras alanları ziyaretçilerin kollektif hafızasına ve tarihsel kimliğini temsil eden çekiciliklerdir Meneghello, 2023; Hall vd., 2025). Bu nedenle kültürel miras destinasyonlarının ziyaretçi deneyimleri üzerinden analiz edilmesi turizm faaliyetlerinin kültürel değerlerle nasıl ilişkilendirildiğini anlamak açısından önemli bir araştırma alanı oluşturmaktadır. Özellikle tarihi yapılar ve simgesel mimari unsurlar ziyaretçilerin destinasyonla kurduğu duygusal ve bilişsel bağın önemli bir parçası olarak görülmektedir (Li vd., 2025; Levy ve Shani, 2025).

Ziyaretçilerin deneyimi ve turizm algısını inceleyen çalışmalar genellikle memnuniyet ve hizmet kalitesi konularıyla ilişkilendirilmiştir. Yenileyici turizm yaklaşımı açısından ise ziyaretçilerin çevresel farkındalığı, yer bağlılığı ve kültürel öğrenme boyutlarıyla ilişkisi ortaya konmuştur (Pung vd., 2024; Corral-Gonzalez vd., 2024). Turistlerin destinasyon deneyimlerini nasıl yorumladığı ve bu deneyimlerin çevresel ve kültürel değerlerle nasıl ilişkilendirildiği yenileyici turizm araştırmaları açısından önemli bir konu haline gelmiştir. Özellikle kültürel miras destinasyonlarında ziyaretçi deneyimlerinin analizi, turizmin kültürel değerlerin korunması ve yeniden yorumlanması üzerindeki rolünü anlamaya katkı sağlayabilmektedir (Li vd., 2025; Hall vd., 2025).

Destinasyon ziyaretçileri son yıllarda yaşamış oldukları seyahat deneyimlerini çevrimiçi platformlarda paylaşmaktadırlar. Bu platformlar turizm araştırmaları açısından önemli bir veri kaynağı olmuştur. Özellikle TripAdvisor gibi platformlar ziyaretçilerin deneyimlerini anlamayı kolaylaştırmaktadır (Omma, 2024; Lovelock vd., 2026). Bu nedenle çevrimiçi yorumların analiz edilmesi turist algılarının ve

deneyimlerinin anlaşılmasına katkı sağlayabilmektedir (Gerke vd., 2024; Hussain ve Fusté-Forné, 2026). Bu bağlamda metin madenciliği, duygu analizi ve konu modelleme gibi veri analitiği yöntemleri turizm araştırmalarında giderek daha yaygın olarak kullanılmaktadır.

Yenileyici turizm yaklaşımına yönelik çalışmalar incelendiğinde, çalışmaların önemli bir kısmının kavramsal tartışmalara odaklandığı görülmektedir. Ziyaretçi deneyimlerinin yenileyici turizm açısından nasıl algılandığını inceleyen çalışmaların sınırlı olduğu görülmektedir (Bellato ve Pollock, 2025; Yeoman ve Sheldon, 2024). Ayrıca mevcut çalışmaların büyük bölümü doğa temelli turizm destinasyonlarına odaklanırken, kentsel kültürel miras alanlarında ziyaretçi deneyimlerini inceleyen çalışmaların sınırlı olduğu görülmektedir (Lemarié ve Bellato, 2025; Li vd., 2025). Yöntemsel olarak da metin madenciliği yöntemi ile yorumların incelenmesi yenileyici turizm literatüründe henüz yaygın değildir (Omma, 2024; Gerke vd., 2024).

Bu araştırma, İstanbul'daki Galata Kulesi ve Bakü'deki Maiden (Kız) Kulesi ile ilgili TripAdvisor yorumlarını metin madenciliği teknikleri kullanarak analiz etmektedir. Çalışma ziyaretçi deneyimlerinde öne çıkan temaları belirlemeyi ve bu temaları yenileyici turizm perspektifi doğrultusunda yorumlamayı amaçlamaktadır. Bu yönüyle araştırma yenileyici turizm literatürüne kültürel miras destinasyonları, ziyaretçi deneyimleri ve dijital veri analizi bağlamında katkı sunmaktadır.

2.YÖNTEM

Bu araştırmada veri analiz süreci, çevrimiçi kullanıcı yorumlarının sistematik biçimde incelenmesine dayanan çok aşamalı bir metin madenciliği yaklaşımı çerçevesinde gerçekleştirilmiştir. Günümüzde kullanıcı tarafından oluşturulan içerikler, turistlerin deneyimlerini, algılarını ve duygularını doğal biçimde yansıttıkları için turizm araştırmalarında önemli bir veri kaynağı olarak kabul edilmektedir. Özellikle çevrimiçi seyahat platformlarında yer alan yorumlar, ziyaretçilerin destinasyon deneyimlerini ve destinasyon algılarını anlamak açısından önemli bilgiler sunmaktadır (Zaman, 2024; Vegas-Macias vd., 2026). Bu doğrultuda araştırma kapsamında Galata Kulesi ve Maiden (Kız) Kulesi için TripAdvisor platformunda yer alan ilk 500 ziyaretçi yorumu toplanmış ve toplamda 1000 yorumdan oluşan bir veri seti oluşturulmuştur.

