ULUSLARARASI MÜHENDİSLİK VE BÖLGESEL KALKINMA KONFERANSI

INTERNATIONAL CONFERENCE ON ENGINEERING AND REGIONAL DEVELOPMENT

IntCERD 2026

BOOK OF PROCEEDINGS

Abstracts and Full Papers

- Electrical and Electronics Engineering
- Civil Engineering
- Materials and Metallurgy Engineering
- Precision Agriculture & Food Processing
- Computer Engineering
- Mechanical Engineering
- Energy Systems Engineering
- Food Engineering

ISBN: 978-605-80193-6-2

OCTOBER 22-24, 2026

KARAMANOĞLU MEHMETBEY UNIVERSITY
KARAMAN, TÜRKİYE

www.kmu.edu.tr/intcerd

INTERNATIONAL CONFERENCE ON
ENGINEERING AND REGIONAL DEVELOPMENT

IntCERD 2026

October 22–24, 2026 • Karaman, Türkiye

BOOK OF PROCEEDINGS

Abstracts and Full Papers

ISBN: 978-605-80193-6-2

Editors

Prof. Dr. Nevzat AYDIN
Prof. Dr. Selami BALCI
Prof. Dr. Yalçın COŞKUNER

Karamanoğlu Mehmetbey University
Karaman, Türkiye

IntCERD 2026 – International Conference on Engineering and Regional Development
Book of Proceedings (Abstracts and Full Papers)
October 22–24, 2026, Karaman, Türkiye

Publisher:	Karamanoğlu Mehmetbey University
Publication Date:	November 2026
ISBN:	978-605-80193-6-2
Editors:	Prof. Dr. Nevzat AYDIN Prof. Dr. Selami BALCI Prof. Dr. Yalçın COŞKUNER
Web:	www.kmu.edu.tr/intcerd
E-mail:	intcerd@kmu.edu.tr

All papers published in this book were evaluated through the scientific review process of the conference and are published as submitted by their authors. The authors are solely responsible for the scientific content, language, accuracy and originality of their papers and for obtaining any necessary permissions. The views expressed in the papers do not necessarily reflect those of the editors, the organizing committee or Karamanoğlu Mehmetbey University.

Conference Committees

Honorary Board

Prof. Dr. Mehmet GAVGALI	Rector, Karamanoğlu Mehmetbey University, Türkiye

General Chairs

Prof. Dr. Nevzat AYDIN	Vice Rector, Karamanoğlu Mehmetbey University, Türkiye
Prof. Dr. Selami BALCI	Dean of the Faculty of Engineering, Karamanoğlu Mehmetbey University, Türkiye
Prof. Dr. Yalçın COŞKUNER	Coordinator of the HASTUYIT (Precision Agriculture Applications and Innovative Processing Technologies), Karamanoğlu Mehmetbey University, Türkiye

Organizing Committee

Prof. Dr. Özlem ATEŞ SÖNMEZOĞLU	Karamanoğlu Mehmetbey University, Türkiye
Prof. Dr. Buket ÇARBAŞ	Karamanoğlu Mehmetbey University, Türkiye
Prof. Dr. Serdar ÇARBAŞ	Karamanoğlu Mehmetbey University, Türkiye
Prof. Dr. Ahmet KAYABAŞI	Karamanoğlu Mehmetbey University, Türkiye
Prof. Dr. Uğur KÖKLÜ	Karamanoğlu Mehmetbey University, Türkiye
Prof. Dr. Murat MAYDA	Karamanoğlu Mehmetbey University, Türkiye
Prof. Dr. Kadir SABANCI	Karamanoğlu Mehmetbey University, Türkiye
Prof. Dr. Abdulvahit SAYASLAN	Karamanoğlu Mehmetbey University, Türkiye
Prof. Dr. Savaş SÖNMEZOĞLU	Karamanoğlu Mehmetbey University, Türkiye
Prof. Dr. Metin TOZ	Karamanoğlu Mehmetbey University, Türkiye
Prof. Dr. Sadık Alper YILDIZEL	Karamanoğlu Mehmetbey University, Türkiye
Asst. Prof. Dr. Şinasi ÇETİNKAYA	Karamanoğlu Mehmetbey University, Türkiye
Asst. Prof. Dr. Sezgin ESER	Karamanoğlu Mehmetbey University, Türkiye

Scientific Committee

Prof. Dr. Ahu ALTINKUT UNCUOĞLU	Marmara University, Türkiye
Prof. Dr. Hüseyin ARSLAN	İstanbul Medipol University, Türkiye
Prof. Dr. Muhammad ASIM	Sivas Science and Technology University, Türkiye
Prof. Dr. Mehmet AVCAR	Süleyman Demirel University, Türkiye
Prof. Dr. Fatih BAŞÇİFTÇİ	Selçuk University, Türkiye
Prof. Dr. Ramazan BAYINDIR	Vice President of the YÖKAK, Türkiye
Prof. Dr. Nicu BIZON	Pitești University, Romania
Prof. Dr. Mehmet CANBAZ	Eskişehir Osmangazi University, Türkiye
Prof. Dr. Yunus A. ÇENGEL	OSTIM Technical University, Türkiye
Prof. Dr. Aysel ÇİMEN	Karamanoğlu Mehmetbey University, Türkiye
Prof. Dr. İhsan DAĞTEKİN	Fırat University, Türkiye
Prof. Dr. Mustafa Kürşat DEMİR	Necmettin Erbakan University, Türkiye
Prof. Dr. Mustafa ERSÖZ	Selçuk University, Türkiye
Prof. Dr. Nilgün ERTAŞ	Necmettin Erbakan University, Türkiye
Prof. Dr. İsa GÖKÇE	Tokat Gaziosmanpaşa University, Türkiye
Prof. Dr. Sabri GÖKMEN	Selçuk University, Türkiye
Prof. Dr. Arif HEPBAŞLI	Yaşar University, Türkiye
Prof. Dr. Nejdet KANDEMİR	Ankara Hacı Bayram Veli University, Türkiye
Prof. Dr. Mustafa KARAKAYA	Selçuk University, Türkiye
Prof. Dr. Erol KURT	Gazi University, Türkiye

Prof. Dr. Jose Manuel LOPEZ GUEDE	University of the Basque Country, Spain
Prof. Dr. Hakan ÖZKAN	Çukurova University, Türkiye
Prof. Dr. Mehmet SALTAN	Süleyman Demirel University, Türkiye
Prof. Dr. Cemalettin SARIÇOBAN	Selçuk University, Türkiye
Prof. Dr. İbrahim SEFA	Gazi University, Türkiye
Prof. Dr. Ali SEVİM	Kırşehir Ahi Evran University, Türkiye
Prof. Dr. Farhad SHAHNIA	Murdoch University, Australia
Prof. Dr. Bahattin TANYOLAÇ	Ege University, Türkiye
Prof. Dr. Serdal TERZİ	Süleyman Demirel University, Türkiye
Prof. Dr. Metin TOZ	Karamanoğlu Mehmetbey University, Türkiye
Prof. Dr. Yunus UZUN	Aksaray University, Türkiye
Prof. Dr. Mustafa YILDIRIM	Kahramanmaraş Sütçü İmam University, Türkiye
Assoc. Prof. Dr. Cenk ALBAYRAK	Trabzon University, Türkiye
Assoc. Prof. Dr. Ali CİNGÖZ	Tokat Gaziosmanpaşa University, Türkiye
Assoc. Prof. Dr. Mutlu ÇEVİK	Munzur University, Türkiye
Assoc. Prof. Dr. Uğur ERKAN	Ankara University, Türkiye
Assoc. Prof. Dr. Esma Nur GEÇER	Tokat Gaziosmanpaşa University, Türkiye
Assoc. Prof. Dr. Serap KARAGÖL	Ondokuz Mayıs University, Türkiye
Assoc. Prof. Dr. Serdal SABANCI	Munzur University, Türkiye
Assoc. Prof. Dr. Mustafa ÜNAL	Health Institutes of Türkiye (TÜSEB), Clinical Trials, Türkiye
Asst. Prof. Dr. Erkin AKDOĞAN	Karamanoğlu Mehmetbey University, Türkiye
Asst. Prof. Dr. Hüseyin DUYSAK	Karamanoğlu Mehmetbey University, Türkiye
Asst. Prof. Dr. Ayşe ELDEM	Karamanoğlu Mehmetbey University, Türkiye
Asst. Prof. Dr. Hüseyin ELDEM	Karamanoğlu Mehmetbey University, Türkiye
Asst. Prof. Dr. Ömer Faruk FIRAT	Recep Tayyip Erdoğan University, Türkiye
Asst. Prof. Dr. Hasan İbrahim KOZAN	Necmettin Erbakan University, Türkiye
Asst. Prof. Dr. Muhammed Furkan KÜÇÜK	Bursa Technical University, Türkiye
Asst. Prof. Dr. Seda KÜL	Karamanoğlu Mehmetbey University, IEEE Senior Member, Türkiye
Asst. Prof. Dr. Emine OLUM	İstanbul Medipol University, Türkiye
Asst. Prof. Dr. Ali ÖZCAN	Karamanoğlu Mehmetbey University, Türkiye
Asst. Prof. Dr. Güliz TOZ	Karamanoğlu Mehmetbey University, Türkiye
Asst. Prof. Dr. Berat YILDIZ	Karamanoğlu Mehmetbey University, Türkiye
Asst. Prof. Dr. Ahmet YILMAZ	Karamanoğlu Mehmetbey University, Türkiye
Satish KABADE	Enterprise Software Architect, IEEE Senior Member, USA
Anup KAGALKAR	Technical Product Expert, IEEE Senior Member, USA

Conference Secretariat

Assoc. Prof. Dr. Sabire YERLİKAYA	Karamanoğlu Mehmetbey University, Faculty of Engineering, Department of Food Engineering
Assoc. Prof. Dr. Gamze KARANFİL KAÇMAZ	Karamanoğlu Mehmetbey University, Faculty of Engineering, Department of Energy Systems Engineering
Asst. Prof. Dr. Kemal ARMAĞAN	Karamanoğlu Mehmetbey University, Faculty of Engineering, Department of Civil Engineering
Asst. Prof. Dr. Muhammet Rojhat KARA	Karamanoğlu Mehmetbey University, Faculty of Engineering, Department of Electrical and Electronics Engineering
Res. Asst. Dr. Hasan YÜNGEVİŞ	Karamanoğlu Mehmetbey University, Faculty of Engineering, Department of Metallurgical and Materials Engineering
Lect. Dr. Deniz YİĞİT KOÇAK	Karamanoğlu Mehmetbey University, Karamanoğlu Mehmetbey University, Quality Assurance and Accreditation Coordination Office
Lect. Bilal COŞKUN	Karamanoğlu Mehmetbey University, Karamanoğlu Mehmetbey University, Technology Transfer Office (TTO)
Res. Asst. Zehra ERKOL	Karamanoğlu Mehmetbey University, Faculty of Engineering, Department of Bioengineering
Res. Asst. Merve ÇALIŞKAN	Karamanoğlu Mehmetbey University, Faculty of Engineering, Department of Mechanical Engineering
Res. Asst. Gül Cihan HABEK	Karamanoğlu Mehmetbey University, Faculty of Engineering, Department of Computer Engineering
Res. Asst. Ahmet Tolgahan ŞENLİ	Karamanoğlu Mehmetbey University, Faculty of Engineering, Department of Computer Engineering
Res. Asst. İbrahim ARI	Karamanoğlu Mehmetbey University, Faculty of Engineering, Department of Electrical and Electronics Engineering

Preface

Dear Colleagues,

It is our great pleasure to present the Book of Proceedings of the International Conference on Engineering and Regional Development (IntCERD 2026), held on October 22–24, 2026 at Karamanoğlu Mehmetbey University in Karaman, Türkiye.

IntCERD 2026 was organized to bring together researchers, academicians, industry representatives and professionals from different engineering disciplines to share theoretical approaches, applied research findings and innovative engineering solutions that support regional and sustainable development. The conference covered a wide range of fields, including electrical and electronics, computer, civil, mechanical, materials and metallurgy, energy systems and food engineering, as well as precision agriculture applications and innovative food processing technologies.

All submissions were evaluated through the scientific review process of the conference, and the accepted studies were presented in person and online during the conference sessions. This book contains the abstracts and full papers of the studies presented at IntCERD 2026, arranged by conference track in the order of the conference programme. The papers are published as submitted by their authors.

We would like to thank our keynote speakers for their valuable contributions, all authors and presenters for sharing their work, the members of the Scientific Committee for their careful evaluations, the session chairs, and the members of the Organizing Committee and the Conference Secretariat for their dedicated efforts. We also thank the Rectorate of Karamanoğlu Mehmetbey University and the Coordinatorship for Precision Agriculture Applications and Innovative Processing Technologies (HASTUYİT) for their support.

We hope that this book will serve as a useful reference for researchers and practitioners and contribute to future collaborations between academia, industry and public institutions.

On behalf of the Organizing Committee,

Prof. Dr. Nevzat AYDIN	**Prof. Dr. Selami BALCI**	**Prof. Dr. Yalçın COŞKUNER**
General Chair	*General Chair*	*General Chair*

Keynote Speakers

Prof. Dr. Hüseyin ARSLAN

Istanbul Medipol University
Dean of the School of Engineering and Natural Sciences
Head of Communication, Signal Processing, and Networking Center (CoSiNC)
Istanbul, Türkiye

Keynote Title: *5G and Beyond for the Digitalization of the World*

Biography

Dr. Hüseyin Arslan (IEEE Fellow, NAI Fellow, Member of Turkish Academy of Science) received his B.S. degree from Middle East Technical University (METU), Ankara, Turkey in 1992, and his M.S. and Ph.D. degrees from Southern Methodist University (SMU), Dallas, TX, in 1994 and 1998, respectively. From 1998 to 2002, he worked with the research group at Ericsson. Between 2002 and 2022, Dr. Arslan was a faculty member in the Electrical Engineering Department at the University of South Florida, where he served as a Professor. In 2013, he joined Istanbul Medipol University to establish the Engineering College and currently serves as the Dean of the School of Engineering and Natural Sciences. He has provided expert consulting for organizations such as Anritsu, Savronik, and The Scientific and Technological Research Council of Turkey (TÜBİTAK). Dr. Arslan was the founding Chairman of the Board of Directors of ULAK Communications and served as a member of the TÜBİTAK Scientific Board. From 2021 to 2024, he was a member of the Board of Directors of Turkcell. His current research interests are on 6G and beyond radio access technologies, physical layer security, interference management, cognitive radio, multi-carrier wireless technologies, dynamic spectrum access, non-terrestrial communications, joint radar (sensing) and communication designs. He has collaborated extensively with global companies including InterDigital, NTT DoCoMo, Raytheon, Honeywell, ASELSAN, VESTEL, Turkcell, Anritsu, and Keysight Technologies. He also pioneered wireless education by developing innovative courses and the NSF-funded “Wireless Systems Laboratory.” He has served in numerous IEEE leadership roles and editorial boards.

Dr. Saadoon Obaid EYADA

Postdoctoral Researcher, Al Mansour University College
Senior Chief Engineer, Anbar Health Directorate, Iraqi Ministry of Health
Anbar, Iraq

Keynote Title: *The Implementation of Mechanistic Empirical Pavement Design Method in Türkiye*

Biography

Dr. Saadoon Obaid Eyada is a civil engineer specializing in highway engineering and transportation. He is a Postdoctoral Researcher at Al Mansour University College, Iraq, and has served as Senior Chief Engineer at the Anbar Health Directorate of the Iraqi Ministry of Health since 2003. He received his B.Sc. (1996) and M.Sc. (2009) degrees in Civil Engineering from the University of Anbar, and his Ph.D. in Civil Engineering (Highway Engineering & Transportation) from Selçuk University, Türkiye, in 2021, with a dissertation on the evaluation of Turkish standard flexible pavement sections using the mechanistic-empirical approach and life cycle cost analysis. His research interests include mechanistic-empirical pavement design, sustainable pavement materials and the recycling of waste materials, asphalt binder and mixture performance, life cycle cost analysis of pavements, and the engineering requirements of health facilities. He has authored more than 30 journal and conference papers and two books on the architectural and engineering requirements of health facilities, and holds an Iraqi patent (No. 5381, 2018) on producing building bricks from cement kiln waste. Dr. Eyada serves on the editorial board of the Advanced Transportation Journal (Mersin University, Türkiye) and has been a consultant engineer of the Iraqi Engineering Union since 1996.

Contents

Conference Committees ..... **iii**

Conference Secretariat ..... **v**

Preface ..... **vi**

Keynote Speakers ..... **vii**

PART I – ABSTRACTS **1**

Transformer Design, Manufacturing, and Testing Suitable for a Solar Power Plant System with Dual Voltage (High Voltage 1 to High Voltage 2 Ratio Equal to 2) 2
Duhan SAYGIN, Burçin DOĞRU KARCI

Providing a Dataset for Torque Ripple Analysis in Permanent Magnet Synchronous Motors Using FEA Parametric Solver 3
İbrahim ARI, Ahmet KAYABAŞI, Selami BALCI

Investigation of the Reflection Characteristics of Pancreatic Tissues Using a Split Ring Resonator Based Microwave Sensor on Different Substrates 4
Rabia TOPRAK

Development and Validation of a MATLAB-Based Integrated Computational Framework for Wireless Power Transfer System Analysis 5
Mehmet CİCEK, Ibrahim SEFA

Comparative Performance Analysis of Statistical and Deep Learning Models for Temperature Forecasting in Konya 6
Ayşe ELDEM, Habibe KARBEYAZ

Multi-Temporal Drought Analysis of Latakia Station Using the Standardized Precipitation Index 7
Akın DUVAN, Mohamed ZEINEDOIN

High Temperature Performance of Boron Waste Modified Mortars for Rapid Service Restoration in Fire Damaged Tunnels and Concrete Pavements 8
Kemal ARMAĞAN, Mehmet CANBAZ

Valorization of Date Palm Leaf Biomass as a Sustainable Biosorbent for Cationic Dye Removal 9
Siham BOULKAIBET, Chebli DERRADJI

Development and Commercialization of Legume and Vegetable Based Gluten Free Pasta Through University and Industry Collaboration 10
Mehmet KOYUNCU, Recep ÇİFÇİ, Şeyma AKIN

Improvement of Bread-Making Quality in Bread Wheat Lines Carrying a Rye Translocation 11
Bedrettin DEMİR, Zehra ERKOL, Nevzat AYDIN

New Recombinations Detected in High-Molecular-Weight Glutenin Subunits Encoded by the Glu-B1 Allele in Bread Wheat 12
Bedrettin DEMİR

Monitoring Biodiversity for Ecosystem Health and Technological Applications 13
Yasir DOĞAN, Elif YAVUZASLANOĞLU

A Study on Some Rheological Properties of Composite Flours Made from Wheat and Molokhia (Corchorus olitorius L.) Leaf Powder 14
Büşra DADUŞ, Yalçın COŞKUNER

Bibliometric and Content Analysis of Plant Attached Sensors for Agriculture in the Konya Plain Project Region 15
Ali OZCAN, Omer PEKTAS, Muhammed Emin BEDİR

Development of New Durum Wheat Cultivars Using Speed Breeding and Biochemical Marker-Assisted Selection 16
Bedrettin DEMİR, Ziya DURU, Aziz DOĞANAY, Nevzat AYDIN

Contaminants Formed During Food Processing 17
Muharrem GELEN, Nizam NİZAMLIOĞLU

Effects of Nixtamalization on Cooking Quality, Color, and Antioxidant Properties of Gluten-Free Pasta from Legumes and Pseudocereals 18
Mehmet KOYUNCU, Şeyma AKIN, Recep ÇİFÇİ, Muharrem GELEN, Abdulvahit SAYASLAN

PART II – FULL PAPERS 19

Adaptive Fault Prognosis for Ground Based Radio Navigation Equipment Using Multisensor Fusion 20
Huseynaga N. HAQVERDİYEV

Day-Ahead Solar PV Power Generation Forecasting Using Physics-Based Synthetic Data Generation and XGBoost 24
Nihat YILMAZ, İsmail Vedat YILMAZ, Osman Kaan YILDIRIM, Haydar DOĞAN, Süleyman Sungur TEZCAN

Comparison of Gradient Boosting and Deep Tabular Models for Synchronous Generator Performance Prediction Under Geomagnetically Induced Current Conditions 30
Şevket CANTÜRK, Oktay KARAKAYA

Shared Battery Storage Planning for Renewable Source Coalitions: PSO-Based Sizing and Shapley Attribution of Pooling Benefits 35
Emine İŞİN, Nurettin ÇETİNKAYA

Frequency Regulation Using a Cascade T2F-PIDDN Controller for LFC of a 2-Area Multi-Source Power System 41
Mehmet KARAYEL, Turan ÇAKIL

When Does Feature-Selection Stability Matter? Evidence Across Low- and High-Dimensional Classification Regimes 47
Ramazan EŞMELİ

Classification of Coffee Beans Under Non-Ideal Conditions 52
Ali BİRCAN, Nurdan AKHAN BAYKAN

A Model-Independent Upper Bound on Prediction Accuracy for M- and NM-Polynomial Topological Indices 59
Talha KIRCA, Merve GÖRGÜLÜ, Hasan YILDIRIM

BI-RADS Risk Group Classification Using Four-View Mammogram Fusion and Deep Learning 65
Güliz TOZ, Asiye SÖZERİ, Nisanur ERDAL, Merve YAZICI

Lightweight Convolutional Neural Networks for Wheat Yellow Rust Severity Classification: An Accuracy and Efficiency Benchmark 69
Hüseyin ELDEM, Alper DEMİR

Seasonal Characteristics of Long-Term Precipitation Trends in Douala, Cameroon, Based on ERA5 Reanalysis 74
Abdel Karim Kouotou NDAM, Djiguiba Abdallah Kone GOGBE, Soner Çağatay BAĞÇACI

Sand Column Test for Evaluating the Performance of Grout Injection in Historic Structures 78
Murat ERDEMİR, Sadık Alper YILDIZEL

Continuous 3D-Printed Polymer Reinforcement in Foam Concrete: A Review of Materials, Structures, Interface Behavior, and Research 81
Özge Zor AVCAN, Sadık Alper YILDIZEL, Ülkü Sultan KESKİN

Bird Strike Investigation of Bio-Inspired Laminated Composites with Different End Conditions and Cutouts 92
Meryem Dilara KOP, Aman GARG, Mehmet AVCAR

Free Vibration Analysis of Functionally Graded Porous Nanobeams 98
Burak İKİNCİ, Mehmet AVCAR

TTHM and HAA5 as Related but Non-Equivalent Disinfection By-Product Groups: Implications for Drinking-Water Monitoring and Data-Driven Modelling 105
Berkin ÜSTÜNYILDIZ

Machine Learning in Struvite Recovery: A Focused Review of Process Inputs and Prediction Targets 110
Berkin ÜSTÜNYILDIZ

Ağır Betonların TPU veya PLA Hasır Donatı Kullanılarak Optimizasyonu 115
Ömer ŞİMŞEK, Sadık Alper YILDIZEL, Selin ERZİN

A Comparative Study of the Number of Stories, Structural Rigidity, and LRB–FPS Hybrid Seismic Isolation Performance in Multi-Story Reinforced Concrete Frame Buildings 120
Rifat SEZER, Muhammet YANGÖZ

IRQMEPAVE: A Knowledge-Based Intelligent Mechanistic–Empirical Tool for Flexible Pavement Design in Iraq and Its Validation Against AASHTO 1993 Method and MnPAVE 127
Saadoon Obaid EYADA, Adil N. ABED

Design, Fabrication, and Electromechanical Integration of an Original Experimental Device for Measuring Tissue Friction Force in Medical Needle–Tissue Interactions 132
Erman ÇELİK, Alparslan Ulaş ÇAYDAŞ, Süreyya Elif KÖM

Development and Validation of a Generalized Decision-Support Tool for Engineered Air-Nozzle Retrofits in Compressed-Air Blow-Gun Systems 135
Orhan KAYA

Accuracy and Computational Cost of Modeling Choices in Computational Fluid Dynamics Simulations of Gas-Solid Cyclone Separators: A Comparative Review 140
Nurettin ÇAY, Erman ÇELİK

Optimization of Automotive B Pillar Crash Performance Through Local Carbon Fiber Composite Reinforcement 147
Ege UĞUR, Samet Deniz MAYDA, Mahmut PEKEDİS

Designing Soft Body Armor for End of Life: Reliability-Based Selection of Fiber and Areal Density Under Aging 150
Muhammed Zahid KOÇ

Is Biodegradable Greener? Robust Fiber Selection for Single-Use Nonwoven Wipes Under Data and Preference Uncertainty 155
Muhammed Zahid KOÇ

Integrating Environmental Performance into Hotel Pricing and Financial Outcomes: A Comparative Review of Quantitative and Analytical Studies Published from 2021 to 2026 159
Sarper SERİNTÜRK, Pırıl TEKİN, Tolunay GÖÇKEN

Determination of the Heating Energy Demand of a Public Building According to the TS 825:2008 and TS 825:2024 Standards: A Comparative Assessment for Four Different Provinces 163
Enes KAVUN, Ertaç HÜRDOĞAN

The Effect of Different Thermal Insulation Materials on the Heating-Related Carbon Emissions of a Building 169
Enes KAVUN, Ertaç HÜRDOĞAN

Feasibility of LPG Conversion for Passenger Cars in Türkiye: A Total Cost of Ownership Comparison of Gasoline and LPG Operation 174
Cemre POLAT, Coşkun ÖZALP

Active Packaging Systems for Extending Food Shelf Life and Reducing Food Waste: A Sustainable Food Systems Perspective 183
Muhammed Yusuf ÇAĞLAR, Abdullah ÇAĞLAR

The Role of "Best Before" and "Use By" Dates in Food Waste: Consumer Perception, Food Safety, and Sustainability 188
Muhammed Yusuf ÇAĞLAR, Abdullah ÇAĞLAR

Nanotechnology Applications in Foods: Nanoencapsulation of Phenolic Compounds 193
Serdar YEŞİL

Food Non-Compliances in Türkiye and Their Distribution by Public Health Risk: An Analysis of 559 Official Public Disclosure Records 199
Serdar YEŞİL

Sponsor 203

PART I

ABSTRACTS

IntCERD 2026 – International Conference on Engineering and Regional Development

Transformer Design, Manufacturing, and Testing Suitable for a Solar Power Plant System with Dual Voltage (High Voltage 1 to High Voltage 2 Ratio Equal to 2)

Duhan SAYGIN[1] Burçin DOĞRU KARCI[1]

[1]Beta Enerji ve Teknoloji A.Ş., Research and Development Department, Adana, Türkiye

***Abstract*—This study presents the design, manufacturing, and testing of a dual high voltage input transformer developed for solar power plant applications, enabling energy from two different high voltage levels to be transferred through a single transformer to a common low voltage level. A prototype transformer rated at 3300 kVA with high voltage input levels of 10.5 kV and 21 kV and a low voltage output of 0.8 kV was developed. During the design process, the magnetic core dimensions, winding structures, conductor selection, insulation system, cooling channels, electrical clearances, tank dimensions, and accessory arrangements were determined in accordance with the relevant IEC 60076 standards. Following the electrical design, the mechanical design was completed and a prototype was manufactured according to the defined design parameters. The manufacturing process included magnetic core and winding production, fabrication of mechanical components, active part assembly, drying, vacuum oil filling, and final assembly. Routine tests were performed on the developed prototype, including winding direct current resistance, load loss and short circuit impedance, no load loss and no load current, turns ratio and polarity, applied voltage, and induced voltage tests. The study demonstrates a feasible design approach for integrating two energy sources with different high voltage levels into a single transformer. The proposed configuration can provide a more compact solution for solar power plant applications and may offer potential advantages in installation area and overall system cost compared with the use of separate transformer units.**

***Keywords*— Dual input transformer, solar power plant, multiple voltage levels, renewable energy integration**

 ISBN: 978-605-80193-6-2

Providing a Dataset for Torque Ripple Analysis in Permanent Magnet Synchronous Motors Using FEA Parametric Solver

İbrahim Arı[1], Ahmet Kayabaşı[2] and Selami Balcı[3]

[1,2,3]Karamanoglu Mehmetbey University, Faculty of Engineering, Department of Electrical and Electronics Engineering, Karaman, Türkiye.

* Corresponding author e-mail: ibrahimari@kmu.edu.tr

***Abstract*— Permanent Magnet Synchronous Motors (PMSMs) are frequently preferred in electric vehicles due to critical factors such as their high power-to-weight ratio (kW/kg) and the ability to perform speed control over a wide range. However, the tooth-slot structure of the motor, saturation effects, and current harmonics make torque ripple inevitable. This situation leads to problems such as mechanical vibration, acoustic noise, and control instability. It is difficult to accurately calculate torque ripple using analytical methods due to magnetic saturation. Finite element simulations, on the other hand, provide accurate results but require a significant amount of computation time. Therefore, reliable simulation data is needed to investigate the motor's nonlinear responses and, when necessary, to use them in data-driven predictive models. In this study, a 0.55 kW interior PMSM operating at 1500 rpm was modeled using the Ansys Maxwell 2D transient solver, and parametric torque ripple data were generated. The motor geometry configured with a 24-slot, 4-pole interior rotor topology and nonlinear core saturation effects was directly transferred to the model. To accurately capture the magnetic flux distribution, the air gap and the mesh structure at the pole tips were condensed. Three key variables affecting torque ripples were selected for parametric analysis. These are the armature current (I_a), which determines magnetic saturation; the input voltage (V_i), which defines the operating limits; and the Pulse-Width Modulation (PWM) switching frequency (f_{PWM}), which influences the inverter harmonics. The defined operating conditions include an armature current of 0.83 A to 7.50 A, an input DC voltage of 110 V to 330 V and a PWM switching frequency of 5 kHz to 60 kHz. Across these conditions, a comprehensive dataset consisting of 628 distinct operating points was collected. Torque ripple ratios were calculated from the time-domain torque waveforms obtained from transient mode simulations. Quantitative results demonstrate that peak-to-peak torque ripple varies between 0.07 Nm and 1.53 Nm across the operating envelope, primarily affected by operational voltage and armature current dynamics. The findings are verified through transient electromagnetic FEA formulations incorporating nonlinear material saturation. The results provide a solid data foundation for torque ripple analysis, design optimization, and alternative data-driven prediction studies.**

***Keywords*—Permanent magnet synchronous motor (PMSM), torque ripples, Ansys Maxwell, parametric solver, finite element analysis (FEA), dataset creation.**

 ISBN: 978-605-80193-6-2

Investigation of the Reflection Characteristics of Pancreatic Tissues Using a Split Ring Resonator Based Microwave Sensor on Different Substrates

Rabia Toprak[1]

[1]Karamanoglu Mehmetbey University, Faculty of Engineering, Department of Electrical-Electronics Engineering, Karaman, Turkiye

* Corresponding author e-mail: rabiatop@kmu.edu.tr

***Abstract*— Pancreatic cancer is among the malignancies with the lowest survival rates, largely because it is diagnosed at a late stage, which creates a need for rapid and label free diagnostic approaches. Differences in water content and ionic conductivity between healthy and malignant tissues produce a measurable dielectric contrast in the microwave band, and this contrast can be sensed through the reflection characteristics of split ring resonator structures. In this study, one normal and two cancerous formalin-fixed pathological pancreatic tissue specimens were measured using split-ring resonator-based sensors realized on two substrates, FR4 and RT Duroid 5880, both 1.6 mm thick. Measurements were carried out with a Keysight N5224A vector network analyzer from 0.5 to 10 GHz for FR4 and RT Duroid 5880. On RT Duroid 5880, the resonance observed at 6.618 GHz for the normal specimen, with a return loss of 23.15 dB, shifted to 6.751 GHz for both cancerous specimens, a change of 133 MHz, with return losses of 22.15 dB and 22.09 dB. Across the band from 3.6 to 10 GHz the mean deviation of the cancerous specimens from the normal tissue was 0.25 dB, and the correlation between their deviation patterns was 0.981. On FR4 the two cancerous specimens behaved inconsistently, deviating by 2.52 dB and 17.42 dB at the dominant resonance of 4.655 GHz, with a correlation of 0.50 between their deviation patterns. The low-loss RT Duroid 5880 substrate therefore provides more consistent discrimination of pancreatic tissues than FR4. As all specimens were formalin-fixed, the reported contrast characterizes fixed pathological tissue. This preliminary study, limited to three specimens, will be extended with a larger specimen set and machine learning-based classification.**

***Keywords*—Split ring resonator, pancreatic cancer, microwave sensing, reflection coefficient, dielectric characterization**

 ISBN: 978-605-80193-6-2

Development and Validation of a MATLAB-Based Integrated Computational Framework for Wireless Power Transfer System Analysis

Mehmet Cicek[1] and Ibrahim Sefa[2]

[1]Gazi University, Graduate School of Natural and Applied Sciences, Electrical and Electronics Engineering, Ankara, Türkiye

[2]Gazi University, Faculty of Technology, Department of Electrical-Electronic Engineering, Ankara, Türkiye

Article note: This study was conducted as part of the first author's ongoing Ph.D. research under the supervision of the second author.

* Corresponding author e-mail: mehmet.cicek2@gazi.edu.tr

***Abstract*— The design and analysis of wireless power transfer (WPT) systems involve multiple interrelated computational steps, such as coil geometry modeling, self- and mutual-inductance calculation, coupling coefficient calculation, operating medium characterization, and compensation topology evaluation. These computational steps often require different methods and tools. Their separate implementation complicates the analysis of WPT systems. In this study, Wireless Power Transfer Analysis Studio (WPT Studio) was developed to integrate all these processes into a unified analysis environment. WPT Studio is a modular, MATLAB-based computational framework for WPT system analysis. It also accounts for high-frequency conductor and operating medium effects and supports system-level analysis of series-series (SS), series-parallel (SP), parallel-series (PS), and parallel-parallel (PP) compensation topologies. The self-inductance calculation capability of WPT Studio was validated using two experimental planar square coil cases reported in the literature. For a 40-turn coil with a measured self-inductance of 399 microhenries, the Greenhouse model implemented in WPT Studio yielded 397 microhenries, corresponding to a relative error of 0.35 percent. For a second 5-turn coil with a measured self-inductance of 4.38 microhenries, the Grover model yielded 4.42 microhenries, corresponding to a relative error of 1.00 percent. These results demonstrate that the developed WPT Studio produces results in good agreement with independent experimental reference values.**

***Keywords*— Wireless power transfer, WPT Studio, electromagnetic modelling, inductance calculation.**

ISBN: 978-605-80193-6-2

Comparative Performance Analysis of Statistical and Deep Learning Models for Temperature Forecasting in Konya

Ayşe ELDEM[1] Habibe KARBEYAZ[2]

[1]Karamanoğlu Mehmetbey University, Faculty of Engineering, Department of Computer Engineering, Karaman, Turkey

[2]Karamanoğlu Mehmetbey University, Institute of Science, Department of Computer Engineering, Karaman, Turkey

* Corresponding author e-mail: ayseeldem@kmu.edu.tr

***Abstract*—This study examines the effectiveness of baseline reference methods, classical statistical models, and deep learning-based approaches in the short- and medium-term forecasting of daily average temperatures in Konya province. To compare the forecasting performance of the models using daily average temperature data from the Open-Meteo archive for the 2015–2026 period, the Persistence, Seasonal Naive, and Climatology reference approaches, along with the ARIMA, Fourier term SARIMAX, and multi-output LSTM models, were selected. To examine the effect of the amount of training data on accuracy, all models were trained using 70%, 80%, and 90% of the series. They were also tested using a rolling-origin protocol with 7-day steps. Forecast horizons ranged from 1 to 50 days (1, 5, 10, …, 45, 50). Prediction performance was evaluated using RMSE, MAE, and the Skill score. Additionally, the Diebold-Mariano test was used to determine whether the differences between the models were statistically significant. According to the results obtained, SARIMAX-Fourier was the model that produced the lowest error value in each horizon. On the common test set, the RMSE was approximately 1.47°C for the 1-day forecast and approximately 3.0°C for the 50-day horizon. Although the model's superiority over the Climatology reference decreased slightly as the prediction period increased, it managed to maintain its statistical significance ($p < 0.05$) across all horizons and different training data ratios. On the other hand, while LSTM produces fewer errors than Climatology in the medium and long term, it has only been able to achieve similar performance to SARIMAX-Fourier and has not been able to establish a significant advantage over it. LSTM's superior performance relative to Climatology persisted up to 50 days with a 70% training ratio, 40 days with an 80% ratio, and 20 days with a 90% ratio. Additionally, it was observed that LSTM performed quite poorly in very short-term (1-day) forecasts. ARIMA and Persistence models have fallen behind even Climatology after periods of 5-10 days. Another noteworthy finding is that the performance of SARIMAX-Fourier remained virtually unaffected by the amount of training data used. In long-term forecasts, skill scores approaching zero indicate that predictability is naturally quite limited over 50-day periods.**

***Keywords*— Temperature forecasting, time series analysis, ARIMA, SARIMAX-Fourier, LSTM**

 ISBN: 978-605-80193-6-2

Multi-Temporal Drought Analysis of Latakia Station Using the Standardized Precipitation Index

Akın DUVAN[1] Mohamed ZEINEDOIN[1]

[1]Karamanoglu Mehmetbey University, Engineering Faculty, Civil Engineering Department, Karaman, Türkiye

* Corresponding author e-mail: akinduvan@kmu.edu.tr

***Abstract*—Drought is a major hydroclimatic extreme event that occurs when precipitation remains significantly below long-term averages, exerting multidimensional impacts on ecological systems, agricultural activities, water resources, and socioeconomic structures. The effects of climate change on precipitation regimes and variability necessitate detailed investigations into the occurrence dynamics and temporal variability of drought events. The Eastern Mediterranean Basin is considered a region of significant drought risk due to its sensitivity to changing hydroclimatic conditions. Latakia, one of the important coastal regions of Syria under the influence of the Mediterranean climate, requires comprehensive assessments of meteorological drought due to changes in its precipitation regime. This study aims to characterize the temporal properties of meteorological drought at Latakia station using the Standardized Precipitation Index (SPI) based on monthly total precipitation data for the period 1992–2026. SPI values were calculated at 1-, 3-, 6-, 9-, 12-, and 24-month timescales to assess short- and long-term drought conditions. Dry and wet periods were identified, while drought severity, duration, frequency, and temporal variability were examined through graphical analyses. The results indicate substantial variability in precipitation across years and periods, resulting in temporally heterogeneous meteorological drought conditions. SPI analyses revealed the occurrence of mild, moderate, severe, and extreme drought categories with varying durations. Furthermore, differences were observed in drought persistence and the temporal integration of hydroclimatic impacts as the SPI timescale increased. Overall, the findings demonstrate that SPI is an effective and applicable method for identifying and characterizing meteorological drought at Latakia station. The results are expected to contribute to sustainable water resources management, agricultural planning, drought risk reduction, and regional drought monitoring efforts**

***Keywords*— Drought, SPI, Latakia, Syria, Eastern Mediterranean, precipitation variability**

ISBN: 978-605-80193-6-2

High Temperature Performance of Boron Waste Modified Mortars for Rapid Service Restoration in Fire Damaged Tunnels and Concrete Pavements

Kemal ARMAĞAN[1*], Mehmet CANBAZ[2]

[1]Karamanoğlu Mehmetbey University, Civil Engineering Department, 70100, Karaman, Turkey

[2]Eskişehir Osmangazi University, Civil Engineering Department, 26480, Eskişehir, Turkey.