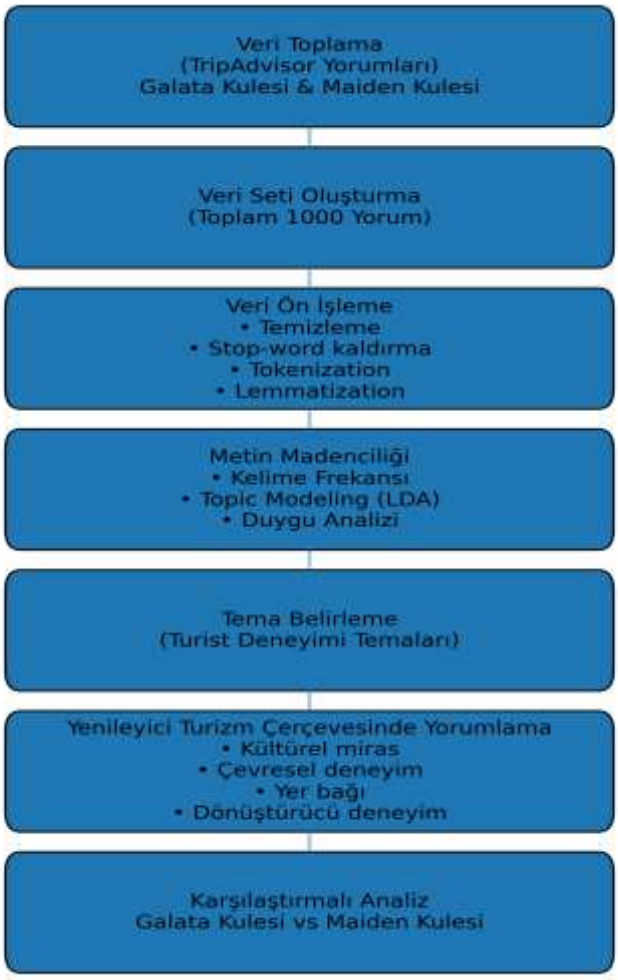
Veri analiz sürecinin ilk aşamasında metinsel veriler analiz için uygun hale getirilmek amacıyla çeşitli ön işleme adımlarından geçirilmiştir. Metin madenciliği çalışmalarında veri ön işleme süreci, analiz sonuçlarının güvenilirliği ve doğruluğu açısından kritik bir aşama olarak değerlendirilmektedir. Metin madenciliği analizi yapmak için Rapidminer veri analiz programından faydalanılmıştır. Bu kapsamda metinlerde yer alan noktalama işaretleri ve sayısal ifadeler veri setinden çıkarılmış, tüm metinler küçük harfe dönüştürülmüş ve analitik anlam taşımayan sık kullanılan kelimeler (stop words) veri setinden temizlenmiştir. Ayrıca metinler kelime temelli analiz için "tokenlere" ayrılmış ve kelimelerin kök biçimlerine indirgenmesi amacıyla "lemmatization" işlemi uygulanmıştır. Bu işlemler sayesinde veri seti standartlaştırılmış ve metin madenciliği analizine uygun hale getirilmiştir.

Ön işleme aşamasının ardından metin madenciliği teknikleri kullanılarak veri analizi gerçekleştirilmiştir. Analizin ilk aşamasında kelime frekansı analizi uygulanarak ziyaretçi yorumlarında en sık kullanılan kelimeler belirlenmiştir. Kelime frekansı analizi, ziyaretçi deneyimlerinde öne çıkan destinasyon özelliklerinin ve algıların belirlenmesine katkı sağlayan temel metin analizi yöntemlerinden biridir. Bu analiz, ziyaretçilerin destinasyon deneyimlerinde hangi unsurlara odaklandıklarını ortaya koyarak turizm deneyimlerinin anlaşılmasına katkı sağlamaktadır.

Bir sonraki aşamada metinlerdeki gizli tematik yapıları ortaya çıkarmak amacıyla Latent Dirichlet Allocation (LDA) konu modelleme yöntemi kullanılmıştır. LDA yöntemi, metinlerde birlikte kullanılan kelimeleri belirleyerek bu kelime kümelerini temalar halinde gruplandıran denetimsiz bir makine öğrenmesi tekniğidir. Turizm araştırmalarında konu modelleme yöntemleri, ziyaretçi deneyimlerinin ve destinasyon algılarının analiz edilmesinde giderek daha yaygın biçimde kullanılmaktadır (Zaman, 2024). Tematik analiz sürecine ek olarak ziyaretçi yorumlarının duygusal tonunu belirlemek amacıyla duygu analizi uygulanmıştır. Bu analiz kapsamında yorumlar pozitif, nötr ve negatif olmak üzere üç kategori

altında sınıflandırılmıştır. Duygu analizi, ziyaretçilerin destinasyonlara yönelik memnuniyet düzeylerini ve deneyimlerinin olumlu ya da olumsuz yönlerini ortaya koymaya yardımcı olan önemli bir analiz yöntemidir. Elde edilen temalar ve analiz sonuçları daha sonra yenileyici turizm yaklaşımı çerçevesinde yorumlanmıştır.

Bu doğrultuda metin madenciliği sonucunda elde edilen temalar yenileyici turizm literatüründe öne çıkan boyutlar çerçevesinde değerlendirilmiştir. Bu boyutlar arasında kültürel miras algısı, çevresel ve mekânsal deneyim, ziyaretçilerin destinasyonla kurduğu bağ (place attachment) ve dönüştürücü turizm deneyimi yer almaktadır. Analizin son aşamasında ise Galata Kulesi ve Maiden (Kız) Kulesi için elde edilen bulgular karşılaştırmalı olarak incelenmiş ve iki destinasyona ilişkin ziyaretçi deneyimlerinin benzerlikleri ve farklılıkları ortaya konulmuştur. Araştırmanın veri analiz sürecine ilişkin aşamalar Şekil 1'de sunulmaktadır.



Şekil 1. Metin Madenciliği Tabanlı Veri Analiz Süreci

4. BULGULAR

4.1 Kelime Frekansı Analizi ve kelime bulutu

Ziyaretçi yorumlarında öne çıkan kavramları belirlemek amacıyla kelime frekansı analizi gerçekleştirilmiş ve sonuçlar kelime bulutu görselleri ile görselleştirilmiştir. Kelime bulutu analizi, metin verilerinde en sık kullanılan kelimelerin görsel olarak sunulmasını sağlayarak ziyaretçi deneyimlerinde öne çıkan unsurların hızlı biçimde anlaşılmasına yardımcı olmaktadır. Bu yöntem turizm araştırmalarında kullanıcı yorumlarından elde edilen büyük metin veri setlerinin incelenmesinde yaygın olarak kullanılan bir metin madenciliği tekniğidir. Analiz kapsamında TripAdvisor platformundan elde edilen Galata Kulesi ve

Maiden (Kız) Kulesi ziyaretçi yorumları anlamlı kelimeler analiz sürecine dahil edilmiştir. Elde edilen sonuçlar Şekil 2'de karşılaştırmalı olarak sunulmaktadır.