* Corresponding author e-mail: kemalarmagan@kmu.edu.tr

Abstract*—Fire-induced damage to transportation infrastructure, particularly road tunnels and rigid concrete pavements, can result in significant deterioration, prolonged traffic closures, and disruptions to transportation networks. Rapid restoration of damaged infrastructure is therefore essential for maintaining network functionality and improving the resilience of transportation systems. This study investigates the high-temperature performance of rapid-repair mortars produced using calcium aluminate cement (CAC), basalt sand (0–2 mm), and industrial Kırka boron plant waste as a partial cement replacement at 0%, 1%, 3%, and 5% by weight. Prismatic mortar specimens (40 × 40 × 160 mm), cured for 28 days, were exposed to 20, 600, and 800 °C for 3 h to represent severe thermal conditions that may occur in fire-damaged transportation facilities. Physical, non-destructive, and mechanical performance was evaluated through unit weight, ultrasonic pulse velocity (UPV), compressive strength, and flexural strength tests. The results showed progressive reductions in physical and mechanical properties with increasing temperature. Nevertheless, the CAC–basalt mortar matrix maintained considerable residual mechanical performance after severe thermal exposure. The incorporation of 1% boron waste provided performance close to that of the control mixture, whereas higher replacement ratios resulted in more pronounced reductions. These findings indicate the potential of boron waste-modified CAC mortars as rapid-repair materials for the post-fire rehabilitation of road tunnels, rigid concrete pavements, and related transportation infrastructure. Their application may contribute to reducing infrastructure recovery time and supporting more resilient and sustainable transportation networks.

***Keywords*— Transportation infrastructure, tunnel fire, rapid service restoration, calcium aluminate cement, boron waste, high-temperature performance, ultrasonic pulse velocity.**

 ISBN: 978-605-80193-6-2

Valorization of Date Palm Leaf Biomass as a Sustainable Biosorbent for Cationic Dye Removal

Siham BOULKAIBET[1*], Chebli DERRADJI[2]

[1]Kasdi Merbah University of Ouargla

[2]Ferhat Abbas University Setif 1

* Corresponding author e-mail: sisiboulk5@gmail.com

***Abstract*—Date palm leaves represent an abundant lignocellulosic residue that can be valorized as a low-cost material for water treatment. In this study, untreated date palm leaf biomass was investigated as a sustainable biosorbent for the removal of methylene blue (MB) and crystal violet (CV) from aqueous solutions. The material was used without chemical activation and characterized by scanning electron microscopy (SEM), Fourier-transform infrared spectroscopy (FTIR), and X-ray diffraction (XRD). The results revealed a heterogeneous lignocellulosic surface containing oxygen-bearing functional groups and both ordered and amorphous structural domains. The point of zero charge was determined at pH 6.46. Batch adsorption experiments showed maximum removal efficiencies of 98% for MB and 97% for CV, while near-equilibrium conditions were reached within approximately 45–50 min. The highest experimentally measured equilibrium capacities were 47.64 mg g^{-1} for MB and 29.83 mg g^{-1} for CV. Nonlinear kinetic analysis indicated that the pseudo-first-order model provided the best statistical description for both dyes, whereas equilibrium data were better described by the Langmuir model. These findings support the valorization of date palm leaf residues as a sustainable biosorbent for water-treatment applications.**

***Keywords*— Date palm leaves, biosorption, Methylene blue, crystal violet, adsorption kinetics.**

 ISBN: 978-605-80193-6-2

Development and Commercialization of Legume and Vegetable Based Gluten Free Pasta Through University and Industry Collaboration

Mehmet Koyuncu[1,2*] Recep Çifçi[1,2] Şeyma Akın[2]

[1]Karamanoğlu Mehmetbey University, Faculty of Engineering, Department of Food Engineering, Karaman, Türkiye

[2]Sadre Gıda Ar-Ge Danışmanlık San. ve Tic. Ltd. Şti., Karaman Teknopark, Karaman, Türkiye

Article note: This study was supported by TÜBİTAK 1512 BiGG, Project No. 2220267.

* Corresponding author e-mail: mkoyuncu@kmu.edu.tr

***Abstract*—Gluten free pasta on the Turkish market is largely imported, expensive, and often of limited technological and nutritional quality, which restricts access for individuals with celiac disease. This study aimed to develop a legume and vegetable based gluten free pasta and to transfer it from the laboratory to the market through a university and industry collaboration carried out by a technology based start up in a regional technopark. The initial approach, based on corn and rice flours, was planned to be enriched with extruded red lentil and chickpea flours. Contrary to expectations, pasta containing extruded legume flours could not maintain its structural integrity during drying and disintegrated during cooking, probably because extrusion reduced the binding capacity of the flours. The formulation was therefore redesigned using only native red lentil and chickpea flours, without hydrocolloids or other structuring agents, together with a modified drying protocol and vegetable purees. The resulting pasta contained 20 to 25 percent protein and showed a cooking loss of approximately 14 percent, which is comparable to values reported for pasta made from native legume flours but higher than the limit defined for durum wheat pasta. Seven commercial prototypes were developed and introduced to domestic and international markets through online sales channels. The study shows that testing and abandoning an initial hypothesis can guide a regional start up toward a marketable product, whereas reducing cooking loss remains the main objective of future work.**

***Keywords*—Gluten free pasta, legume flour, extrusion, cooking loss, university and industry collaboration, regional development**

 ISBN: 978-605-80193-6-2

Improvement of Bread-Making Quality in Bread Wheat Lines Carrying a Rye Translocation

Bedrettin DEMİR[1], Zehra ERKOL[2], Nevzat AYDIN[1]

[1]Karamanoğlu Mehmetbey University, Vocational School of Technical Sciences, Chemical Technology Program, Karaman, Türkiye

[2]Karamanoğlu Mehmetbey University, Faculty of Engineering, Department of Bioengineering, Karaman, Türkiye

* Corresponding author e-mail: bedrettindemir@kmu.edu.tr

***Abstract*—Although the rye translocation provides many agronomic advantages in bread wheat, it reduces bread-making quality due to the secalin protein. In order to improve the gluten quality of a line carrying the rye translocation (RILs44), this study aimed to develop near-isogenic lines (NILs) into which the overexpressed *Glu-B1al* (Bx7OE) subunit and the *Glu-A3b, Glu-A3c and Glu-A3g* alleles were transferred. The RILs44 line was used as the recurrent parent in the backcrossing program. The study was conducted in a controlled greenhouse using the speed breeding method, with a photoperiod of 22 h light / 2 h dark. Backcrossing was advanced to the BC4 generation, and 1637 hybrid seeds were screened by SDS-PAGE and selected on the basis of their high- and low-molecular-weight glutenin subunits. Gluten Swelling Index (GSI) and protein content analyses were performed to determine the gluten quality of the resulting BC4F3 lines. With the speed breeding protocol, the generation time of the backcross progeny ranged from 60 to 73 days. GSI analyses showed that lines carrying the *Glu-A3e* and *Glu-B3j* alleles, which negatively affect gluten quality, had a GSI value of 2.986 mL, whereas this value increased to 3.040 mL in lines carrying the *Glu-A3b* and *Glu-B3j* alleles ($p < 0.01$). In terms of protein content, lines carrying the *Glu-A3b and Glu-B3j* alleles also ranked highest, with 20.8% on a dry matter basis ($p < 0.01$). The results showed that protein content and quality increased in rye translocation lines combining high- and low-molecular-weight glutenin subunits that positively affect bread-making quality.**

***Keywords*— Bread wheat, speed breeding, rye translocation (1BL.1RS), GSI (Gluten Swelling Index), HMW-GS and LMW-GS, near-isogenic lines (NILs)**

** This research was derived from a doctoral thesis project supported by the Scientific Research Projects (BAP) Coordination Unit under project no. 01-D-20.*

 ISBN: 978-605-80193-6-2

New Recombinations Detected in High-Molecular-Weight Glutenin Subunits Encoded by the Glu-B1 Allele in Bread Wheat

Bedrettin DEMİR[1]

[1]Karamanoğlu Mehmetbey University, Vocational School of Technical Sciences, Chemical Technology Program, Karaman, Türkiye

* Corresponding author e-mail: bedrettindemir@kmu.edu.tr

***Abstract*—In bread wheat (Triticum aestivum L.), raw material quality is an important breeding criterion in addition to grain yield. The main genotypic factor determining the technological performance, dough rheology and viscoelastic balance of wheat flour is the High-Molecular-Weight Glutenin Subunits (HMW-GS). Therefore, HMW-GS are of great importance in deciding for which raw material group a bread wheat will be used. In our bread wheat breeding studies, HMW-GS information of the genetic material is determined by the SDS-PAGE (Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis) method. In these studies, recombinations were detected in the subunits encoded by genes located at the Glu-B1 locus on chromosome 1B. These recombinations are 7x + 18y and 17x + 8y. In the literature review, no article reporting these recombinations in bread wheat was found. The recombinations obtained remained stable in subsequent generations and were also observed in the new progeny. In addition to being an important novelty as a source of biodiversity and genetic resources, the effect of these new recombinations on bread wheat quality is still being investigated.**

***Keywords*— Bread wheat, HMW-GS, recombination, SDS-PAGE, wheat breeding**

 ISBN: 978-605-80193-6-2

Monitoring Biodiversity for Ecosystem Health and Technological Applications

Yasir DOĞAN*[1] Elif YAVUZASLANOĞLU[1]

[1]Karamanoğlu Mehmetbey University, Vocational School of Technical Sciences, Department of Crop and Animal Production, Karaman, Turkey

* Corresponding author e-mail: yasirdogan@kmu.edu.tr

***Abstract*—Ecosystems, broadly categorized into terrestrial and aquatic systems on Earth, harbor a wide range of biodiversity. In ecological studies concerning the interactions of living organisms with each other and their abiotic environment, the population monitoring of living entities provides insights into ecosystem health and trends. Monitoring biodiversity is gaining importance for optimizing crop yields in both natural and agricultural systems within terrestrial ecosystems. More stable data can be obtained by monitoring organisms known as bioindicators within the soil environment, which is a relatively sheltered habitat. However, because the identification of these organisms requires specialized expertise and is costly, traditional methods remain slow and have limited practical applications. In this context, technological applications based on molecular diagnostics, geographic information systems, and remote sensing techniques offer effective monitoring opportunities over large areas. Next-generation sequencing studies and metagenetic analyses enable the large-scale and rapid assessment of soil biological assets. On the other hand, technological applications based on the use of vegetation indices provide detailed information regarding plant health. Early warning and decision-support systems based on the development of diseases and pest agents encountered in agricultural production areas not only offer practical solutions for reducing economic crop losses, but also make significant contributions to ensuring food safety and protecting human health. The collected data can be utilized to generate future projections, and these projection data contribute to shaping governmental policies. This study will present information on technological applications used in monitoring ecosystem health through biodiversity.**

***Keywords*— Decision support systems, geographic information systems, bioindicator organisms, ecosystem management**

 ISBN: 978-605-80193-6-2

A Study on Some Rheological Properties of Composite Flours Made from Wheat and Molokhia (*Corchorus olitorius* L.) Leaf Powder*

Büşra DADUŞ[1] and Yalçın COŞKUNER[1]

[1]Karamanoğlu Mehmetbey University, Engineering Faculty, Department of Food Engineering

***Abstract*—In this study, the rheological properties of the dough were investigated by adding molehiya (*Corchorus olitorius* L.) leaf powder (MLP) to the dough formulation in different proportions (0.0%, 2.5%, 5.0%, 7.5%, 10.0%, 12.5%, 15.0%, 17.5% and 20.0%). For this purpose, the rheological properties of the doughs were evaluated using Mixolab, farinograph and extensograph analyses. According to the results of the Mixolab analysis, no statistically significant change was determined in the dough stability while the water absorption of the flour increased with the addition of MLP ($p>0.05$). In farinograph analysis, water absorption and development time increased with the increase in MLP rate, and a decrease in stability value was observed. When the results of the extensograph analysis were examined, it was seen that the energy and resistance to extensibility values generally increased with the addition of MLP to the flour, while the dough extensibility values decreased. According to the results obtained, it was shown that the addition of MLP significantly affected the rheological behavior of the dough and this effect was especially evident on water absorption, dough development, stability and extensibility properties. Overall, increasing levels of MLP enhanced the water-holding capacity of the dough and resulted in changes in its viscoelastic properties.**

***Keywords*— Molehiya, (*Corchorus olitorius* L.), leaf powder, rheological properties, Farinograph, Extensograph, Mixolab.**

**This study was prepared by summarizing of a portion of Büşra Daduş's Master of Science thesis titled "Investigation of the usage possibilities of molekhia (Corchorus olitorius L.) leaf powder in homemade pasta production" and was supported by Karamanoğlu Mehmetbey University Scientific Research Projects Commission (KMU-BAP) within the scope of the project numbered 04-YL-25.*

ISBN: 978-605-80193-6-2

Bibliometric and Content Analysis of Plant Attached Sensors for Agriculture in the Konya Plain Project Region

Ali Ozcan[1,*] Omer Pektas[2] Muhammed Emin Bedir[2]

[1]Karamanoglu Mehmetbey University, Kamil Ozdag Faculty of Science,
Department of Chemistry, 70100 Karaman, Türkiye

[2]Karamanoglu Mehmetbey University, Vocational School of Technical Sciences,
Department of Electricity and Energy, 70200 Karaman, Türkiye

* Corresponding author e-mail: aliozcan@kmu.edu.tr

***Abstract*—Plant-attached sensors, including wearable, implanted and nanobionic devices, can measure plant water status and stress responses, but evidence for their use in irrigation management remains limited. This study assessed the literature's relevance to agriculture in the Konya Plain Project region of Türkiye, focusing on Karaman. Bibliometric and content analyses used large language model assistance and author checks on blinded samples. Two independent screening passes with Cohen's kappa of 0.95 retained 555 documents from 985 Scopus records. Of these, 415 original studies measuring signals on or in plants were coded for test environment, monitoring duration, power supply, wireless data transmission and crop. Results were compared with provincial crop and irrigation statistics and regional reports. Publication counts increased from 22 in 2020 to 128 in 2025, with China and the United States contributing most. Field testing was reported in 40 of the 415 studies, or 9.6 percent. One study met the combined criteria of field or greenhouse testing, at least one week of monitoring, and wireless or autonomous operation. Among studies published before 2019 and since 2023, the shares claiming agricultural applicability were 22 and 47 percent, respectively; the shares reporting field or greenhouse testing were 20 and 11 percent. Wheat, barley, maize, sugar beet and sunflower accounted for nearly three quarters of 1.3 million irrigated hectares in the provinces analysed, but appeared in 23 studies, including two on water status; tomato, a minor regional crop, appeared in 63. Karaman accounted for 47 percent of national malting barley area; maize and apple were major irrigated crops, while 86 percent of wheat area was rainfed. The findings support further field evaluation in regionally important crops. Shared trials involving universities, irrigation cooperatives and equipment manufacturers could assess sensor performance under local conditions.**

***Keywords*—Plant-attached sensors, in planta sensing, precision irrigation, bibliometric analysis, field evaluation, regional agriculture**

 ISBN: 978-605-80193-6-2

Development of New Durum Wheat Cultivars Using Speed Breeding and Biochemical Marker-Assisted Selection

Bedrettin DEMİR[1], Ziya DURU[2], Aziz DOĞANAY[2], Nevzat AYDIN[1*]

[1]Karamanoglu Mehmetbey Universitesi

[2]Duru Bulgur A.Ş.

* Corresponding author e-mail: nevzataydin@kmu.edu.tr

***Abstract*—Durum wheat (Triticum durum Desf.) is a hard-grained, tetraploid wheat species used in the bulgur and pasta industries. Türkiye is among the richest countries in the world in terms of wheat landraces. Ahmet wheat is a durum wheat landrace widely used in the bulgur industry and valued for its quality attributes. It is known to be the preferred raw material of Duru Bulgur, a company with production experience in bulgur dating back to 1935. In this study, the characteristics of Ahmet wheat as a bulgur raw material were investigated, and these traits were used in the development of new wheat cultivars. In this research, conducted within the framework of university–industry collaboration, a speed breeding method was employed, enabling four generations per year to be obtained under greenhouse conditions. High- and low-molecular-weight glutenin subunits of the lines derived from crosses were identified by SDS-PAGE, and selections were made accordingly. The developed materials were tested under field conditions for two years. Following registration trials, two durum wheat cultivars, named Beken and Fertilis, were registered. The cultivars are high-yielding and possess superior characteristics as bulgur raw material. Moreover, they are the first cultivars registered in Türkiye through the speed breeding method and the cultivars developed in the shortest time.**

***Keywords*— Durum wheat, speed breeding, wheat landraces, bulgur quality, cultivar development.**

 ISBN: 978-605-80193-6-2

Contaminants Formed During Food Processing

Muharrem Gelen[1*] Nizam Mustafa Nizamlıoğlu[1]

[1]Karamanoğlu Mehmetbey University, Faculty of Engineering, Department of Food Engineering, Karaman, Türkiye

* Corresponding author e-mail: biomuharremgelen@gmail.com

***Abstract*—Process-induced food contaminants are genotoxic, mutagenic, and carcinogenic compounds formed as a result of chemical reactions (such as the Maillard reaction, pyrolysis, and lipid peroxidation) occurring during processes like heating, cooking, fermentation, and refining. Numerous toxic compounds can emerge, notably acrylamide in starchy products; heterocyclic aromatic amines and polycyclic aromatic hydrocarbons in protein-rich foods such as meat and fish; 3-MCPD and glycidyl esters in refined vegetable oils; and ethyl carbamate, furan, and nitrosamines in fermented or canned products. At the cellular level, these contaminants trigger oxidative stress, DNA adduct formation, and organ damage, thereby posing severe public health risks including cancer, neurotoxicity, nephropathy, and cardiovascular disorders. Because the formation levels of these contaminants vary depending on the food matrix, temperature, processing time, and pH, mitigation strategies play a critical role. These strategies include pre-treatments (e.g., blanching, selecting raw materials with low precursor levels), exogenous additives (e.g., asparaginase enzyme, natural antioxidants), optimization of cooking conditions, and the use of innovative technologies such as air frying, high hydrostatic pressure, and ohmic heating. Consequently, rigorous implementation of holistic management systems - incorporating HACCP, GMP, and source-control mechanisms - is essential to keep exposure to process-induced contaminants within legal tolerance limits and to ensure food safety.**

***Keywords*— Food processing contaminants, Maillard reaction, acrylamide, toxicity**

 ISBN: 978-605-80193-6-2

Effects of Nixtamalization on Cooking Quality, Color, and Antioxidant Properties of Gluten-Free Pasta from Legumes and Pseudocereals

Mehmet Koyuncu[1,2] Şeyma Akın[1] Recep Çifçi[1] Muharrem Gelen[1] Abdulvahit Sayaslan[2]

[1]Sadre Gıda Arge Danışmanlık San. Tic. Ltd. Şti., Karaman Teknopark, Karaman, Türkiye

[2]Karamanoğlu Mehmetbey University, Faculty of Engineering, Department of Food Engineering, Karaman, Türkiye

Article note: This study was supported by TÜBİTAK 1507 KOBİ AR-GE (Project No: 7258014) and conducted in collaboration with Karamanoğlu Mehmetbey University (KMÜ GIDAMER).

* Corresponding author e-mail: mkoyuncu@kmu.edu.tr

Abstract*—This study investigated the effects of nixtamalization ($Ca(OH)_2$ solution, boiling 30 min, steeping 12 h) on the cooking quality, textural properties, color characteristics, and antioxidant capacity of gluten-free pasta produced from chickpea (Cicer arietinum L.), red lentil (Lens culinaris Medik.), buckwheat (Fagopyrum esculentum Moench), and quinoa (Chenopodium quinoa Willd.). This represents one of the first systematic applications of nixtamalization to these raw materials in Türkiye for gluten-free pasta production. Nixtamalized and untreated (control) whole grains were dried, hammer-milled, and sieved to optimized particle size distributions. Pasta was produced at 30% moisture content using a laboratory-scale extruder (Bottene PM35) and dried at 60 °C. Results revealed strikingly contrasting responses across raw materials. Quinoa pasta achieved the lowest cooking loss (4.64%) among all samples, representing a 57% reduction compared to its control (10.75%), while chickpea exhibited the highest cooking loss (39.33%), a nearly 5-fold increase over its control (8.44%). Nixtamalization substantially increased dry hardness of buckwheat pasta (from 1161 to 4600 g) while drastically reducing it in chickpea (from 9770 to 2125 g). Color analysis demonstrated pronounced darkening in buckwheat (L* from 45.27 to 28.31) and dramatic redness reduction in red lentil (a* from 29.98 to 12.31). Regarding antioxidant capacity (DPPH), nixtamalization enhanced activity in chickpea (from 25.94% to 30.19%) but substantially reduced it in buckwheat (from 77.26% to 28.31%) and quinoa (from 58.68% to 34.07%). All nixtamalized pasta maintained protein contents above 21%. These findings demonstrate that nixtamalization effects are fundamentally raw-material-dependent, showing particular promise for quinoa-based formulations while requiring careful consideration of trade-offs in chickpea and buckwheat applications.

***Keywords*—Nixtamalization, gluten-free pasta, chickpea, red lentil, buckwheat, quinoa, cooking quality, antioxidant, color**

 ISBN: 978-605-80193-6-2

PART II

FULL PAPERS

IntCERD 2026 – International Conference on Engineering and Regional Development

Adaptive Fault Prognosis for Ground Based Radio Navigation Equipment Using Multisensor Fusion

Huseynaga N. Haqverdiyev

National Aviation Academy, Department of Aerospace Devices, Baku, Azerbaijan

Article note: This paper is derived from the author's doctoral dissertation on intelligent monitoring of the technical condition of ground facilities used for flight radio support.

* Corresponding author e-mail: huseynaga1999@gmail.com

***Abstract*—Ground based radio navigation facilities must maintain stable radio frequency, electrical, and thermal parameters because gradual degradation can reduce service quality before a hard failure is detected by conventional threshold checks. This study proposes an adaptive multisensor prognosis method for early identification of degradation in radio navigation ground equipment. The method combines normalized radio frequency power deviation, carrier frequency error, modulation deviation, feeder mismatch, power supply ripple, and cabinet thermal gradient. A reliability score is calculated for every channel from signal stability, missing data, and estimated noise, after which the features are fused with adaptive weights. An interpretable health index and a trend based warning rule are then used to detect incipient degradation and estimate the remaining warning horizon. A reproducible simulation experiment was created with twenty four thousand observation windows representing normal operation and four fault mechanisms at three degradation levels. The proposed method was compared with a best single channel baseline, static feature concatenation, and principal component analysis with a support vector machine. In the simulated test set, adaptive fusion achieved 97.1 percent classification accuracy, a macro F1 score of 0.968, and an average early warning lead time of 26.7 minutes. It also preserved a macro F1 score above 0.93 at a signal to noise ratio of 10 decibels. The results indicate that sensor reliability weighting can improve robustness when individual measurements become noisy or temporarily unavailable. The method is intended as a decision support layer for ground radio technical maintenance and requires field validation before operational certification.**

***Keywords*—Condition monitoring, fault prognosis, multisensor fusion, radio navigation equipment, predictive maintenance**

I. INTRODUCTION

Ground based radio navigation equipment is part of the technical infrastructure that supports safe and regular flight operations. Its serviceability depends not only on the presence or absence of a complete failure, but also on the stability of transmitter power, carrier frequency, modulation parameters, antenna feeder matching, supply quality, and thermal conditions. Small deviations in several channels can develop simultaneously and may remain below individual alarm limits while the overall technical condition is already deteriorating. For this reason, a maintenance strategy based only on periodic inspection and fixed single parameter thresholds can detect some faults later than a condition based approach that evaluates the joint evolution of several signals.

Condition based maintenance and prognostics have been studied extensively for industrial assets. Jardine et al. described the diagnostic and prognostic logic of condition based maintenance [1], while Lei et al. systematized the complete prognostic chain from data acquisition to remaining useful life prediction [2]. Deep learning has expanded data driven health management, although model complexity, interpretability, and benchmarking remain practical limitations [3]. Statistical remaining useful life models also emphasize the importance of uncertainty and degradation trend information rather than isolated measurements [4]. These findings are directly relevant to aviation support equipment, where false alarms and missed degradation events both carry operational costs.

Multisensor fusion provides an additional advantage because heterogeneous measurements do not fail or degrade in the same way. Recent studies have shown that multisensor diagnosis can outperform single channel solutions under noise and nonstationary conditions [5]-[7]. However, direct concatenation treats all channels as equally reliable. In radio technical facilities this assumption is weak: a temperature sensor may be stable while an RF measurement channel is temporarily noisy, or a feeder measurement may be unavailable during a short maintenance interval. A fixed fusion rule can therefore transfer poor quality information into the diagnostic decision.

This paper develops a different contribution from general technical condition monitoring: it focuses specifically on adaptive early prognosis. The proposed method assigns a quality dependent weight to each sensor channel, creates a fused health representation, and estimates an interpretable health index whose trend is used to produce an early warning. The study has four objectives: to define a compact multisensor feature model for ground radio navigation equipment, to formulate a reliability weighted fusion rule, to evaluate its robustness with simulated fault injection and sensor degradation, and to translate the results into practical deployment recommendations. The numerical experiment is intentionally identified as simulation based; it is not presented as a substitute for certified field trials.

II. PROPOSED METHOD

A. Signal Model and Feature Preparation

Each observation window is represented by six condition variables: relative RF output power deviation, carrier frequency error, modulation deviation, antenna

 ISBN: 978-605-80193-6-2

feeder mismatch represented by voltage standing wave ratio, normalized power supply ripple, and internal thermal gradient. The variables were selected because they correspond to different physical degradation paths and can be acquired without changing the main control logic of the radio facility. Before fusion, every feature is standardized using parameters estimated only from the normal training subset. The normalization applied to channel i at window k is given in (1).

$$z_{i,k} = \frac{x_{i,k} - \mu_i}{\sigma_i + \varepsilon} \tag{1}$$

Here x is the measured feature, mu and sigma are the normal state mean and standard deviation, and epsilon is a small positive constant that prevents numerical instability. This transformation makes electrical, radio frequency, and thermal measurements comparable without removing their direction of change.

B. Adaptive Reliability Weighting

For each channel a quality score is estimated from three observable properties: short term stability, valid sample ratio, and signal to noise ratio. A channel with frequent missing values or rapidly increasing noise receives a smaller influence even when its instantaneous value is large. The normalized reliability weight is defined in (2), where q is the positive quality score and m is the number of sensor channels.

$$w_{i,k} = \frac{q_{i,k}}{\sum_{j=1}^{m} q_{j,k}} \tag{2}$$

The adaptive fused feature vector is then calculated by (3). The formulation preserves channel level interpretability because every contribution can be traced to its weight and standardized deviation. It also allows a channel to re enter the fusion process automatically after its quality recovers.

$$F_k = \sum_{i=1}^{m} w_{i,k}\, z_{i,k} \tag{3}$$

C. Health Index and Prognostic Warning

The fused representation is compared with the normal reference center. A probabilistic classifier supplies the probability of the normal state, while the distance term captures deviations that may not match a known fault class. The anomaly score in (4) combines both components. Parameter alpha was set to 0.55 in the experiments after validation on the development subset.

$$A_k = \alpha \parallel F_k - c_N \parallel_2 + (1-\alpha)\left(1 - p_N(F_k)\right) \tag{4}$$

For maintenance interpretation, the anomaly score is converted to a bounded health index using (5). Values close to one describe normal operation, whereas sustained movement toward zero indicates increasing degradation.

$$H_k = \exp(-A_k) \tag{5}$$

A linear robust slope is estimated over the latest twelve windows. When the health index crosses a warning level and the slope remains negative for three consecutive windows, the algorithm issues an early warning. The approximate warning horizon in minutes is obtained from (6), where Hmin is the maintenance intervention level, beta is the fitted health slope, and delta prevents division by a value close to zero.

$$T_{warn} = \frac{H_k - H_{min}}{\max(-\beta_k, \delta)} \tag{6}$$

Figure 1 summarizes the complete processing chain. The design intentionally separates measurement quality estimation from fault classification so that a noisy sensor affects the fusion weight before it affects the maintenance decision.

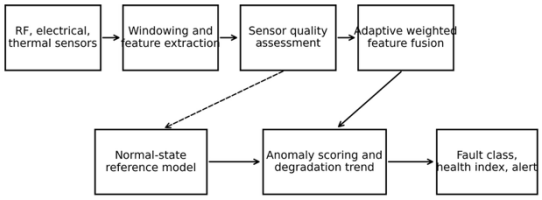


Fig. 1. Adaptive multisensor prognosis architecture for ground radio navigation equipment.

III. SIMULATION DESIGN

A. Fault Injection and Data Partitioning

Because an operational ground radio navigation dataset with synchronized fault labels was not available for this study, a reproducible synthetic experiment was used. Twenty four thousand ten second observation windows were generated. Forty percent represented healthy operation and the remaining windows represented four degradation mechanisms: RF power amplifier loss, oscillator drift, antenna feeder mismatch, and thermal power supply stress. Each fault mechanism was generated at early, moderate, and advanced degradation levels. Noise, ambient temperature, and load factors were randomized independently within each run. The purpose of the experiment was methodological validation of the fusion logic rather than certification of operational thresholds.

Runs rather than individual windows were separated into training, validation, and test groups using a sixty, twenty, twenty percent split. This prevented adjacent windows from the same simulated degradation trajectory from appearing in both training and testing. Ten random seeds were used, and the reported values are means across repetitions. The classifier was a compact support vector machine with a radial basis kernel trained on the fused representation. Hyperparameters were selected on the validation subset. The same classifier settings were used for the static feature baselines to isolate the effect of fusion.

Table I lists the simulated condition variables and the physical interpretation of their changes. The ranges are normalized engineering test ranges and must not be interpreted as certified maintenance limits for a specific VOR, ILS, DME, or other radio navigation model.

TABLE I. SIMULATED CONDITION VARIABLES

Variable	Normal band	Principal sensitivity
RF power deviation	-1.0 to 1.0 %	Amplifier loss
Frequency error	-8 to 8 Hz	Oscillator drift

ISBN: 978-605-80193-6-2

Modulation deviation	-1.5 to 1.5 %	RF path changes
VSWR	1.00 to 1.18	Feeder mismatch
Supply ripple	0.0 to 1.2 %	Power stress
Thermal gradient	-1.5 to 1.5 C/min	Thermal stress

B. Evaluation Metrics and Baselines

Four approaches were compared: the best individual sensor channel selected on the validation set, static concatenation of all standardized channels, principal component analysis followed by the same support vector machine, and the proposed adaptive reliability weighted fusion. Classification accuracy, macro F1 score, and one versus rest area under the receiver operating characteristic curve were calculated. For degradation trajectories, early warning lead time was measured from the first stable warning to the simulated intervention point. A separate robustness test added Gaussian noise corresponding to signal to noise ratios between 10 and 25 dB. Another test removed one sensor channel for a continuous five minute interval.

IV. RESULTS AND DISCUSSION

The main results are shown in Table II. Adaptive fusion produced the highest mean accuracy and macro F1 score while also increasing the average early warning lead time. The improvement over static concatenation is important because both approaches used the same underlying six measurements and classifier. The difference therefore results from the quality dependent contribution of each channel rather than from a larger input set.

TABLE II. COMPARATIVE PERFORMANCE ON THE SIMULATED TEST SET

Method	Accuracy	Macro F1	Lead time
Best single channel	88.6 %	0.874	12.4 min
Static concatenation	93.4 %	0.925	18.2 min
PCA plus SVM	95.0 %	0.943	20.6 min
Adaptive fusion	97.1 %	0.968	26.7 min

Fault specific analysis showed that oscillator drift was the easiest class to identify because frequency error developed monotonically and remained relatively independent of thermal noise. Early RF amplifier degradation was more difficult because power deviation partly overlapped with normal load variation. The fusion rule reduced this ambiguity by combining RF power, modulation, and thermal evidence. Antenna feeder mismatch also benefited from multisensor information because voltage standing wave ratio changes were accompanied by smaller but consistent RF power and temperature effects.

Figure 2 reports the noise robustness experiment. At 10 dB, the single channel baseline lost substantial discriminative power, whereas the proposed method retained a macro F1 score of 0.93. The advantage decreased as signal quality improved, which is expected because adaptive weighting becomes less necessary when all channels are reliable. This behavior is desirable: the method does not force complex weighting when measurements already agree.

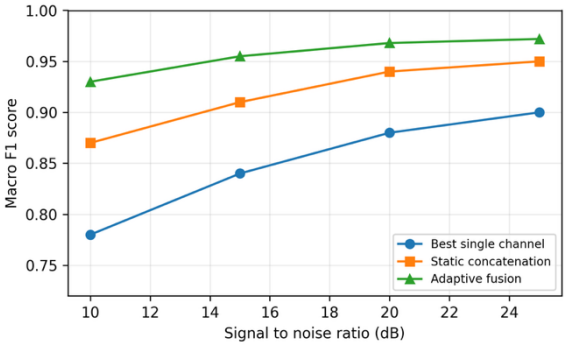


Fig. 2. Macro F1 score under increasing measurement noise in the simulation experiment.

An ablation analysis further examined the source of the gain. Replacing adaptive weights with equal weights reduced macro F1 from 0.968 to 0.937. Keeping adaptive weights but removing the temporal trend rule reduced average warning lead time from 26.7 to 19.1 minutes, although classification accuracy changed only slightly. When one sensor was unavailable for five minutes, macro F1 decreased to 0.949 rather than collapsing to the level of the single channel baseline. This confirms that the quality score acts as a practical fault containment mechanism for the monitoring layer.

The results are consistent with the broader literature on condition based maintenance, data driven prognosis, and multisensor diagnosis [1]-[8]. They also support the aviation prognostic principle that degradation trajectories and sensor selection should be treated jointly rather than independently [9]. Nevertheless, the numerical values should be interpreted with caution. Synthetic data cannot reproduce all environmental interference, calibration drift, electromagnetic coupling, maintenance actions, or subsystem specific protection logic found in operational radio navigation stations. The next research stage should therefore use recorded maintenance histories and controlled bench faults to recalibrate the normal reference distribution and warning thresholds.

V. PRACTICAL DEPLOYMENT RECOMMENDATIONS

For a ground radio technical facility, the proposed method should be implemented as a supervisory decision support layer rather than as a direct replacement for certified protection and alarm functions. The first recommended step is read only acquisition from existing monitoring interfaces. Time stamps from RF, electrical, and thermal channels should be synchronized, and a rolling buffer should retain both raw measurements and calculated health indicators. This makes every warning auditable and allows maintenance engineers to reconstruct the sequence that produced it.

A three level alarm logic is recommended. The information level can indicate a temporary health reduction without maintenance action. The warning level can request engineering inspection when both the health index and degradation slope remain abnormal. The critical level should continue to rely on the certified equipment protection limits. In this architecture, the prognosis layer adds anticipation but does not weaken established safety barriers.

ISBN: 978-605-80193-6-2

The reliability weights should also be logged. A persistent low weight on one channel can identify an instrumentation problem even when the radio facility itself is healthy. This feature helps separate equipment degradation from sensor degradation. Model retraining should be performed only after reviewed maintenance events and controlled calibration records have been added to the reference dataset. Automatic online retraining is not recommended for the first operational implementation because it can absorb abnormal states into the normal model.

Finally, field validation should be organized by fault mechanism. Controlled tests can introduce small frequency offsets, feeder mismatch, thermal stress, and supply ripple within safe laboratory limits, while the prognosis output is compared with calibrated reference instruments. Detection probability, false alarm rate, warning lead time, and repeatability should be reported separately for each equipment family. This validation path converts the simulation based proof of concept into evidence suitable for engineering acceptance.

VI. CONCLUSION

This paper proposed an adaptive multisensor fault prognosis method for ground based radio navigation equipment. Unlike static monitoring approaches, the method estimates sensor quality before fusion and therefore reduces the influence of noisy or temporarily unavailable channels. A fused anomaly score is converted into an interpretable health index, and its temporal trend is used to generate an early warning horizon. In a reproducible simulation with twenty four thousand observation windows and four degradation mechanisms, the proposed method achieved 97.1 percent accuracy, a macro F1 score of 0.968, and an average simulated warning lead time of 26.7 minutes. Robustness tests showed a clear advantage under noise and sensor dropout. The principal contribution is not the specific numerical threshold but the architecture that combines channel reliability, health estimation, and prognosis in one explainable workflow. Future work will focus on bench measurements and operational maintenance records from ground radio technical facilities so that the parameters can be calibrated for real equipment and assessed against aviation maintenance requirements.