Şekil 2. Galata Kulesi ve Maiden (Kız) Kulesi ziyaretçi yorumlarına ilişkin karşılaştırmalı kelime bulutları

Şekil 2 incelendiğinde Galata Kulesi'ne ilişkin ziyaretçi yorumlarında özellikle *galata, güzel, manzara, tarihi, ziyaret, kule ve şehir* gibi kelimelerin ön plana çıktığı görülmektedir. Bu kelimeler ziyaretçilerin Galata Kulesi deneyimini büyük ölçüde panoramik şehir manzarası, tarihi yapı özellikleri ve görsel deneyim üzerinden değerlendirdiklerini göstermektedir. Özellikle *manzara* ve *güzel* kelimelerinin yüksek frekansta kullanılması, ziyaretçilerin Galata Kulesi'ni İstanbul'un panoramik görünümünü deneyimleyebilecek önemli bir gözlem noktası olarak algıladıklarını ortaya koymaktadır. Bu durum Galata Kulesi'nin ziyaretçi deneyiminde mekânsal deneyim ve estetik algı boyutlarının öne çıktığını göstermektedir.

Maiden (Kız) Kulesi yorumlarında ise *bakü, kule, şehir, tarihi, ziyaret ve güzel* gibi kelimelerin öne çıktığı görülmektedir. Bu kavramların Bakü kentinin tarihsel ve kültürel kimliğini temsil eden önemli bir simge olarak değerlendirildiğini göstermektedir. Özellikle *bakü* ve *şehir* kelimelerinin yüksek frekansta kullanılması, bu yapının ziyaretçi deneyiminde güçlü bir yer ve kent kimliği ile ilişkilendirildiğini ortaya koymaktadır.

İki destinasyon karşılaştırıldığında, Galata Kulesi yorumlarında manzara ve görsel deneyim odaklı ifadelerin, Maiden Kulesi yorumlarında ise kent kimliği ve tarihsel bağlamı vurgulayan ifadelerin daha belirgin olduğu görülmektedir. Bu bulgu, her iki kulenin ziyaretçi deneyiminde farklı algısal boyutların ön plana çıktığını göstermektedir. Galata Kulesi daha çok panoramik deneyim ve görsel çekicilik, Maiden Kulesi ise tarihsel ve kentsel simgesellik üzerinden değerlendirilmektedir.

Bu sonuçlar, ziyaretçi deneyimlerinin destinasyonun tarihsel, kültürel ve sembolik anlamlarıyla da şekillendiğini göstermektedir. Elde edilen bulgular, tarihi kulelerin ziyaretçi deneyiminde hem kültürel miras algısı hem de mekânsal deneyim boyutları açısından önemli bir rol oynadığını ortaya koymaktadır.

4.2 Konu Modellemesi (Topic Modeling) Analizi

Ziyaretçi yorumlarında ortaya çıkan tematik yapıları belirlemek amacıyla Latent Dirichlet Allocation (LDA) konu modelleme yöntemi uygulanmıştır. Analiz sonucunda ziyaretçi yorumlarında dört temel tematik yapı ortaya çıkmıştır. Elde edilen kelime kümeleri ve yorum içerikleri dikkate alınarak bu temalar yorumlanmış ve yenileyici turizm literatüründe öne çıkan kavramsal boyutlar çerçevesinde değerlendirilmiştir. Bu temalar sırasıyla kültürel miras algısı, çevresel ve mekânsal deneyim, destinasyonla kurulan bağ ve dönüştürücü turizm deneyimi olarak adlandırılmıştır.

Tablo 1. LDA konu modelleme sonuçları

Yenileyici Turizm Boyutu	Anahtar Kelimeler	Tematik Yorum
Kültürel miras algısı	tarih, eski, mimari, miras, kültür	Tarihsel ve kültürel değerlerin vurgulanması
Çevresel ve mekânsal deneyim	manzara, şehir, panorama, yukarı, seyir	Manzara ve mekânsal deneyim
Destinasyonla kurulan bağ	simge, ünlü, önemli, şehir sembolü	Destinasyon kimliği ve sembolik değer
Dönüştürücü turizm deneyimi	deneyim, etkileyici, unutulmaz, güzel	Ziyaret deneyiminin duygusal etkisi