References

[1] Jardine, A. K. S., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. Mechanical Systems and Signal Processing, 20(7), 1483-1510. https://doi.org/10.1016/j.ymssp.2005.09.012

[2] Lei, Y., Li, N., Guo, L., Li, N., Yan, T., & Lin, J. (2018). Machinery health prognostics: A systematic review from data acquisition to RUL prediction. Mechanical Systems and Signal Processing, 104, 799-834. https://doi.org/10.1016/j.ymssp.2017.11.016

[3] Khan, S., & Yairi, T. (2018). A review on the application of deep learning in system health management. Mechanical Systems and Signal Processing, 107, 241-265. https://doi.org/10.1016/j.ymssp.2017.11.024

[4] Si, X. S., Wang, W., Hu, C. H., & Zhou, D. H. (2011). Remaining useful life estimation: A review on the statistical data driven approaches. European Journal of Operational Research, 213(1), 1-14. https://doi.org/10.1016/j.ejor.2010.11.018

[5] Huang, H., Tang, B., Luo, J., Pu, H., & Zhang, K. (2021). Residual gated dynamic sparse network for gearbox fault diagnosis using multi-sensor data. IEEE Transactions on Industrial Informatics. https://doi.org/10.1109/TII.2021.3099060

[6] Yang, J., Gao, T., Zhang, H., & Li, Y. (2023). A multi-sensor fault diagnosis method for rotating machinery based on improved fuzzy support fusion and self-normalized spatio-temporal network. Measurement Science and Technology, 34(12), 125112. https://doi.org/10.1088/1361-6501/acef4a

[7] Liu, J., Xie, F., Zhang, Q., Lyu, Q., Wang, X., & Wu, S. (2024). A multisensory time-frequency features fusion method for rotating machinery fault diagnosis under nonstationary case. Journal of Intelligent Manufacturing, 35. https://doi.org/10.1007/s10845-023-02198-x

[8] Mao, G., Wang, J., Sun, Y., & Kim, J. (2026). An explainable intelligent fault diagnosis for rotating machinery via multi-source information fusion under noisy environments and small sample conditions. Sensors, 26(5), 1713. https://doi.org/10.3390/s26051713

[9] Saxena, A., Goebel, K., Simon, D., & Eklund, N. (2008). Damage propagation modeling for aircraft engine run-to-failure simulation. In 2008 International Conference on Prognostics and Health Management (pp. 1-9). IEEE. https://doi.org/10.1109/PHM.2008.4711414

[10] Liu, K., Gebraeel, N. Z., & Shi, J. (2013). A data-level fusion model for developing composite health indices for degradation modeling and prognostic analysis. IEEE Transactions on Automation Science and Engineering, 10(3), 652-664.

ISBN: 978-605-80193-6-2

Day-Ahead Solar PV Power Generation Forecasting Using Physics-Based Synthetic Data Generation and XGBoost

Nihat Yılmaz[1] İsmail Vedat Yılmaz[1] Osman Kaan Yıldırım[1] Haydar Doğan[1] Süleyman Sungur Tezcan[2]

[1]Pratikus Electrical Consultancy Ltd. Co. Ankara, Turkey

[2]Gazi University Department of Electrical - Electronic Engineering Ankara, Turkey

* Corresponding author e-mail: stezcan@gazi.edu.tr

***Abstract*— In this study, a forecasting framework is proposed to predict the day-ahead hourly energy generation of solar power plants (SPPs). This framework combines physics-based synthetic data generation with the XGBoost algorithm and is designed to prevent temporal data leakage. A dataset comprising 26,304 hours of observations covering the 2023–2025 period was generated for a synthetic reference SPP located in Ankara, featuring an installed capacity of 1.2 MWp DC and 1.0 MW AC. The day-ahead forecasting problem was defined as predicting the energy generation for the 24 hours of the following day, with forecasts generated daily at 18:00. To evaluate the contributions of different information groups, three XGBoost models were compared: M1 (utilizing meteorology, time, and solar geometry), M2 (incorporating historical generation features), and M3 (augmented with SCADA and operational data). The study's primary contribution is the integration of physics-based synthetic PV plant and SCADA data generation, an explicit operational forecasting setup, data leakage control based on information availability, and staged feature analysis within a reproducible experimental framework. The results are limited to synthetic data conditions; validation using actual PV plant and meteorological forecast data is required to determine real-world field performance.**

***Keywords*— Solar power plant, day-ahead energy forecasting, XGBoost, physics-based synthetic data, SCADA**

I. INTRODUCTION

The increasing share of renewable energy sources in electricity generation brings about new requirements for the generation planning and operation of electric power systems. Furthermore, the variable nature of PV generation—dependent on meteorological conditions—complicates the management of generation uncertainty in power systems with high PV penetration [1].

The output of a solar power plant (SPP) is influenced by factors such as solar irradiance, cloud cover, ambient temperature, wind speed, PV module temperature, and plant operating conditions. Therefore, accurate forecasting of PV generation is crucial for grid operation, energy management, and the reliable integration of renewable energy sources into the electrical system [1], [2].

Day-ahead PV generation forecasting is operationally significant as it enables the determination of the generation profile for the following day in advance. Recent studies treat day-ahead PV forecasting as a multi-step forecasting problem aligned with actual operating conditions and examine the impact of the forecast horizon on model performance [2]–[4].

Under real-world operating conditions, the actual solar irradiance, temperature, wind speed, or cloud cover values for the following day are unknown at the time of forecasting. Consequently, day-ahead models utilize Numerical Weather Prediction (NWP) or similar meteorological forecast data [2], [4]. However, significant discrepancies can exist between NWP values and actual local meteorological conditions, and these discrepancies can degrade PV power forecasting performance [5]. The fact that NWP irradiance errors exhibit varying behavior under different weather conditions indicates that the meteorological uncertainty inherent in day-ahead solar forecasting must also be addressed.

Depoortere et al. [6] highlighted that high-quality PV data is not always available in sufficient quantities in practice and addressed this issue by employing transfer learning with synthetic data generated via PVGIS. Their work demonstrates that synthetic PV data can be utilized to develop data-driven forecasting models, particularly in scenarios where the volume of actual observational data is limited.

Physical models, statistical methods, and machine learning-based approaches are employed for PV generation forecasting. In recent years, methods such as artificial neural networks, random forests, LSTM, GRU, Transformers, and gradient boosting have been widely investigated for PV forecasting tasks, largely due to their ability to model non-linear relationships [1], [7], [8]. XGBoost is also utilized for PV generation forecasting, serving both as a standalone forecasting algorithm and as a component of hybrid architectures combined with other learning methods [7], [9].

The problem of data scarcity is also addressed in contemporary studies that integrate physical modeling with machine learning. Hybrid approaches combining physics-based models with machine learning are proposed for PV forecasting applications where data availability is limited [10]. However, it must be noted that synthetic data does not substitute for real-world field validation, and forecasting performance achieved using synthetic data does not guarantee that the same level of performance will be attained at an actual power plant.

This study proposes an XGBoost-based day-ahead solar power plant (SPP) energy generation forecasting method that utilizes physics-based synthetic data generation for scenarios where actual plant data is unavailable. Synthetic data generation is carried out within a causal framework that encompasses time and solar geometry, meteorological conditions, PV module

ISBN: 978-605-80193-6-2

temperature, DC power generation, system losses, degradation, soiling, inverter conversion, power clipping, operational events, and SCADA variables.

The proposed method employs a three-stage XGBoost model to examine the contribution of different information groups to forecasting performance. Model M1 incorporates day-ahead meteorological forecasts, temporal variables, and solar geometry; Model M2 adds historical energy generation features to these; and Model M3 further includes SCADA and plant operational status variables. Thus, the aim is not only to determine the final forecast accuracy but also to systematically analyze the incremental contribution of different information groups to forecasting performance.

The primary contributions of this study can be summarized as follows:

1. A three-year hourly synthetic solar power plant (SPP) dataset was generated by integrating variables related to solar geometry, meteorology, PV thermal behavior, DC/AC energy conversion, degradation, soiling, clipping, derating, and SCADA systems into a unified data generation chain that respects physical relationships.

2. The day-ahead forecasting problem was defined operationally not as a fixed (t+24) target, but as the prediction of the 24-hour generation profile for day D+1, performed at 18:00 on day D.

3. Historical generation and SCADA features were constructed based on information availability at the time of the forecast rather than the target time, thereby aiming to prevent temporal data leakage.

4. The model was not provided with realized future meteorological values; instead, imperfect synthetic day-ahead forecast surrogates were constructed using an explicitly defined stochastic error model.

5. The incremental contributions of meteorology and solar geometry, historical energy generation, and SCADA/operational data to forecasting performance were compared using a staged experimental design (M1–M3).

II. MATERIALS AND METHODS

In this study, physics-based synthetic data generation is combined with XGBoost-based machine learning for day-ahead solar power plant (SPP) energy generation forecasting. The proposed approach consists of two main stages. In the first stage, a three-year hourly synthetic SPP dataset was generated by considering the physical relationships between solar geometry, meteorological conditions, PV module temperature, DC power generation, system losses, inverter conversion, and operational SCADA variables. In the second stage, the 24-hour energy generation profile for the following day was predicted using XGBoost models, with the forecast start time set at 18:00 each day.

A. Reference PV Plant and Dataset

A synthetic reference solar power plant (SPP) defined for Ankara, Turkey, was used in the study. The plant's location was assumed to be at approximately 39.93° N latitude, 32.86° E longitude, and an elevation of 900 m. The total installed DC capacity of the PV system was set at 1.2 MWp, and the installed AC capacity at 1.0 MW; consequently, the system's DC/AC ratio is 1.20. It was assumed that there are five inverters on the AC side, each with a nominal power of 200 kW.

The synthetic dataset covers a three-year period from 00:00 on January 1, 2023, to 23:00 on December 31, 2025, with hourly resolution. Since 2024 is a leap year, the total number of observations was calculated as 365×24 + 366×24 + 365×24 = 26,304.

B. Physics-Based Synthetic Data Generation

The objective of synthetic data generation is to preserve the fundamental physical dependencies within the PV energy conversion process, rather than simply generating a dataset composed of independent random variables. To this end, the data generation chain has been structured to sequentially encompass solar geometry, meteorological conditions, PV module temperature, DC power, losses, inverter conversion, AC power, and operational SCADA variables.

The main conversion models were specified explicitly to support reproducibility. Module temperature was calculated using a simplified Faiman-type solar irradiance–wind relationship [11]:

$$T_m = T_a + \frac{k_G G}{1 + k_w v} = T_a + \frac{0.030G}{1 + 0.10v} \quad (1)$$

In this equation; T_m: PV module temperature (°C), T_a: Ambient temperature (°C), G: Global horizontal irradiance (W/m²), v: Wind speed (m/s). The synthetic coefficients were (k_G = 0.030) °C·m²/W and (k_w = 0.10) s/m. DC power was calculated from irradiance scaling and a linear temperature correction:

$$P_{DC} = P_{DC,r} \frac{G}{1000} f_T (1 - L_b)(1 - d)^y (1 - L_s) \quad (2)$$

where; P_{DC}: DC power after losses (kW), $P_{DC,r}$: nominal installed DC power ($P_{DC,r} = 1200\, kWp$), $G/1000$: ratio of irradiance to reference irradiance, f_T: temperature correction factor ($f_T = 1 - 0.004(T_m - 25)$), L_b: base system loss ($L_b = 0.03$), d: annual degradation rate ($d = 0.005$), y: year elapsed since the start of the simulation, L_s: soiling loss rate. A simplified load-dependent inverter efficiency was used:

$$\eta_{calculation=} 0.975 - 0.035 e^{-8\frac{P_{DC}}{P_{AC,r}}} \quad (3)$$

The nominal installed AC power ($P_{AC,r}$) is 1000 kW. The inverter efficiency ($\eta_{inverter}$) is limited;

$$\eta_{inverter} = \begin{cases} 0.94 & \eta_{calculation} < 0.94 \\ \eta_{calculation} & 0.94 \le \eta_{calculation} \le 0.985 \\ 0.985 & \eta_{calculation} > 0.985 \end{cases} \quad (4)$$

During night-time hours, inverter efficiency is defined as zero.

C. Generation of Day-Ahead Meteorological Forecasts

In a real-world day-ahead PV forecasting system, actual meteorological measurements for the following day are not available at the time of forecasting. Therefore, directly feeding actual meteorological values

ISBN: 978-605-80193-6-2

into XGBoost models would result in an unrealistic and overly optimistic experimental design.

To address this issue, synthetic day-ahead meteorological forecast series were generated, distinct from the actual meteorological data series. Controlled forecast errors were introduced into the values for solar irradiance, cloud cover, ambient temperature, and wind speed. To prevent the violation of physical limits, the forecast values were constrained within the valid ranges of the respective variables.

D. XGBoost Models and Feature Analysis

Three different feature sets were created using the same XGBoost algorithm to examine the contribution of different information groups to day-ahead forecasting performance. Thus, the fundamental difference between the models was not the learning algorithm, but the amount of information provided to the model.

M1 is defined as the baseline model and includes the following feature groups:

- forecast horizon,
- target hour,
- day of the year,
- sine/cosine cyclical features for the hour,
- sine/cosine cyclical features for the day of the year,
- solar elevation and zenith angles,
- day-ahead forecast of solar irradiance,
- day-ahead forecast of cloud cover,
- day-ahead forecast of ambient temperature,
- day-ahead forecast of wind speed.

Thus, the M1 model represents the explanatory power of meteorological and astronomical conditions without utilizing historical energy or SCADA data.

In the M2 model, historical production data available at the start of the forecast were added to the features of the M1 model. The performance difference between M1 and M2 was used to examine the marginal contribution provided by historical production data, in addition to meteorological and solar geometry data.

The M3 is the most comprehensive model. The M3 model includes AC active power, DC power, inverter efficiency, heatsink temperature, internal inverter temperature, IGBT temperature, active alarm code, active warning code, derating status, and grid connection status.

The aim of this study is not to compare different machine learning algorithms, but rather to examine the values obtained using different data groups with the same learning algorithm.

E. Training, Validation, and Testing Protocol

To prevent future observations from leaking into the training process in time-series data, the dataset was not randomly shuffled but instead divided chronologically into three independent periods.

The year 2023 was used as the training period, 2024 as the validation period, and 2025 as the test period.

Due to the requirement for an initial 168 hours of historical data, the 2023 training set contains 357 valid target days. Accordingly, the data distribution is as follows in Table I.

TABLE I. The data distribution

Dataset	Year	Days	Hourly samples
Training	2023	357	8,568
Validation	2024	366	8,784
Test	2025	365	8,760
Total		1,088	26,112

F. Performance Evaluation Criteria

The prediction performance of the models was evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Bias Error (MBE), and Mean Absolute Percentage Error (MAPE).

In addition to the overall performance of the models, the marginal contribution of historical production and SCADA data groups was examined by evaluating the changes in error during the transitions from M1 to M2 and from M2 to M3.

III. FINDINGS AND DISCUSSION

A. Verification of Day-Ahead Meteorological Forecasts

To accurately represent day-ahead operational forecasting conditions, the XGBoost models were not provided with actual future meteorological values. Instead, day-ahead meteorological forecasts containing controlled errors were generated. The error values between the actual and forecasted meteorological variables are presented in Table II.

TABLE II. The error values in meteorological variables

Meteorological variable	MAE	RMSE	MBE
Solar radiation	19,88 W/m²	41,42 W/m²	−2,53 W/m²
Ambient temperature	1,60 °C	2,00 °C	+0,03 °C
Wind speed	0,93 m/s	1,16 m/s	+0,02 m/s
Cloudiness	9,26%	11,64%	+0,30%

The results indicate that the meteorological data used for forecasting differs from the actual observed weather conditions. In particular, the RMSE of 41.42 W/m² for solar radiation is significant, as it demonstrates that the model was not provided with perfect information regarding incoming radiation.

B. Day-Ahead Energy Forecasting Performance

The final performance of the day-ahead energy forecasting models was evaluated using the 2025 test period, which was kept independent of the training and validation stages. The test set comprises 365 forecast days and 8,760 hours of observations (365 days × 24 hours).

The performance results obtained from models M1, M2, and M3 are presented in Table III.

TABLE III. Day-ahead forecasting performance on the 2025 test dataset

ISBN: 978-605-80193-6-2

Model	Feature structure	MAE (kWh)	RMSE (kWh)	MBE (kWh)	MAPE (%)
M1	Meteorology + time + solar geometry	11,695	29,899	+1,782	5,726
M2	M1 + past production	11,562	29,840	+1,866	5,670
M3	M2 + SCADA/operation	**11,458**	**29,500**	+1,869	**5,630**

The baseline M1 model achieved MAE and RMSE values of 11,695 kWh and 29,899 kWh, respectively, using only day-ahead meteorological forecasts, temporal variables, and solar geometry. The MAPE calculated for positive generation observations is 5.726%.

In the M2 model, created by incorporating historical energy features available at the start of the forecast period, the MAE decreased to 11,562 kWh and the RMSE to 29,840 kWh.

The M3 model, which also included SCADA and operational status variables, yielded the lowest error values among the three feature configurations compared, with an MAE of 11,458 kWh, an RMSE of 29,500 kWh, and a MAPE of 5.630%.

The positive MBE values across all three models indicate a slight positive forecast bias on the test set. However, the MBE level of approximately 1.8 kWh represents a limited average deviation relative to the maximum hourly generation scale of 1,000 kWh.

C. The Contribution of Past Production and SCADA Data

The primary objective of the staged experimental design is not merely to identify the model with the lowest error, but to examine the additional contribution that different information sets provide to the baseline meteorology-based model.

In the transition from M1 to M2, the MAE decreased from 11.6953 to 11.5624. The relative improvement was calculated as (11.6953 - 11.5624) * 100 / 11.6953 = 1.1363%.

In terms of RMSE, an improvement of approximately 0.20% was achieved during the same transition. This result indicates that historical generation data provides additional yet limited information beyond the meteorological and solar geometry inputs.

In the transition from M2 to M3, the MAE decreased from 11.5624 to 11.4575, yielding an additional improvement of approximately 0.91%.

The RMSE dropped from 29.8397 to 29.5001, resulting in an improvement of approximately 1.14%.

When directly comparing the meteorology-based M1 model with the full-feature M3 model, the total reduction in MAE was calculated at approximately 2.03%, and the total reduction in RMSE at approximately 1.33%.

These results demonstrate that, within the synthetic experimental environment, meteorological forecasts constitute the primary source of information for PV energy forecasting, while historical generation and SCADA variables provide smaller but consistent additional information.

The first significant finding is that the M1 model achieves a low error level on its own. This is a physically expected result, given that PV generation is largely determined by solar irradiance and solar geometry.

The approximately 1.14% reduction in MAE upon the inclusion of historical generation data indicates that the past state of the PV generation series provides only limited additional information for day-ahead forecasting.

Incorporating SCADA and operational variables into the M3 model also reduced error values. The improvements of approximately 0.91% in MAE and 1.14% in RMSE when moving from M2 to M3 suggest that the plant's operational status at the start of the forecast period can provide additional information beyond meteorological and historical generation data. Variables such as inverter power status, temperatures, derating, and grid connection specifically represent the plant's current operational state.

Since the SCADA variables and the energy target are derived from the same physics-based synthetic generation chain, the relationships between them may be more regular than those found in a real solar power plant, which are subject to uncertainties arising from sensor noise, equipment variations, communication errors, and maintenance. Therefore, one should not conclude that the improvement achieved by M3 would be of the same magnitude under real-world field conditions.

Similarly, the MAPE value of 5.630% obtained in this study should not be interpreted as field performance validated on an actual solar power plant. The results demonstrate the comparative performance of XGBoost models within the context of the defined physics-based synthetic system and a controlled meteorological forecast error model.

Another noteworthy aspect of the results is the lack of significant differences between models M1 and M3. This should not be viewed negatively; rather, it suggests that meteorological forecasts and solar geometry already encapsulate the dominant physical information regarding PV generation, while historical generation data and SCADA variables provide only limited refinements to this foundational information. Furthermore, the constraint of using only features available at the forecast start time creates a forecasting problem that is operationally realistic yet more challenging.

Ultimately, the findings support the study's core methodological approach: for day-ahead PV forecasting, it is crucial to explicitly define not only the model algorithm but also the forecast start time, to exclude information that would not be available in the future, and to systematically segregate different information groups in order to reliably evaluate forecasting performance.

ISBN: 978-605-80193-6-2

IV. CONCLUSION

In this study, a forecasting framework was developed that combines physics-based synthetic data generation with the XGBoost algorithm to predict the hourly day-ahead energy generation of solar power plants, while accounting for information availability at the time of forecasting. The primary objective of the study is not merely to achieve low forecasting error, but to examine the contributions of meteorological forecasts, historical generation data, and SCADA/operational information within a controlled experimental setup that prevents temporal data leakage.

To this end, a total of 26,304 hourly observations covering the 2023–2025 period were generated for a synthetic reference solar power plant (SPP) located in Ankara, with a capacity of 1.2 MWp DC and 1.0 MW AC. The synthetic data generation process modeled solar geometry, meteorological conditions, PV module temperature, DC power generation, temperature losses, degradation, soiling, inverter efficiency, clipping, and various operational states within a unified physical data flow. Additionally, SCADA variables were created to represent inverter and plant operational conditions.

The day-ahead forecasting problem was defined by setting the forecast start time at 18:00 each day and predicting energy generation for the subsequent day across the 24-hour period from 00:00 to 23:00. Consequently, the forecast horizon ranges from +6 to +29 hours. Historical generation and SCADA features were derived from information available at the time of forecasting rather than relative to the target time. This approach prevented the inclusion of measurements taken after the forecast time in the model inputs, thereby establishing a controlled experimental protocol against temporal data leakage. Furthermore, instead of realized meteorological values for the future, synthetic day-ahead meteorological forecasts containing controlled errors were used.

Three XGBoost models were compared to examine the contributions of different information groups. Model M1 utilized meteorological forecast, time, and solar geometry features; Model M2 incorporated historical energy features available at the start of the forecast in addition to these; and Model M3 included SCADA and operational status variables alongside the features used in M2. The models were evaluated on data chronologically divided into training (2023), validation (2024), and testing (2025) periods.

On the 2025 test set, Model M1 achieved an MAE of 11,695 kWh and an RMSE of 29,899 kWh; in Model M2, which incorporated historical generation data, these values improved to 11,562 kWh and 29,840 kWh, respectively. Model M3, which also utilized SCADA and operational data, achieved the lowest overall error levels among the three feature configurations, with an MAE of 11,458 kWh, an RMSE of 29,500 kWh, and a MAPE of 5.630%. A comparison between M1 and M3 revealed reductions of approximately 2.03% in MAE and 1.33% in RMSE.

The findings indicate that meteorological forecasts and solar geometry provide the fundamental explanatory information for day-ahead PV energy forecasting, while historical generation and SCADA variables offer more limited contributions to overall performance beyond this baseline information. However, the primary contribution of this study is neither the development of a novel XGBoost algorithm nor solely the improvement of forecasting accuracy.

The primary limitation of the study is that the model development and evaluation processes were conducted using a synthetic SPP dataset. Deriving synthetic SCADA and energy variables from the same physical generation chain fails to capture the full range of sensor errors, communication interruptions, data losses, maintenance activities, and inter-equipment variations encountered in real-world field systems. Similarly, the day-ahead meteorological forecasts employed were not the outputs of an actual numerical weather prediction model but rather represented a controlled simulation of forecast errors. Therefore, the absolute error values obtained should not be interpreted as representing the field performance of an actual SPP.

References

[1] L. de Oliveira Santos, T. AlSkaif, G. C. Barroso, and P. C. M. de Carvalho, “Photovoltaic power estimation and forecast models integrating physics and machine learning: A review on hybrid techniques,” Solar Energy, vol. 284, Art. no. 113044, 2024, doi: 10.1016/j.solener.2024.113044.

[2] L. Massidda, F. Bettio, and M. Marrocu, “Probabilistic day-ahead prediction of PV generation. A comparative analysis of forecasting methodologies and of the factors influencing accuracy,” Solar Energy, vol. 271, Art. no. 112422, 2024, doi: 10.1016/j.solener.2024.112422.

[3] K. Tao, J. Zhao, Y. Tao, Q. Qi, and Y. Tian, “Operational day-ahead photovoltaic power forecasting based on transformer variant,” Applied Energy, vol. 373, Art. no. 123825, 2024, doi: 10.1016/j.apenergy.2024.123825.

[4] Z. Hu, Y. Gao, S. Ji, M. Mae, and T. Imaizumi, “Improved multistep ahead photovoltaic power prediction model based on LSTM and self-attention with weather forecast data,” Applied Energy, vol. 359, Art. no. 122709, 2024, doi: 10.1016/j.apenergy.2024.122709.

[5] M. Liu, Z. Lai, Y. Fang, and Q. Ling, “Day-ahead photovoltaic power forecasting based on corrected numeric weather prediction and domain generalization,” Energy and Buildings, vol. 329, Art. no. 115212, 2025, doi: 10.1016/j.enbuild.2024.115212.

[6] J. Depoortere, J. Driesen, J. Suykens, and H. S. Kazmi, “SolNet: Open-source deep learning models for photovoltaic power forecasting across the globe,” International Journal of Forecasting, vol. 41, no. 3, pp. 1223–1236, 2025, doi: 10.1016/j.ijforecast.2024.12.003.

[7] S. T. Asiedu, F. K. A. Nyarko, S. Boahen, F. B. Effah, and B. A. Asaaga, “Machine learning forecasting of solar PV production using single and hybrid models

 ISBN: 978-605-80193-6-2

over different time horizons," Heliyon, vol. 10, no. 7, Art. no. e28898, 2024, doi: 10.1016/j.heliyon.2024.e28898.

[8] J. Kim, J. Obregon, H. Park, and J.-Y. Jung, "Multi-step photovoltaic power forecasting using transformer and recurrent neural networks," Renewable and Sustainable Energy Reviews, vol. 200, Art. no. 114479, 2024, doi: 10.1016/j.rser.2024.114479.

[9] A. Khayat, M. Kissaoui, L. Bahatti, A. Raihani, Y. Atif, and K. Errakkas, "A novel hybrid GRU-XGBoost model for day-ahead photovoltaic generation forecasting in microgrids," Scientific African, vol. 29, Art. no. e02884, 2025, doi: 10.1016/j.sciaf.2025.e02884.

[10] H. Nourollahi Hokmabad, O. Husev, and J. Belikov, "Day-Ahead Solar Power Forecasting Using LightGBM and Self-Attention Based Encoder-Decoder Networks," IEEE Transactions on Sustainable Energy, vol. 16, no. 2, pp. 866–879, Apr. 2025, doi: 10.1109/TSTE.2024.3486907.

[11] D. Faiman, "Assessing the outdoor operating temperature of photovoltaic modules," Progress in Photovoltaics: Research and Applications, vol. 16, no. 4, pp. 307–315, 2008, doi: 10.1002/pip.813.

ISBN: 978-605-80193-6-2

Comparison of Gradient Boosting and Deep Tabular Models for Synchronous Generator Performance Prediction Under Geomagnetically Induced Current Conditions

Şevket Cantürk[1] Oktay Karakaya[2]

[1]Balıkesir University, Sındırgı Vocational School, Department of Electricity and Energy, Balıkesir, Türkiye

[2]Balıkesir University, Bigadiç Vocational School, Department of Electricity and Energy, Balıkesir, Türkiye

* Corresponding author e-mail: scanturk@balikesir.edu.tr

***Abstract*—Electromagnetic transient (EMT) simulations can represent the nonlinear effects of geomagnetically induced currents (GICs) on transformer saturation, harmonic generation, and synchronous generator behavior in detail. However, their computational cost becomes substantial when large parameter spaces are explored. This study compares gradient boosting and deep tabular learning approaches for rapid prediction of synchronous generator performance from physics-based EMT data. A total of 600 scenarios were generated using an EMT model developed in MATLAB/Simulink. The scenarios cover GIC magnitude, grounding resistance, load level, load power factor, and transmission-line impedance unbalance. The target variables are equivalent negative-sequence current, stator-current total harmonic distortion, terminal voltage, active power, and reactive power. XGBoost, LightGBM, and CatBoost were compared with FT-Transformer, NODE, and TabNet, while Ridge regression was used as a linear baseline. The models were evaluated using nested cross-validation with a five-fold outer loop and a three-fold inner loop. On the held-out test set, CatBoost achieved the highest accuracy with an R-squared value of 0.9999 and a MAPE of 0.23%, whereas FT-Transformer was the best-performing deep learning model with an R-squared value of 0.9996 and a MAPE of 0.44%. The marked reduction in Ridge performance for harmonic-related outputs highlights the nonlinear nature of GIC-driven behavior. Learning curves further showed that performance approached a plateau at approximately 400 training samples. All surrogate models provided a speedup of four to five orders of magnitude compared with the EMT simulations. These findings indicate that well-optimized gradient boosting models provide an effective balance between predictive accuracy and computational cost for small, deterministic, physics-based datasets.**

***Keywords*—Geomagnetically induced currents, synchronous generator, surrogate modeling, gradient boosting, deep tabular learning, CatBoost.**

I. INTRODUCTION

Geomagnetically induced currents (GICs) that arise in power transmission networks during geomagnetic disturbances can cause half-cycle saturation in grounded transformers, harmonic distortion, and increased reactive power consumption [1]. These effects can also propagate to synchronous generators via generator step-up transformers, altering negative-sequence currents, harmonics, terminal voltage, and active and reactive power behaviour [2], [3]. EMT models are therefore important tools for detailed analysis of generator behaviour under GIC conditions because they can directly represent nonlinear electromagnetic phenomena such as transformer saturation and harmonic generation.

A major limitation of EMT models is their high computational cost in parametric studies where several operating and network parameters vary simultaneously. In the system considered here, one EMT scenario requires approximately 16 min. Generating 600 operating points therefore corresponds to about 160 h of total simulation time. This computational burden motivates the use of data-driven surrogate models that retain information from detailed physics-based simulations while enabling rapid exploration of the parameter space.

Machine-learning-based surrogate models have been used in power systems for tasks such as harmonic estimation, transient stability assessment, and approximation of computationally expensive simulations [4]-[6]. Gradient boosting algorithms such as XGBoost, LightGBM, and CatBoost can provide high predictive accuracy for tabular datasets of limited size [7]-[9]. In contrast, more recent deep tabular architectures, including FT-Transformer, NODE, and TabNet, use attention mechanisms or learnable decision structures to capture complex relationships among features [10]-[12].

However, the conditions under which gradient boosting or deep tabular models offer an advantage for small, deterministic physics-based simulation datasets remain unclear. In particular, harmonic-related outputs and power-flow quantities may exhibit varying degrees of nonlinearity in synchronous generator performance under GIC conditions. A single aggregate accuracy metric may therefore be insufficient to characterise model behaviour. Both overall predictive accuracy and performance across physically distinct output groups should be evaluated.

In this study, three gradient boosting models (XGBoost, LightGBM, and CatBoost), three deep tabular models (FT-Transformer, NODE, and TabNet), and a linear Ridge baseline are compared using a physics-based dataset comprising 600 EMT scenarios. Nested cross-validation is used for model selection and performance assessment. In addition to test accuracy, sample efficiency and computational cost are examined. The main objective is to identify a surrogate model class that can provide both high accuracy and low computational cost for small-scale GIC-based physical datasets.

ISBN: 978-605-80193-6-2

II. DATASET AND SURROGATE MODELING METHOD

To evaluate surrogate modelling approaches under GIC conditions, a physics-based dataset was first generated using EMT simulations. The following subsections describe the dataset, the selected models, and the training procedure.

A. EMT-Based Dataset

The dataset was derived from a detailed EMT model developed in MATLAB/Simulink. The test system comprises a 250 MVA, 13.8 kV synchronous generator, a 13.8/154 kV generator step-up transformer, a 100 km, 154 kV transmission line, a 154/33 kV step-down transformer, and a load. The transmission line was modelled in the phase domain to capture phase unbalance. The five-dimensional input space was sampled using Latin Hypercube Sampling (LHS), yielding 600 scenarios [13], [14]. The dataset contains no missing or duplicate scenarios.

TABLE I. Input Parameters Used in the EMT Dataset

Input variable	Unit	Range
Induced DC voltage level	V	0-500
Grounding resistance, R_G	Ω	0.1-0.5
Load active-power level	pu	0.6-1.0
Load power factor (lagging)	-	0.9-1.0
Line-impedance unbalance coefficient	-	0-0.2

For each scenario, five generator performance indicators were recorded as outputs: equivalent negative-sequence current (I_{2eq}), stator-current total harmonic distortion ($THDI_G$), terminal voltage (Vt), active power (P_G), and reactive power (Q_G). I2eq and THDI reflect harmonic and rotor thermal behaviour, whereas V_t, P_G, and Q_G reflect voltage and power-flow performance. The dataset was randomly split into 80% training data (480 samples) and 20% test data (120 samples).

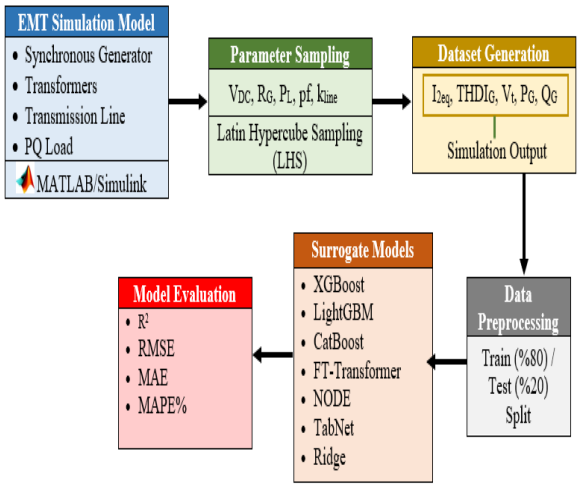


Fig. 1. Workflow for EMT-based data generation and surrogate model evaluation

B. Compared Models

Three gradient-boosting algorithms were included in the comparison: XGBoost [7], LightGBM [8], and CatBoost [9]. For the multi-output problem, these models were implemented as independent target regressors. The deep tabular group comprised FT-Transformer [10], NODE [11], and TabNet [12]. These architectures learn the five targets jointly via a shared multi-output layer. Ridge regression was included as a reference model to quantify the performance achievable with linear relationships.

C. Training and Evaluation Protocol

The input variables were standardised using the mean and standard deviation calculated solely from the training data. The target variables were also transformed in each outer cross-validation fold using scaling parameters derived solely from the corresponding training subset. This procedure prevented information leakage from the validation and test data into the training process.

Hyperparameters were selected via Bayesian optimisation using Optuna with the Tree-structured Parzen Estimator [15]. An unbiased estimate of generalisation performance was obtained through nested cross-validation with a five-fold outer loop and a three-fold inner loop [16]. Within each outer fold, hyperparameters were selected in the inner loop, after which the model was evaluated on the outer validation fold. Performance was quantified using R^2, RMSE, MAE, and MAPE. Finally, models with the selected hyperparameters were retrained on the complete training set of 480 samples and evaluated on the previously unseen test set of 120 samples.

III. RESULTS AND DISCUSSION

Model performance was assessed not only by overall prediction accuracy but also by its dependence on output type, training sample size, and computational cost. The corresponding results are discussed in the following subsections.

A. Nested Cross-Validation Results

The five-fold nested cross-validation results are presented in Fig. 2. CatBoost achieved the highest and most consistent performance across all metrics, with a mean R^2 of 0.9999 ± 0.0001 and a MAPE of 0.23 ± 0.10%. FT-Transformer was the best-performing deep tabular model, with R^2 = 0.9989 ± 0.0008, while XGBoost and LightGBM also achieved high accuracy. For Ridge, the mean R^2 decreased to 0.9791 and MAPE increased to 4.45%, indicating that the relationships in the dataset cannot be represented fully by a linear model.

B. Test Set and Physical Output Groups

The held-out test-set results are summarised in Table II. CatBoost achieved R^2 = 0.9999 and MAPE = 0.23%, followed by FT-Transformer with R^2 = 0.9996 and MAPE = 0.44%. XGBoost and LightGBM also achieved R^2 values above 0.998, whereas NODE and TabNet produced larger errors. Nevertheless, all nonlinear models retained high predictive accuracy for practical surrogate-model applications.