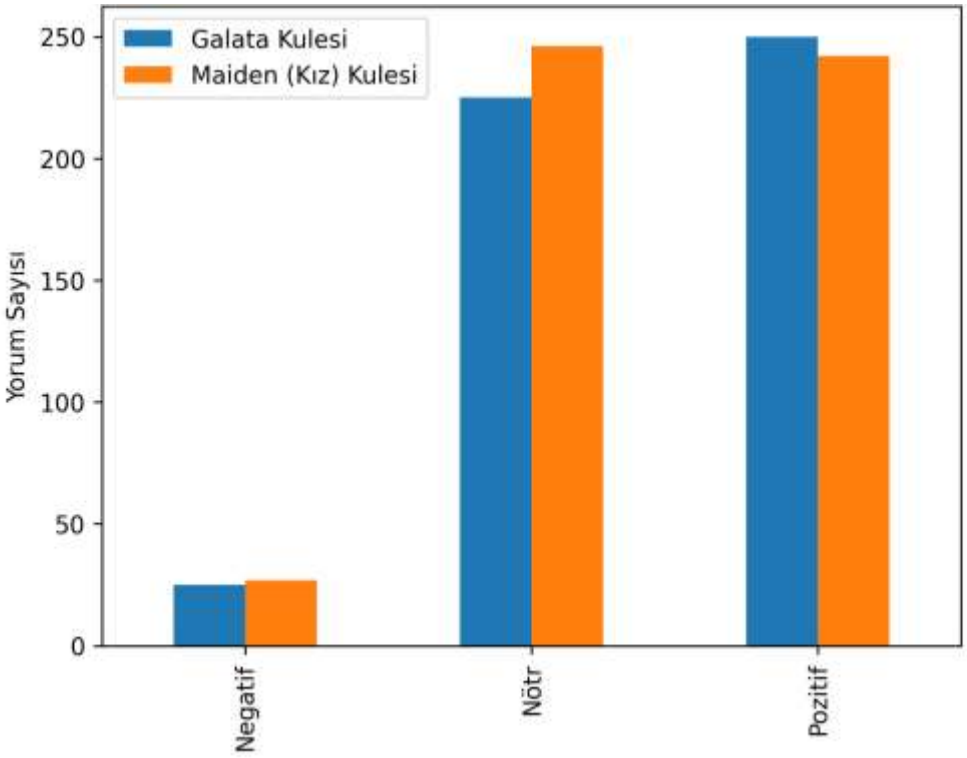
Analiz sonuçları, ziyaretçi yorumlarının deneyimsel ve duygusal boyutlara odaklandığını göstermektedir. Özellikle kültürel miras ve mekânsal deneyim temalarının veri setinde güçlü bir şekilde ortaya çıkması tarihi yapıların ziyaretçi deneyimlerinde önemli bir rol oynadığını göstermektedir. LDA analizi sonucunda elde edilen temalar Galata Kulesi ve Maiden (Kız) Kulesi ziyaretçi yorumları bağlamında karşılaştırmalı olarak incelenmiştir. Analiz sonuçları iki destinasyonun ziyaretçi deneyimlerinin farklı tematik odak noktaları etrafında şekillendiğini göstermektedir.

Galata Kulesi yorumlarında özellikle çevresel ve mekânsal deneyim temasının daha baskın olduğu görülmektedir. Ziyaretçi yorumlarında *manzara, şehir, panorama, seyir noktası* gibi ifadelerin sık kullanılması, kulenin ziyaretçiler tarafından büyük ölçüde İstanbul'un panoramik manzarasını deneyimleyebilecekleri bir gözlem noktası olarak algılandığını göstermektedir. Bu durum Galata Kulesi ziyaret deneyiminde özellikle mekânsal algı ve görsel deneyim boyutunun öne çıktığını ortaya koymaktadır. Buna karşılık Maiden (Kız) Kulesi yorumlarında kültürel miras algısı temasının daha belirgin olduğu görülmektedir. Ziyaretçilerin yorumlarında *tarih, eski yapı, mimari, kültürel miras* gibi kelimelerin daha sık kullanılması kulenin tarihsel ve mimari özelliklerinin ziyaretçi deneyiminde önemli bir rol oynadığını göstermektedir. Bu bulgu Maiden Kulesi'nin ziyaretçiler tarafından tarihsel ve kültürel bir simge olarak algılandığını göstermektedir.

Her iki destinasyonda da destinasyonla kurulan bağ teması dikkat çekmektedir. Ziyaretçilerin yorumlarında yer alan *simge yapı, önemli yer, şehrin sembolü* gibi ifadeler bu yapıların ziyaretçiler açısından güçlü bir mekânsal kimlik oluşturduğunu göstermektedir. Ayrıca her iki kule için de dönüştürücü turizm deneyimi ile ilişkili ifadelerin bulunduğu görülmektedir. Ziyaretçilerin *etkileyici, unutulmaz, harika deneyim* gibi ifadeler kullanması bu yapıların ziyaretçiler üzerinde duygusal ve deneyimsel bir etki yarattığını göstermektedir. Genel olarak değerlendirildiğinde, Galata Kulesi'nin ziyaretçi deneyiminde mekânsal deneyim ve şehir manzarası, Maiden Kulesi'nin ise tarihsel ve kültürel miras boyutunun daha belirgin olduğu görülmektedir. Her iki destinasyon da ziyaretçiler açısından güçlü sembolik anlamlar taşıyan ve deneyimsel değer üreten turistik mekânlar olarak algılanmaktadır.

4.3 Duygu Analizi (Sentiment Analysis)

Ziyaretçi yorumlarının duygusal tonunu belirlemek amacıyla veri setine duygu analizi uygulanmıştır. Bu analiz kapsamında yorumlar üç temel kategori altında sınıflandırılmıştır: pozitif, nötr ve negatif. Duygu analizi, ziyaretçilerin destinasyonlara yönelik memnuniyet düzeylerini ve deneyimlerine ilişkin algılarını belirlemek açısından önemli bir metin madenciliği yöntemidir. TripAdvisor platformundan elde edilen Galata Kulesi ve Maiden (Kız) Kulesi yorumları üzerinde gerçekleştirilen analiz sonucunda yorumların büyük ölçüde pozitif duygular içerdiği görülmüştür. Duygu analizine ilişkin dağılım Şekil 2'de sunulmaktadır.



Şekil 2. Galata Kulesi ve Maiden (Kız) Kulesi ziyaretçi yorumlarının duygu dağılımı