ISBN: 978-605-80193-6-2

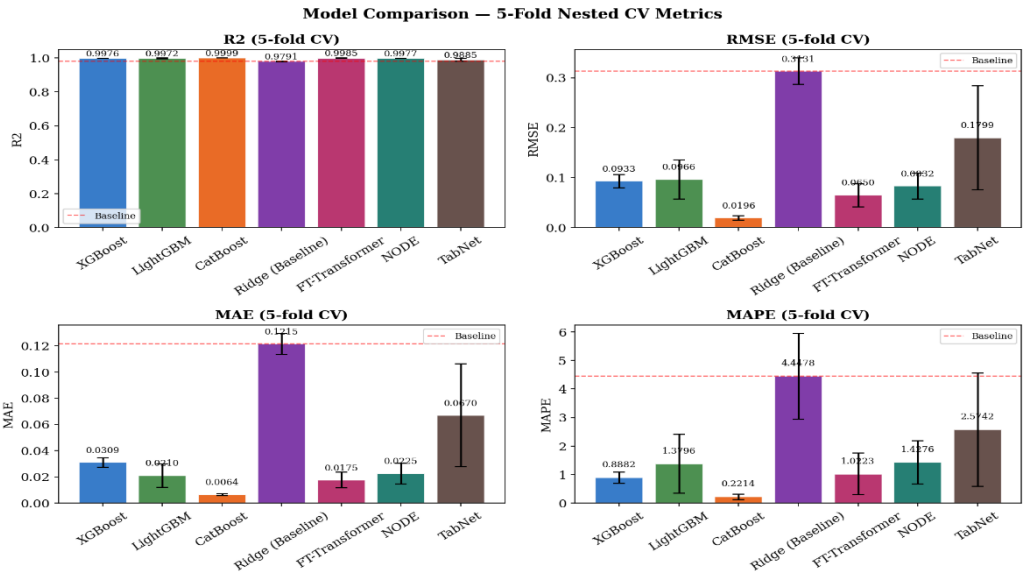


Fig. 2. Comparison of model performance under five-fold nested cross-validation

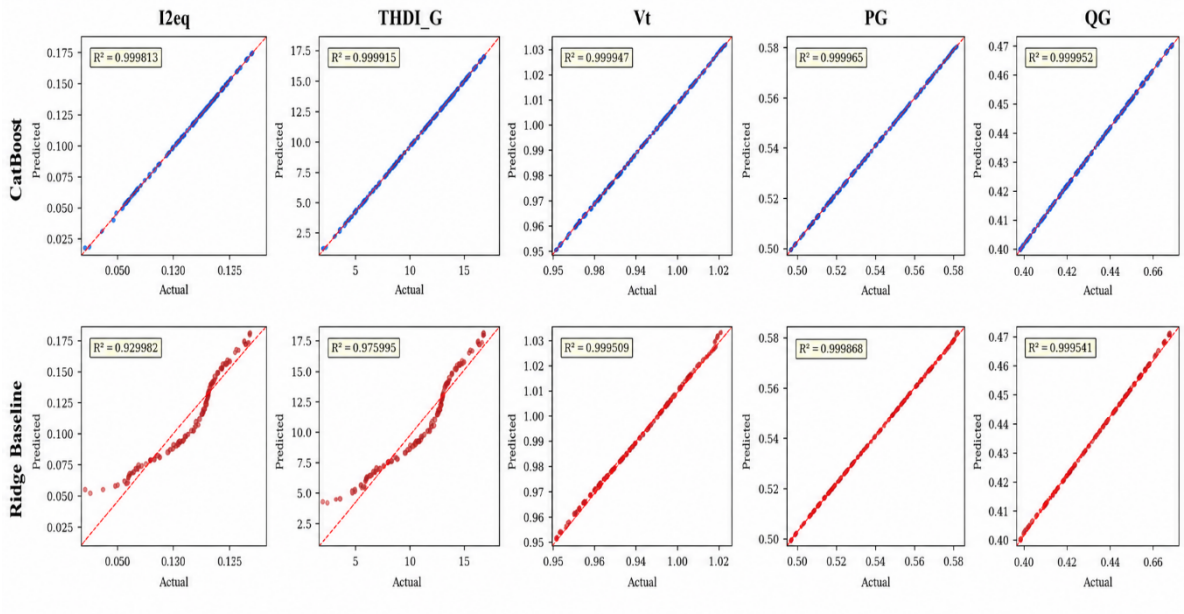


Fig. 3. Actual-versus-predicted comparison of CatBoost and Ridge for the five generator performance outputs

TABLE II. Overall Performance on the Test Set

Model	Type	R^2	RMSE	MAPE (%)
CatBoost	GB	0.9999	0.0159	0.23
FT-Transformer	Deep	0.9996	0.0289	0.44
XGBoost	GB	0.9987	0.0615	0.66
LightGBM	GB	0.9985	0.0748	1.05
NODE	Deep	0.9953	0.1306	2.13
TabNet	Deep	0.9910	0.1851	2.53
Ridge	Linear	0.9796	0.3001	4.59

The differences among the models become clearer when the outputs are divided into two physical groups. All models achieved very high accuracy for terminal voltage, active power, and reactive power, whereas the performance of the linear Ridge model decreased for the harmonic-related I2eq and THDI outputs. The test R^2 values of Ridge were 0.923 for I2eq and 0.976 for THDI, while CatBoost achieved $R^2 > 0.9998$ for both outputs. Fig. 3 illustrates this difference through actual-versus-predicted distributions. The result indicates that the nonlinear harmonic mechanism associated with transformer half-cycle saturation under GIC conditions requires nonlinear surrogate models, particularly for rotor-thermal and harmonic indicators.

C. Sample Efficiency

Because EMT data generation is computationally expensive, the sensitivity of model performance to the number of training samples was also examined. Fig. 4 presents three-fold cross-validation R^2 curves as the training-set size increases from 100 to 400 samples. CatBoost and XGBoost retain high accuracy even with relatively small training subsets, whereas FT-Transformer approaches a similar performance level as the amount of data increases. TabNet shows more variable behavior, particularly around 200 samples. Most models approach a performance plateau at approximately 400 training samples. This observation provides a practical indication of the number of simulations that may be required in comparable EMT-based surrogate studies.

ISBN: 978-605-80193-6-2

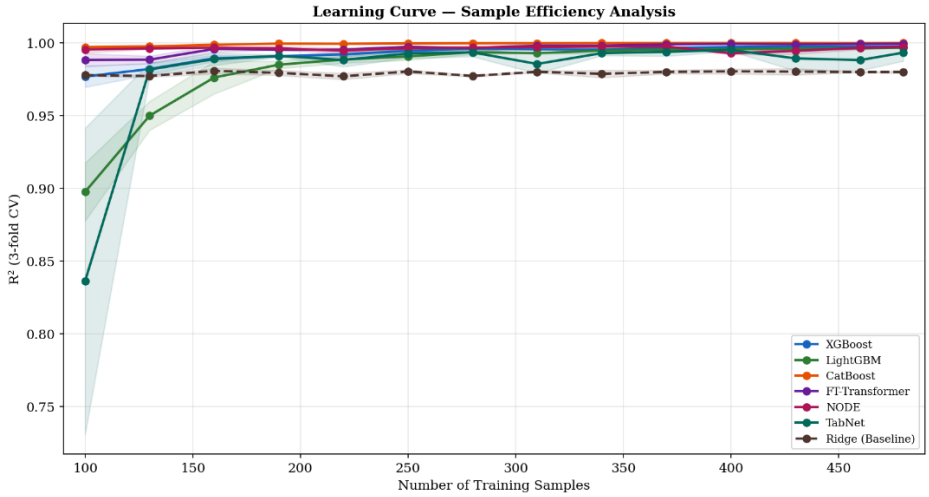


Fig. 4. Model performance as a function of training-set size (three-fold CV R^2)

D. Computational Cost

The primary purpose of surrogate models is to reduce the need to rerun a detailed EMT solution for every new operating point. Table III compare model training and inference times. Deep models generally require longer training times, with TabNet having the longest training time at 54.92 s. CatBoost requires 4.61 s for training, while its mean inference time over the test/validation group is 29.20 ms. All models achieve inference times below 160 ms and provide a speedup of approximately four to five orders of magnitude relative to the approximately 16 min EMT simulation. The reported batch inference times for XGBoost and LightGBM include the parallelisation overhead of MultiOutputRegressor; prediction time per individual sample remains below one millisecond. When accuracy and computational cost are considered together, CatBoost provides a strong balance between the two.

TABLE III. Model Accuracy and Computational Cost

Model	Test R^2	MAPE (%)	Training (s)	Inference (ms)
CatBoost	0.9999	0.23	4.61	29.20
FT-Transformer	0.9996	0.44	37.21	9.79
XGBoost	0.9987	0.66	1.07	155.32
LightGBM	0.9985	1.05	2.47	157.44
NODE	0.9953	2.13	19.63	8.73
TabNet	0.9910	2.53	54.92	44.23
Ridge	0.9796	4.59	0.001	0.17

IV. CONCLUSION

This study compared gradient boosting and deep tabular learning models for rapid prediction of synchronous generator performance under GIC conditions, using a physics-based EMT dataset comprising 600 scenarios. Nested cross-validation and independent test results showed that CatBoost achieved the highest overall predictive accuracy, while FT-Transformer was the best-performing deep tabular method. Power-flow-related outputs, including Vt, P, and Q, were predicted with high accuracy even by the linear model. In contrast, the marked reduction in Ridge performance for harmonic-related outputs such as I2eq and THDI confirmed the nonlinear character of GIC-driven electromagnetic behaviour. The learning curves indicated that model performance largely saturated at approximately 400 training samples. All surrogate models achieved inference times below 160 ms and provided a speed-up of four to five orders of magnitude relative to the EMT solution. These results suggest that complex deep architectures are not always necessary for small, deterministic, physics-based tabular datasets. Properly optimised gradient boosting models can provide a strong alternative in terms of accuracy, sample efficiency, and computational cost. However, the results remain dependent on the electromagnetic characteristics of the test system. Application to different systems would therefore require retraining with system-specific EMT data.

References

[1] Rajput, V. N., Boteler, D. H., Rana, N., Saiyed, M., Anjana, S., & Shah, M. (2021). *Insight into impact of geomagnetically induced currents on power systems:* Overview, challenges and mitigation. *Electric Power Systems Research, 192*, 106927. https://doi.org/10.1016/j.epsr.2020.106927

[2] Boguslawsky, I., Korovkin, N., Popov, V., & Sokolova, O. (2021). *Synchronous machine performance under geomagnetic disturba*nces. *High Voltage, 6*(5), 894-903. https://doi.org/10.1049/hve2.12107

[3] Jankee, P., Oyedokun, D. T. O., & Chisepo, H. K. (2022). *Dynamic response of power systems with real GICs:* Impact on generator excitation control. *IEEE Transactions on Power Delivery, 37*(6), 4911-4922. https://doi.org/10.1109/TPWRD.2022.3162881

[4] Taghvaie, A., Warnakulasuriya, T., Kumar, D., Zare, F., Sharma, R., & Vilathgamuwa, D. M. (2023). A comprehensive review of harmonic issues and estimation techniques in power system networks based on traditional and artificial intelligence/machine learning. IEEE Access, 11, 31417-31442. https://doi.org/10.1109/ACCESS.2023.3260768

[5] Chen, M., Liu, Q., Chen, S., Liu, Y., Zhang, C. H., & Liu, R. (2019). XGBoost-based algorithm interpretation and application on post-fault transient stability status prediction of power system. IEEE Access, 7, 13149-13158. https://doi.org/10.1109/ACCESS.2019.2893448

ISBN: 978-605-80193-6-2

[6] Li, X., Liu, C., Guo, P., Liu, S., & Ning, J. (2022). Deep learning-based transient stability assessment framework for large-scale modern power system. International Journal of Electrical Power & Energy Systems, 139, 108010. https://doi.org/10.1016/j.ijepes.2022.108010

[7] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785-794). ACM. https://doi.org/10.1145/2939672.2939785

[8] Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., & Liu, T. Y. (2017). LightGBM: *A highly efficient gradient boosting decision tree*. *In Advances in Neural Information Processing Systems 30* (pp. 3146-3154). Neural Information Processing Systems Foundation.

[9] Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: *Unbiased boosting with categorical features*. *In Advances in Neural Information Processing Systems 31* (pp. 6638-6648). Neural Information Processing Systems Foundation.

[10] Gorishniy, Y., Rubachev, I., Khrulkov, V., & Babenko, A. (2021). Revisiting deep learning models for tabular data. *In Advances in Neural Information Processing Systems 34* (pp. 18932-18943). Neural Information Processing Systems Foundation.

[11] Popov, S., Morozov, S., & Babenko, A. (2020). Neural oblivious decision ensembles for deep learning on tabular data. *In International Conference on Learning Representations*.

[12] Arik, S. Ö., & Pfister, T. (2021). *TabNet:* Attentive interpretable tabular learning. *Proceedings of the AAAI Conference on Artificial Intelligence, 35*(8), 6679-6687. https://doi.org/10.1609/aaai.v35i8.16826

[13] McKay, M. D., Beckman, R. J., & Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometrics, 21(2), 239-245. https://doi.org/10.1080/00401706.1979.10489755

[14] Loeppky, J. L., Sacks, J., & Welch, W. J. (2009). *Choosing the sample size of a computer experiment:* A practical guide. Technometrics*, 51*(4), 366-376. https://doi.org/10.1198/TECH.2009.08040

[15] Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 2623-2631). ACM. https://doi.org/10.1145/3292500.3330701

[16] Cawley, G. C., & Talbot, N. L. C. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. Journal of Machine Learning Research, 11(70), 2079-2107.

ISBN: 978-605-80193-6-2

Shared Battery Storage Planning for Renewable Source Coalitions: PSO-Based Sizing and Shapley Attribution of Pooling Benefits

Emine İŞİN[1] Nurettin ÇETİNKAYA[1]

[1]Konya Technical University, Faculty of Engineering and Natural Sciences, Department of Electrical and Electronics Engineering, Konya, Türkiye

* Corresponding author e-mail: eisin@ktun.edu.tr

***Abstract*—This study examines how pooling solar, wind, hydropower, and biomass generation in a shared battery energy storage system affects the required capacity and annualized cost. The four generation technologies are modeled as players in a cooperative game, and all 15 nonempty coalitions are evaluated. For each coalition, particle swarm optimization jointly determines the energy capacity, power rating, and initial state of charge shared by all design days. The data comprise 2,922 complete days constructed from half-hourly generation records. The 2018–2023 data form the design period, while the 2024 and 2025 data are used to assess the performance of the fixed design in subsequent years. In the profile-controlled analysis, each source is scaled so that the 99.5th percentile of its annual generation distribution equals 100 MW. In the fixed-scaling transfer analysis, the scaling factors derived from the design period are applied unchanged to the subsequent years, thereby retaining changes in generation scale. Coalition savings are defined as the difference between the annualized costs of separately sized batteries and a shared BESS and are allocated among the sources using exact Shapley values. PSO solutions are checked against deterministic capacity references and three independent random-seed batches. Under profile-controlled scaling, the annualized cost of the shared BESS is 54.93% lower than the sum for separate batteries, and wind receives the largest Shapley share at 33.19%. The savings function is superadditive, although the convexity condition fails for 10 coalition pairs. In the fixed-scaling transfer analysis, daily coverage falls from 95.03% in the design period to 91.80% in 2024 and 83.84% in 2025. The results show that pooling can lower the BESS cost, while a capacity fixed in the design period may not retain its initial coverage when the generation scale changes.**

***Keywords*—Battery energy storage system, particle swarm optimization, renewable resource complementarity, Shapley value, shared storage**

I. INTRODUCTION

As the shares of wind and solar energy in power systems increase, greater flexibility is required to balance generation variability and improve the predictability of power delivered to the grid. Hydropower and biomass, which have more regular generation characteristics, can smooth the aggregate renewable generation profile when operated alongside wind and solar. Temporal complementarity among resources and the flexibility and storage requirements arising from variable generation have been examined using different metrics [1]-[4]. However, combining resources does not reduce the required storage capacity in every case. In this study, battery energy storage system (BESS) requirements are determined by accounting for the magnitude and timing of generation deviations and their chronological effect on the battery state of charge. Therefore, renewable resource complementarity is evaluated using the required BESS power and energy capacities in addition to correlation coefficients.

In a shared BESS configuration, multiple generation sources use the same storage system. Pooling can reduce both the capacity requirements and cost relative to operating a separate battery for each source. Previous studies have examined shared storage mainly for energy communities, buildings, prosumers, and distributed energy resources [5]-[10]. The present study treats generation technologies as players and evaluates their effects on shared BESS requirements across all resource combinations. Complementarity among generation sources is governed by meteorological and hydrological conditions as well as their operating characteristics. Treating the sources as players makes it possible to quantify how each technology changes the shared BESS requirement across different resource combinations.

Determining the total savings provided by shared storage does not explain how these savings should be distributed among the sources. Each source's contribution depends on the coalition in which it participates. The Shapley value computes a player's average marginal contribution over all possible participation orders [11]-[15]. In this study, Shapley values are used to attribute the cost savings from a shared BESS to individual sources. The savings function satisfies superadditivity, whereas the convexity condition fails for 10 coalition pairs in the main analysis. Accordingly, the Shapley values are interpreted as the sources' average marginal contributions to pooled savings, without claiming that the allocation lies in the core.

Particle swarm optimization (PSO), a population-based optimization method, was used to size the BESS for each coalition [16]–[19]. The PSO jointly determines the BESS energy capacity, power rating, and common initial state of charge (SoC) for all design days. The solutions were compared with deterministic reference calculations, and the effect of random seeds was evaluated using three independent seed batches. The daily reference power was calculated from the full-day generation profile. Therefore, the analysis assumes that intraday generation is known in advance, and forecast errors and real-time operational uncertainties are outside its scope. Two analyses serve different purposes: one

ISBN: 978-605-80193-6-2

equalizes annual output scales to compare profile timing, while the other fixes scaling to test later-year BESS adequacy. In the profile-controlled analysis, each source and year was scaled such that the 99.5th percentile of the annual generation distribution equaled 100 MW. This removes the differences in generation magnitude across sources and years while preserving the profile shape. In the fixed-scaling transfer analysis, one scaling factor was calculated for each source from the 2018–2023 design period and applied unchanged to the 2024 and 2025 datasets. This procedure retained subsequent changes in the generation scale and tested the adequacy of the design-period BESS capacity.

The main contribution is the transfer of the shared BESS capacity determined for 2018–2023 to 2024 and 2025 without re-optimization and the assessment of its adequacy under changes in generation scale. In the profile-controlled case, pooling reduces the annualized cost by 54.93%, with wind receiving the largest Shapley share. Fixed-scaling transfer shows that the design-period capacity loses daily coverage in subsequent years. These findings distinguish the economic benefits of pooling from the ability of a fixed design to maintain service over time.

II. MATERIALS AND METHODS

A. Dataset and Resource Scaling

Half-hourly generation data for Great Britain published by the National Energy System Operator are used [20]. Solar, hydropower, and biomass use the corresponding dataset columns, and wind is the sum of centralized and embedded wind. The NESO corrects missing or irregular observations. Transmission-connected solar is classified as OTHER and excluded from the solar series used here [20]. Time stamps are ordered in Coordinated Universal Time (UTC), duplicate and invalid records are screened, and complete days containing 48 half-hourly observations are retained. The 2018–2025 period, for which complete annual data are available for all four sources, comprises 140,256 time steps and 2,922 complete days. The 2,191 days from 2018 to 2023 were used for BESS sizing. The selected BESS design was evaluated using the 2024 and 2025 data without re-optimization.

In the profile-controlled analysis, each source is scaled such that the 99.5th percentile of its annual generation distribution equals 100 MW:

$$\tilde{P}_{i,y}(t) = 100\,\frac{P_{i,y}(t)}{Q_{99.5}(P_{i,y})} \tag{1}$$

Here, $P_{i,y}(t)$ denotes the output of source i in year y, while $Q_{99.5}(P_{i,y})$ denotes the 99.5th percentile of the same generation series. Linear scaling removes absolute differences in generation magnitude across sources and years while preserving the temporal shape of each annual series. In the fixed-scaling transfer analysis, a single scaling factor is calculated for each source from the 2018–2023 design period and applied unchanged to the 2024 and 2025 data. This preserves subsequent changes in generation scale and tests the adequacy of the design-period BESS capacity under the same scaling factors. A third approach, which equalizes annual mean generation across sources, is used only to assess sensitivity to profile shape.

Solar, wind, hydropower, and biomass are each defined as players in the cooperative model. The four sources form 15 nonempty coalitions. The aggregate output of each coalition is obtained by summing the synchronized generation series of its members.