Analiz sonuçlarına göre her iki destinasyon için de pozitif yorumların diğer duygu kategorilerine kıyasla daha yüksek olduğu görülmektedir. Pozitif yorumlar genellikle ziyaretçilerin kulelerin sunduğu panoramik manzara, tarihi atmosfer ve turistik çekicilik gibi özelliklerden duydukları memnuniyeti ifade etmektedir. Özellikle "güzel", "harika", "muhteşem", "şaşırtıcı" ve "büyük" gibi ifadelerin sık kullanılması ziyaretçilerin genel deneyimlerinden memnun olduklarını göstermektedir. Nötr yorumlar ise çoğunlukla destinasyonlara ilişkin bilgilendirici veya açıklayıcı ifadelerden oluşmaktadır. Bu tür yorumlarda ziyaretçiler genellikle kulenin tarihi, mimari özellikleri veya ziyaret sürecine ilişkin gözlemlerini paylaşmaktadır. Negatif yorumların ise sınırlı sayıda olduğu görülmektedir. Negatif yorumlar çoğunlukla ziyaretçi yoğunluğu, uzun merdivenler, fiziksel erişim zorlukları ve giriş ücretleri gibi operasyonel unsurlarla ilişkilidir. Özellikle bazı ziyaretçilerin kuleye çıkma sürecinde merdivenlerin dar veya zorlayıcı olduğunu ifade ettikleri görülmektedir. Genel olarak değerlendirildiğinde, duygu analizi sonuçları, Galata Kulesi ve Maiden (Kız) Kulesi'nin ziyaretçiler tarafından olumlu turistik deneyimler sunan önemli kültürel miras alanları olarak algılandığını göstermektedir. Bu durum tarihi kulelerin destinasyon imajı ve ziyaretçi memnuniyeti açısından önemli bir rol oynadığını ortaya koymaktadır. Yenileyici turizm perspektifi açısından değerlendirildiğinde ise pozitif ziyaretçi deneyimleri, tarihi ve kültürel miras alanlarının ziyaretçilerde destinasyona yönelik olumlu algılar oluşturduğunu ve bu yapıların destinasyonun kültürel değerleriyle bağ kurmayı destekleyen deneyimler sunabildiğini göstermektedir.

4.4 Yenileyici Turizm Perspektifinde Bulguların Değerlendirilmesi

Bu çalışmada gerçekleştirilen metin madenciliği analizleri, araştırmanın giriş bölümünde ortaya konulan araştırma sorularına bütüncül bir yanıt sunmaktadır. Elde edilen bulgular, Galata Kulesi ve Maiden (Kız) Kulesi ziyaretçi yorumlarında öne çıkan deneyim temalarının büyük ölçüde tarihsel yapı algısı, mekânsal deneyim, kent kimliği ve ziyaret deneyimi etrafında şekillendiğini göstermektedir. Kelime frekansı analizi ve konu modellemesi sonuçları ziyaretçilerin her iki yapıyı da öncelikle tarihsel ve kültürel miras unsuru olarak değerlendirdiklerini ortaya koymaktadır. Elde edilen temaların yenileyici turizm yaklaşımının temel boyutları ile anlamlı biçimde ilişkilendirilebildiği görülmektedir. Ziyaretçi yorumlarında yer alan tarihsel ve kültürel vurgular kültürel miras algısı boyutunu yansıtırken, manzara ve mekânsal deneyime ilişkin

ifadeler mekânsal deneyim boyutunu ortaya koymaktadır. Ayrıca ziyaretçilerin destinasyon ile kurdukları sembolik ve duygusal ilişki yerel bağ kurma boyutu ile ilişkilendirilebilirken, deneyimin etkileyici veya unutulmaz olarak tanımlanması dönüştürücü turizm deneyimi boyutuna işaret etmektedir. Karşılaştırmalı analiz sonuçları iki destinasyon arasında bazı belirgin farklılıkların bulunduğunu göstermektedir. Galata Kulesi yorumlarında panoramik manzara ve görsel deneyim vurgusu daha güçlü biçimde ortaya çıkarken, Maiden (Kız) Kulesi yorumlarında kent kimliği ve tarihsel simgesellik vurgusunun daha belirgin olduğu görülmektedir.

5. TARTIŞMA VE SONUÇ

Bu çalışmada Galata Kulesi (İstanbul) ve Maiden (Kız) Kulesi (Bakü) ziyaretçi yorumları metin madenciliği yöntemleri kullanılarak analiz edilmiş ve elde edilen bulgular yenileyici turizm yaklaşımı çerçevesinde değerlendirilmiştir.

Araştırma bulguları öncelikle tarihi ve kültürel simge yapıların ziyaretçi deneyimlerinde güçlü bir kültürel miras algısı oluşturduğunu ortaya koymaktadır. Özellikle Maiden Kulesi yorumlarında tarihsel ve mimari özelliklere yapılan vurguların daha belirgin olduğu görülmektedir. Bu bulgu literatürde kültürel miras alanlarının destinasyonların tarihsel kimliğini ve kolektif hafızasını temsil eden mekânlar olduğunu vurgulayan çalışmalarla örtüşmektedir (Hall vd., 2025; Meneghello, 2023; Li vd., 2025). Tarihi yapıların ziyaretçi deneyimleri aracılığıyla kültürel değerlerin yeniden yorumlanmasına katkı sağlayabileceği söylenebilir.

Araştırma bulguları ayrıca Galata Kulesi yorumlarında çevresel ve mekânsal deneyim temasının daha belirgin olduğunu göstermektedir. Ziyaretçilerin şehir manzarası, panoramik görünüm ve seyir deneyimi gibi unsurlara sıkça vurgu yapmaları, kulenin ziyaretçiler tarafından özellikle mekânsal deneyim açısından değerlendirildiğini göstermektedir. Bu sonuç yenileyici turizm yaklaşımının vurguladığı yer temelli deneyim ve mekânsal farkındalık kavramları ile uyumludur. Literatürde yenileyici turizmin turizm deneyimlerinin ziyaretçilerin mekânla daha anlamlı ilişkiler kurmasına katkı sağlayan yer temelli süreçler olarak ele alınması gerektiği vurgulanmaktadır (Bellato vd., 2022; Dredge, 2022; Yeoman ve Sheldon, 2024).

Çalışmada ortaya çıkan bir diğer önemli tema, destinasyonla kurulan bağ olmuştur. Ziyaretçi yorumlarında simgesel yapı, şehir sembolü ve önemli turistik nokta gibi ifadelerin yer alması bu kulelerin ziyaretçiler açısından güçlü bir mekânsal kimlik oluşturduğunu göstermektedir. Bu bulgu turizm deneyimlerinin ziyaretçilerin destinasyonla duygusal ve sembolik bağlar kurabildiğini gösteren çalışmalarla paralellik göstermektedir (Pung vd., 2024; Corral-Gonzalez vd., 2024). Yenileyici turizm yaklaşımı açısından değerlendirildiğinde, destinasyonla kurulan bu bağ, turizm faaliyetlerinin ziyaretçiler ile destinasyon arasında daha anlamlı ilişkiler kurulmasına katkı sağlayabileceğini göstermektedir.

Araştırmanın bir diğer önemli bulgusu ziyaretçi deneyimlerinde dönüştürücü turizm deneyimi boyutunun ortaya çıkmasıdır. Ziyaretçi yorumlarında sıkça kullanılan etkileyici, unutulmaz ve özel deneyim gibi ifadeler tarihi kule ziyaretlerinin ziyaretçiler üzerinde duygusal ve deneyimsel etkiler yarattığını göstermektedir. Bu durum turizm deneyimlerinin ziyaretçilerin destinasyon algısını ve mekânla kurdukları ilişkiyi dönüştürebileceğini göstermektedir. Literatürde yenileyici turizm yaklaşımının turizm deneyimlerinin ziyaretçiler ve destinasyonlar arasında dönüştürücü ilişkiler kurabileceğini vurgulayan çalışmalar bulunmaktadır (Bellato vd., 2023; Cave ve Dredge, 2021; Sheldon, 2022).

Duygu analizi sonuçlarının her iki destinasyon için de ziyaretçi yorumlarının büyük ölçüde pozitif olduğunu göstermesi, tarihi kulelerin ziyaretçiler açısından güçlü deneyimsel değerler ürettiğini ortaya koymaktadır. Bu durum kültürel miras destinasyonlarının ziyaretçi memnuniyeti ve deneyim kalitesi açısından önemli turistik çekicilikler olduğunu gösteren çalışmalarla da uyumludur (Li vd., 2025; Levy ve Shani, 2025). Bazı yorumlarda ziyaretçi yoğunluğu veya fiziksel erişim koşullarına ilişkin olumsuz değerlendirmelerin bulunması, destinasyon yönetimi açısından deneyim kalitesinin sürdürülebilir şekilde yönetilmesi gerektiğini göstermektedir.

Bu çalışmanın literatüre önemli katkılarından biri yenileyici turizm yaklaşımını kentsel kültürel miras destinasyonları bağlamında ele almasıdır. Yenileyici turizm literatüründe yapılan çalışmaların önemli bir kısmı doğa temelli turizm veya kırsal destinasyonlara odaklanmaktadır (Bellato ve Pollock, 2025; Yeoman ve Sheldon, 2024). Buna karşılık, bu çalışma tarihi ve simgesel yapıların da yenileyici turizm perspektifi açısından önemli deneyimsel ve kültürel değerler üretebildiğini göstermektedir. Ayrıca çalışma çevrimiçi ziyaretçi yorumlarının metin madenciliği yöntemleri kullanılarak analiz edilmesi yoluyla yenileyici turizm literatürüne metodolojik bir katkı sunmaktadır.

5.1. Teorik Çıkarımlar

Bu çalışmanın bulguları, yenileyici turizm literatürüne çeşitli teorik katkılar sunmaktadır. Öncelikle çalışma, ziyaretçi deneyimlerinin kültürel miras algısı, mekânsal deneyim ve yer ile kurulan bağ gibi sosyo-kültürel boyutlar üzerinden de değerlendirilebileceğini ortaya koymaktadır. Bu durum, yenileyici turizm yaklaşımının destinasyon deneyimini daha bütüncül bir çerçevede ele alan teorik yaklaşımını desteklemektedir.

Ayrıca, metin madenciliği yöntemleri kullanılarak çevrimiçi ziyaretçi yorumlarının analiz edilmesi, yenileyici turizm araştırmalarında büyük veri ve kullanıcı tarafından üretilen içeriklerin (user-generated content) önemli bir veri kaynağı olabileceğini göstermektedir. Bu bulgu, yenileyici turizm çalışmalarında ziyaretçi deneyimlerinin daha geniş veri setleri üzerinden incelenmesine olanak sağlayan yeni metodolojik yaklaşımların kullanılabileceğini ortaya koymaktadır.

Son olarak Galata Kulesi ve Maiden (Kız) Kulesi örnekleri üzerinden elde edilen bulgular, tarihi ve kültürel yapıların ziyaretçi deneyiminde kentsel kimlik, kültürel miras algısı ve deneyimsel turizm bağlamında değerlendirilebilecek çok boyutlu bir destinasyon unsuru olduğunu göstermektedir. Bu durum, yenileyici turizm literatüründe yer alan yer temelli turizm yaklaşımlarını destekleyen bir bulgu olarak değerlendirilebilir.

5.2. Uygulayıcılar İçin Öneriler

Araştırma bulguları, destinasyon yöneticileri ve turizm planlamacıları açısından da çeşitli uygulamaya yönelik öneriler sunmaktadır. Ziyaretçi yorumlarında Galata Kulesi için özellikle manzara ve görsel deneyim, Maiden (Kız) Kulesi için ise tarihsel kimlik ve kent simgeselliği vurgularının öne çıkması, destinasyon yönetiminde bu deneyim boyutlarının daha görünür hale getirilmesi gerektiğini göstermektedir. Destinasyon yöneticileri, ziyaretçi deneyimini geliştirmek amacıyla panoramik manzara noktaları, deneyimsel ziyaret alanları ve kültürel anlatı unsurlarını daha etkin biçimde kullanabilirler.

İkinci olarak, çevrimiçi platformlarda yer alan ziyaretçi yorumlarının sistematik biçimde analiz edilmesi destinasyon yönetimi açısından önemli bir geri bildirim mekanizması sunmaktadır. Turizm yöneticileri, kullanıcı tarafından üretilen içerikleri analiz ederek ziyaretçi beklentilerini daha iyi anlayabilir ve destinasyon yönetimi stratejilerini bu doğrultuda geliştirebilirler.

Üçüncü olarak, tarihi ve kültürel yapıların destinasyon deneyimindeki rolü dikkate alındığında, bu yapıların deneyimsel turizm ve kültürel anlatı unsurları ile ele alınması önem taşımaktadır. Bu yaklaşım, ziyaretçilerin destinasyonla daha güçlü bir bağ kurmasına katkı sağlayarak yenileyici turizm anlayışının temel hedeflerinden biri olan yer ile anlamlı ilişki kurma sürecini destekleyebilir.

6. ARAŞTIRMA SINIRLILIKLARI VE GELECEK ÇALIŞMALAR

Bu çalışmanın bazı sınırlılıkları bulunmaktadır. Öncelikle araştırma verileri yalnızca TripAdvisor platformunda yer alan ziyaretçi yorumları ile sınırlıdır. Çevrimiçi kullanıcı yorumları ziyaretçi deneyimlerine ilişkin önemli bilgiler sunmakla birlikte, tüm ziyaretçilerin görüşlerini temsil etmeyebilir. Ayrıca kullanıcı yorumları bireysel deneyimlere dayandığı için ziyaretçilerin algıları ve değerlendirmeleri kişisel bakış açılarını yansıtmaktadır.

Araştırmanın bir diğer sınırlılığı veri setinin iki kültürel miras destinasyonu ile sınırlı olmasıdır. Çalışmada İstanbul'daki Galata Kulesi ve Bakü'deki Maiden (Kız) Kulesi ziyaretçi yorumları analiz edilmiştir. Bu

nedenle elde edilen bulgular farklı turizm destinasyonlarına veya farklı turizm türlerine doğrudan genellenemeyebilir. Çalışmada kullanılan metin madenciliği yöntemleri ziyaretçi yorumlarında öne çıkan tematik yapıları ortaya koymak açısından önemli avantajlar sağlamaktadır. Ayrıca, bu yöntemler yorumların bağlamsal anlamlarını her zaman tam olarak yansıtamayabilir. Özellikle kısa yorumlarda veya farklı dillerde yazılmış yorumlarda bazı anlam farklılıkları oluşabilmektedir.

Gelecek araştırmalarda farklı turizm destinasyonlarında benzer analizlerin gerçekleştirilmesi yenileyici turizm yaklaşımının farklı bağlamlarda nasıl ortaya çıktığını anlamaya katkı sağlayabilir. Özellikle farklı şehirlerde yer alan kültürel miras alanlarının karşılaştırmalı olarak incelenmesi turizm deneyimlerinin mekânsal ve kültürel farklılıklarını ortaya koyabilir.

Ayrıca gelecekte yapılacak çalışmalar ziyaretçi yorumlarının yanı sıra turist görüşmeleri, anketler veya katılımcı gözlem gibi nitel araştırma yöntemleri kullanarak ziyaretçi deneyimlerini daha derinlemesine inceleyebilir. Bu tür çalışmalar yenileyici turizm yaklaşımının ziyaretçi deneyimleri üzerindeki etkilerini daha kapsamlı biçimde ortaya koyabilir. Metin madenciliği yöntemleri ile sosyal medya verileri, blog yazıları veya farklı çevrimiçi platformlardaki yorumların birlikte analiz edilmesi turizm deneyimlerinin daha geniş veri setleri üzerinden incelenmesine olanak sağlayabilir. Bu tür çok kaynaklı veri analizleri yenileyici turizm araştırmalarının metodolojik çeşitliliğini artırabilir.

KAYNAKLAR

Amorim, M., Brandão, F., Gomes, A. R., & Dias, M. F. (2025). Tourism as a Catalyst for Regeneration: Identifying Enablers and Overcoming Challenges. In *International Conference of the International Association of Cultural and Digital Tourism* (pp. 107-131). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-78471-2_5

Becken, S., & Kaur, J. (2021). Anchoring "tourism value" within a regenerative tourism paradigm–a government perspective. *Journal of Sustainable Tourism*, *30*(1), 52-68. https://doi.org/10.1080/09669582.2021.1990305

Bellato, L., Frantzeskaki, N., & Nygaard, C. A. (2023). Regenerative tourism: A conceptual framework leveraging theory and practice. *Tourism Geographies*, *25*(4), 1026-1046. https://doi.org/10.1080/14616688.2022.2044376

Bellato, L., Frantzeskaki, N., Briceño Fiebig, C., Pollock, A., Dens, E., & Reed, B. (2022). Transformative roles in tourism: adopting living systems' thinking for regenerative futures. *Journal of Tourism Futures*, *8*(3), 312-329. https://doi.org/10.1108/JTF-11-2021-0256

Bellato, L. (2025). Regenerative development approaches involving tourism: Insights derived from place. *Tourism Planning & Development*, *22*(6), 960-980. https://doi.org/10.1080/21568316.2024.2415119

Bellato, L., & Pollock, A. (2025). Regenerative tourism: A state-of-the-art review. *Tourism Geographies*, *27*(3-4), 558-567. https://doi.org/10.1080/14616688.2023.2294366

Boluk, K. A., & Panse, G. (2022). Recognising the regenerative impacts of Canadian women tourism social entrepreneurs through a feminist ethic of care lens. *Journal of Tourism Futures*, *8*(3), 352-366. https://doi.org/10.1108/JTF-11-2021-0253

Cave, J., & Dredge, D. (2021). Regenerative tourism needs diverse economic practices. In *Global tourism and COVID-19* (pp. 49-59). Routledge. https://doi.org/10.4324/9781003223252-6

Corral-Gonzalez, L., Cavazos-Arroyo, J., & García-Mestanza, J. (2024). Turismo regenerativo: Expectativas a priori de valor de consumo tras un primer contacto con el concepto. *Revista De Estudios Empresariales. Segunda Época*, 23-40. https://doi.org/10.17561/ree.n2.2024.8761

Dredge, D. (2022). Regenerative tourism: Transforming mindsets, systems and practices. *Journal of tourism futures*, *8*(3), 269-281. https://doi.org/10.1108/JTF-01-2022-0015

Duarte, C. M., Cousins, R., Ficociello, M. A., Williams, I. D., & Khowala, A. (2024). Advancing global climate and biodiversity goals through regenerative tourism. *Sustainability*, *16*(20), 9133. https://doi.org/10.3390/su16209133

Forte, A., & Mariotti, A. (2024). Applying regenerative thinking in yachting tourism. Insights from the Northern Adriatic Sea. *Tourism Geographies*, 1-18. https://doi.org/10.1080/14616688.2024.2446354

Gerke, M., Adams, M., Ooi, C. S., & Dahles, H. (2024). Entrepreneuring for regenerative tourism. Doing business differently in Tasmania's regional hospitality industry. *Journal of Sustainable Tourism*, *32*(10), 2207-2224. https://doi.org/10.1080/09669582.2023.2273757

Hall, J., Paddison, B., & Crawforth, D. (2025). Slowing-down tourism in heritage cities: The role of independent businesses in achieving just and regenerative futures. *Tourism Planning & Development*, 1-21. https://doi.org/10.1080/21568316.2025.2529524

Higgins-Desbiolles, F. (2024). The end of tourism? Contemplations of collapse. *Journal of Tourism Futures*, *10*(3), 476-485. https://doi.org/10.1108/JTF-11-2023-0259

Hussain, A., & Fusté-Forné, F. (2026). Co-creating regenerative tourism: stakeholder perspectives and pathways from Banks Peninsula, New Zealand. *Current Issues in Tourism*, 1-11. https://doi.org/10.1080/13683500.2026.2623184

Iddawala, J., & Lee, D. (2025). Regenerative tourism: Context and conceptualisations. *Tourism planning & development*, 185-215. https://doi.org/10.1080/21568316.2025.2527614

Lemarié, J., & Bellato, L. (2025). Hawai'i is shifting toward regenerative tourism, yet extractive logics persist. *Tourism Geographies*, 1-21. https://doi.org/10.1080/14616688.2025.2582648

Levy, E., & Shani, A. (2025). Urban tourism as a catalyst for regenerative economic development: The case of eilat, Israel. *Tourism Planning & Development*, 216-240. https://doi.org/10.1080/21568316.2025.2535533

Li, X., Huang, Y., & Cho, K. (2025). A Study on the Role of Tourists' Multidimensional Perceptions in Regenerative Composite Cultural Spaces and Their Influence on Travel Intentions. *Sustainability*, *17*(18), 8332. https://doi.org/10.3390/su17188332

Lovelock, B., Houge Mackenzie, S., & Mei Pung, J. (2026). Linking wellness tourism and community well-being within a regenerative tourism paradigm. *Tourism Recreation Research*, 939-954. https://doi.org/10.1080/02508281.2025.2606533

Mandić, A., & Rastegar, R. (2025). Revisiting tourism taxes: a justice-oriented approach to redistributing revenue. *Current Issues in Tourism*, 1-8. https://doi.org/10.1080/13683500.2025.2542533

Manthiou, A., Klaus, P., Luong, V. H., & Tarquini-Poli, A. (2025). Exploring regenerative tourism: consumer perspectives on inspiration, legacy, and morality. *Journal of Travel & Tourism Marketing*, *42*(1), 118-132. https://doi.org/10.1080/10548408.2024.2437536

Meneghello, S. (2023). Mapping tourist landscapes in pandemic times: a dwelling-in-motion perspective. *Tourism Geographies*, *25*(7), 1730-1745. https://doi.org/10.1080/14616688.2023.2172604

Omma, F. M. (2024). Regenerative nature-based tourism: Tour guides and stakeholder dynamics in Arctic Norway. *Journal of Tourism Futures*. https://doi.org/10.1108/JTF-10-2023-0237

Pung, J. M., Houge Mackenzie, S., & Lovelock, B. (2024). Regenerative tourism: Perceptions and insights from tourism destination planners in Aotearoa New Zealand. *Journal of destination marketing & management*, *32*, 100874. https://doi.org/10.1016/j.jdmm.2024.100874

Schmidt Rojas, N., Sand, M. S., & Gross, S. (2025). Regenerative adventure tourism. Going beyond sustainability–a horizon 2050 paper. *Tourism Review*, *80*(1), 270-285. https://doi.org/10.1108/TR-12-2023-0874

Scheyvens, R., & Kaire Gataulu, T. (2025). Regenerating tourism and regenerating people: how tourism is achieving justice for Indigenous youths. *Journal of Sustainable Tourism*, 1403-1428. https://doi.org/10.1080/09669582.2025.2562435

Sheldon, P. J. (2022). The coming-of-age of tourism: embracing new economic models. *Journal of Tourism Futures*, *8*(2), 200-207. https://doi.org/10.1108/JTF-03-2021-0057

Soma, K., Verweij, P., Francis, A., & van Wonderen, D. (2025). Citizens' support to nature in Aruba—A study on sustainability and wellbeing. *Wellbeing, Space and Society*, *9*, 100281. https://doi.org/10.1016/j.wss.2025.100281

Vegas-Macias, J., Jørgensen, M. T., & Sørensen, F. (2026). Enterprise-mediated social value creation through tourist-resident interaction in regenerative tourism experiences: the case of CopenPay. *Journal of Sustainable Tourism*, 1-28. https://doi.org/10.1080/09669582.2026.2614536

Yeoman, I. S., & Sheldon, P. J. (2024). The future of tourism: interview with emeritus professor Pauline Sheldon on regenerative tourism. *Journal of Tourism Futures*, *10*(3), 402-407. https://doi.org/10.1108/JTF-06-2024-0120

Zaman, U. (2024). Nexus of regenerative tourism destination competitiveness, climate advocacy and visit intention: mediating role of travel FOMO and destination loyalty. *Sustainability*, *16*(17), 7827. https://doi.org/10.3390/su16177827

Zurbarán-Arias, D. J., & Piñeros, S. T. (2025). Community-based initiatives for the development of regenerative tourism: Actors and variables. Case: Southern villages of Fusagasugá (Cundinamarca, Colombia). *Anuario Turismo y Sociedad, 37*