B. Chronological BESS Model

For each coalition, the constant daily reference power $P_{f,d}$ is determined such that the net change in the stored battery energy over each day is zero after accounting for the charging and discharging losses. Defining $r_d(t)=P_S(t)-P_{f,d}$ as the power imbalance, the energy balance is

$$\sum_{t\in d}[\eta_{ch}(r_d(t))^+ - \frac{(-r_d(t))^+}{\eta_{dch}}]\Delta t = 0 \tag{2}$$

where $(x)^+=\max(x,0)$, $\Delta t=0.5$ h, and $\eta_{ch}=\eta_{dch}=0.95$. The cumulative battery energy within a day is

$$X_d(k) = \sum_{\tau=1}^{k}[\eta_{ch}(r_d(\tau))^+ - \frac{(-r_d(\tau))^+}{\eta_{dch}}]\Delta t \tag{3}$$

The daily energy and power requirements are then

$$U_d = max_k\, X_d(k),\;\; D_d = -min_k\, X_d(k),\;\; P_d^{req} = max_t\,|r_d(t)| \tag{4}$$

where E and P denote the BESS energy and power capacities, respectively, and s_0 is the common initial SoC for all the design days. A day is covered if

$$U_d \le E(s_{max} - s_0),\;\; D_d \le E(s_0 - s_{min}),\;\; P_d^{req} \le P \tag{5}$$

All three conditions must hold. The SoC limits are set to $s_{min}=0.10$ and $s_{max}=0.95$. Because the daily reference is energy-neutral, every covered day begins and ends at the common s_0 value. On uncovered days, the battery remains idle and SoC stays at s_0. Design feasibility is defined as

$$Cov(E,P,s_0) = \frac{1}{N_d}\sum_{d=1}^{N_d} I_d(E,P,s_0) \ge 0.95 \tag{6}$$

where Id equals 1 when all conditions in (5) are satisfied. Daily coverage is the proportion of complete days satisfying these conditions, rather than the proportion of half-hourly intervals served. The assumed 95% target permits up to 5% uncovered days instead of sizing for every extreme day; it is not a regulatory standard. As covered days can differ among coalitions, 100% coverage is also examined. Because the daily reference power is calculated from the realized full-day generation profile, the model assumes full knowledge of each day's profile.

C. PSO-Based BESS Sizing

For each coalition, the decision vector $x=[E,P,s_0]$ is defined, and the annualized capacity cost is minimized. Energy- and power-related capital costs are represented separately, consistent with the model structure in [21]. The annualized cost is

$$C_{ann} = (CRF + f_{OM})(c_E E + c_P P),\;\; CRF = \frac{r(1+r)^L}{(1+r)^L - 1} \tag{7}$$

ISBN: 978-605-80193-6-2

In the base case, c_E=209,000 USD/MWh, c_P=76,000 USD/MW, the discount rate r=0.07, the project lifetime L=12 years, and the annual operation and maintenance rate f_{OM}=0.02. The coefficients define the base-case weighting of energy and power capacity for the pooling comparison; they are scenario inputs, not market-price estimates or values taken from [21]. The four-source design is also evaluated for independent ±20% changes in each unit-cost coefficient. A supplementary allocation calculation holds the 15 base-case coalition designs fixed and combines energy- and power-cost multipliers of 0.8, 1.0, and 1.2 in nine scenarios. The costs and Shapley shares are recalculated without re-optimizing capacity.

Particles that fail to satisfy the coverage constraint receive a penalty coefficient of 10^6. Each PSO run uses 42 particles and at most 110 iterations. The inertia weight decreases linearly from 0.90 to 0.40, the cognitive and social coefficients are both 1.70, and the particle velocity is limited to 25% of the decision range. A run terminates when the relative improvement in the best objective value remains below 10^{-10} for 30 consecutive iterations. In the profile-controlled analysis, 717 of the 900 runs terminated at iteration 31, and the mean number of iterations was 45.7.

For initialization, 65 discrete values were evaluated over the common SoC range, and the minimum-cost E-P pair satisfying 95% coverage was calculated at each SoC value. The 17 lowest-cost solutions from the discrete SoC search were used to initialize particles in every PSO run, and the remaining particles were initialized randomly. PSO searches SoC continuously between these discrete values. The seed test measures the variability conditional on these common deterministic initial solutions. It does not establish independence from initialization or certify a global optimum; agreement with the denser reference and the doubled-bound test provides separate numerical checks.

D. Shapley Allocation and Validation

Let C*(S) denote the lowest annualized cost found for coalition S. Shared-storage savings are defined as

$$v_C(S) = \sum_{i \in S} C^*(\{i\}) - C^*(S) \quad (8)$$

Reference [15] evaluated Shapley approximation methods for larger energy communities. With four players, all 15 nonempty coalitions are evaluated here. With $v_C(\emptyset)$=0, the Shapley values are computed without sampling from the coalition costs found by PSO:

$$\varphi_i = \frac{1}{n!} \sum_{\pi \in \Pi(N)} [v_C(P_i^\pi \cup \{i\}) - v_C(P_i^\pi)] \quad (9)$$

Here, π denotes the set of sources that precede source i in the participation order π. For each coalition, PSO is run 20 times in each of the three independent seed batches. Thus, each analysis mode comprises 60 runs per coalition and 900 PSO runs in total. The PSO solutions were compared with a deterministic reference based on an 81-point initial sweep and two successive 81-point refinements of the common initial SoC.

The prespecified acceptance thresholds are a feasible-run fraction of at least 80%, a coefficient of variation in cost no greater than 1%, and a positive cost deviation from the deterministic reference no greater than 1%. Both the range of Shapley shares across seed batches and the difference from the deterministic-reference shares must be no greater than 0.5 percentage points. Energy-bound use must remain below 98.5%, with doubled-bound cost improvement no greater than 0.1%. The energy-capacity upper bound is doubled, and the search is repeated with the main PSO budget. Sensitivity analyses use five runs with 34 particles and 80 iterations per point; the full-coverage coalition game uses six runs with 36 particles and 90 iterations per coalition. The design-period BESS capacity is evaluated unchanged using the 2024 and 2025 data. The savings function is also examined for superadditivity and convexity.

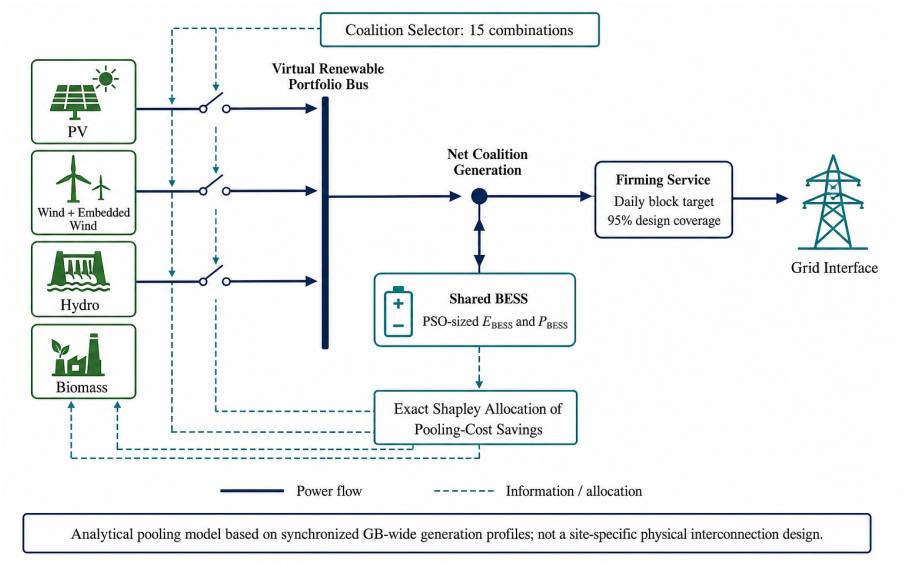


Fig. 1. Conceptual configuration of a shared BESS for renewable source coalitions

III. RESULTS

TABLE I. BESS sizing results for renewable source coalitions scaled to a 100 MW annual 99.5th percentile over 2018–2023

ISBN: 978-605-80193-6-2

Coalition	E* (MWh)	P* (MW)	Annual cost (million USD/year)
PV	480.61	74.97	15.49
WIND	477.49	50.79	15.12
PV+WIND	688.78	88.86	21.99
HYDRO	285.04	61.95	9.38
PV+HYDRO	559.77	78.49	17.94
WIND+HYDRO	468.93	64.69	15.02
PV+WIND+HYDRO	664.80	86.13	21.23
BIOMASS	338.26	50.45	10.87
PV+BIOMASS	610.53	79.52	19.50
WIND+BIOMASS	476.21	64.38	15.24
PV+WIND+BIOMASS	710.51	87.89	22.64
HYDRO+BIOMASS	494.58	71.42	15.87
PV+HYDRO+BIOMASS	710.40	84.45	22.60
WIND+HYDRO+BIOMASS	518.66	89.77	16.81
PV+WIND+HYDRO+BIOMASS	719.65	88.30	22.92

Note: Values refer to the 2018–2023 design period. The sum of the individual system costs is 50.86 million USD/year, the cost of the four-source shared BESS is 22.92 million USD/year, and the resulting savings are 27.94 million USD/year (54.93%).

A. Chronological Operation and Allocation of Cost Savings

In the profile-controlled analysis reported in Table I, the four-source shared BESS was sized at 719.65 MWh and 88.30 MW. Relative to the sum of the separately sized batteries, the energy capacity decreased by 54.49%, the power capacity by 62.92%, and the annualized cost by 54.93%. Adding a source reduces the power requirement in some coalitions: the 89.77 MW required by WIND+HYDRO+BIOMASS fell to 88.30 MW when solar was added. Figure 2 uses 12–18 June 2023 only to illustrate charging, discharging, and SoC, without statistical selection. The BESS balanced generation against constant daily grid delivery within 10%–95% SoC and returned to its initial SoC on covered days without discontinuity. BESS sizing and performance results use the full analysis periods.

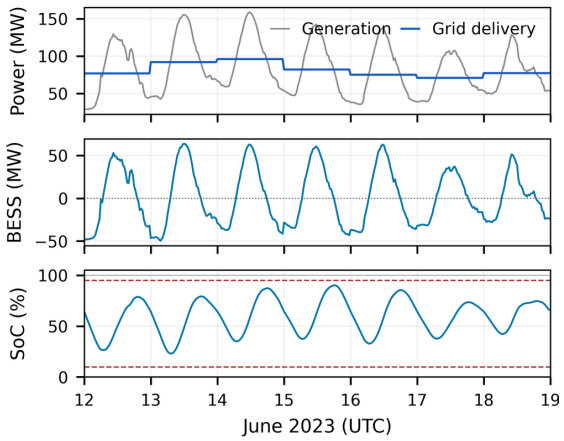


Fig. 2. Illustrative seven-day (12–18 June 2023) chronological operation of the shared BESS under profile-controlled scaling.

Shared storage yields annual cost savings of 27.94 million USD. The exact Shapley allocation assigns 33.19% of these savings to wind, 22.86% to solar, 22.39% to biomass, and 21.56% to hydropower (Figure 3). All shares are positive, satisfying individual rationality for the savings game, whose singleton values are all zero. These allocation results refer to the base case cost coefficients. Across the three independent seed batches, the largest range in savings shares was 0.052 pp, and the largest difference between the PSO and deterministic-reference shares was 0.046 pp.

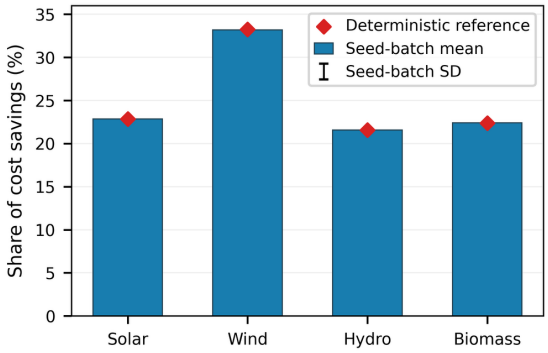


Fig. 3. Shapley shares of cost savings and seed batch variability under profile-controlled scaling.

B. Optimization Checks and Fixed-Scaling Transfer

All 900 PSO runs in the profile-controlled analysis produced feasible solutions. Across the coalitions, the largest coefficient of variation in cost was 0.226%, and the largest positive cost difference from the deterministic reference was 0.154%. PSO found a lower-cost solution than the deterministic reference for 12 of the 15 coalitions. The largest energy capacity obtained was 88.12% of the prescribed upper bound; doubling this bound improved the best cost by only 0.0063%.

At 100% design-day coverage, the shared capacity increases to 1108.39 MWh and 100.06 MW. The shared annualized cost is 34.91 million USD/year, compared with 81.15 million USD/year for separate batteries, a 56.98% reduction. Wind retains the largest Shapley share of the savings (33.27%). Therefore, the pooling benefits persist when all coalitions serve the same complete set of design days. This supplementary experiment used a smaller search budget than the main analysis.

ISBN: 978-605-80193-6-2

For the four-source coalition, the deterministic reference costs increase from 18.53 to 27.31 million USD/year as the energy-cost multiplier increases from 0.8 to 1.2; the corresponding power-cost range is 22.73–23.12 million USD/year. Component efficiencies of 0.90, 0.95, and 0.98 yield costs between 22.82 and 23.12 million USD/year for 95% coverage. These reference results avoid conflating the parameter effects with the variability in the shorter PSO searches. In the nine fixed-capacity allocation scenarios, wind retains the largest Shapley share of the savings, ranging from 32.84% to 33.43%, and total cost savings range from 54.79% to 55.13%. This result concerns the repricing of base-case designs. The sensitivity of the allocation to the cost ratio when every coalition is re-optimized remains to be evaluated.

In the fixed-scaling transfer analysis, the four-source BESS was sized at 698.07 MWh and 91.53 MW. Daily coverage was 95.03% in the design period and decreased to 91.80% in 2024 and 83.84% in 2025 when the same capacity was applied without re-optimization (Figure 4). Under profile-controlled scaling, coverage in the same years was 96.72% and 98.08%, respectively. The difference between the two results arises because annual rescaling removes changes in the generation magnitude.

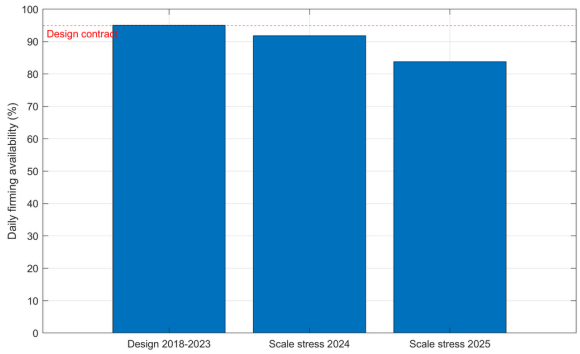


Fig. 4. Daily coverage of the four-source BESS in the fixed-scaling transfer analysis

The savings function satisfies superadditivity: the savings obtained by merging disjoint coalitions never fall below the sum of their separate savings. The convexity condition fails for ten coalition pairs in the main analysis. Accordingly, the Shapley values are treated as the sources' average marginal contributions to pooled savings without claiming core stability.

IV. DISCUSSION

Community storage required 65% of the aggregate individual-household storage capacity in the simulated case reported in [6]. Reference [5] addresses demand-load shifting, while [8] combines household energy management and storage sizing under generation and demand uncertainty. A direct numerical comparison is inappropriate because user profiles, tariffs, operating objectives, and cost definitions differ among studies. The 54.93% reduction reported here is specific to the profile-controlled case, in which the four sources are placed on an equal power scale, and each individual battery is sized to the same coverage target. The savings rate for portfolios with unequal generation scales remains to be evaluated.

The wind allocation reflects its standalone storage cost relative to the additional shared-storage cost when it joins a coalition. The annualized cost of the standalone wind BESS is 15.12 million USD/year, whereas adding wind to the solar-hydropower-biomass coalition raises the shared cost from 22.60 to 22.92 million USD/year. This transition therefore produces marginal savings of 14.80 million USD/year. This coalition transition illustrates the cost difference underlying the Shapley calculation; the reported share averages all 24 participation orders. The fixed-capacity sensitivity calculation preserves wind's leading share within the tested coefficient range; extending this conclusion to re-optimized designs requires a separate coalition analysis.

The decline in coverage to 83.84% in 2025 under fixed-scaling transfer indicates that capacity determined in the design period should be reassessed when generation scale changes. In the model, the constant power delivered to the grid is recalculated from each day's generation, and the battery remains idle on approximately 5% of the design days, during which the constant daily output target is not maintained. This coverage metric does not represent firm capacity available throughout the year or a real-time power guarantee. The findings provide a planning assessment of resource complementarity under the assumption that each day's generation profile is known in advance.

V. CONCLUSION

Shared BESS planning across all 15 renewable-source coalitions identifies a 54.93% annualized cost reduction relative to separate batteries in the profile-controlled case. Shapley values assign the largest share of this saving to wind, and every source receives a positive share. Superadditivity supports pooling between disjoint coalitions in the computed game. The failures of the convexity condition preclude invoking the convex-game guarantee of a core allocation; the reported values describe average marginal contributions.

Fixed-scaling transfer shows that a capacity selected in the design period may lose service coverage as the generation scale changes. The pooling benefit and capacity adequacy therefore require separate assessments. The results are conditional on advance knowledge of daily generation, the specified cost scenario, and the BESS model. Cost-ratio sensitivity has been established only for repriced base-case designs, and re-optimized coalition allocations remain an open test. Future work should examine alternative initializations, forecast errors, battery degradation, network constraints, and year-dependent capacity expansion.

References

[1] Jurasz, J., Canales, F. A., Kies, A., Guezgouz, M., & Beluco, A. (2020). *A review on the complementarity of renewable energy sources: Concept, metrics, application and future research directions*. *Solar Energy, 195*, 703-724. https://doi.org/10.1016/j.solener.2019.11.087

[2] Heide, D., von Bremen, L., Greiner, M., Hoffmann, C., Speckmann, M., & Bofinger, S. (2010). *Seasonal*

ISBN: 978-605-80193-6-2

optimal mix of wind and solar power in a future, highly renewable Europe. *Renewable Energy, 35*(11), 2483-2489. https://doi.org/10.1016/j.renene.2010.03.012

[3] Denholm, P., & Hand, M. (2011). *Grid flexibility and storage required to achieve very high penetration of variable renewable electricity*. *Energy Policy, 39*(3), 1817-1830. https://doi.org/10.1016/j.enpol.2011.01.019

[4] Beluco, A., de Souza, P. K., & Krenzinger, A. (2008). *A dimensionless index evaluating the time complementarity between solar and hydraulic energies*. *Renewable Energy, 33*(10), 2157-2165. https://doi.org/10.1016/j.renene.2008.01.019

[5] Parra, D., Norman, S. A., Walker, G. S., & Gillott, M. (2016). *Optimum community energy storage system for demand load shifting*. *Applied Energy, 174*, 130-143. https://doi.org/10.1016/j.apenergy.2016.04.082

[6] Barbour, E., Parra, D., Awwad, Z., & González, M. C. (2018). *Community energy storage: A smart choice for the smart grid? Applied Energy, 212*, 489-497. https://doi.org/10.1016/j.apenergy.2017.12.056

[7] Koirala, B. P., van Oost, E. C. J., & van der Windt, H. J. (2018). *Community energy storage: A responsible innovation towards a sustainable energy system? Applied Energy, 231*, 570-585. https://doi.org/10.1016/j.apenergy.2018.09.163

[8] Hafiz, F., de Queiroz, A. R., Fajri, P., & Husain, I. (2019). *Energy management and optimal storage sizing for a shared community: A multi-stage stochastic programming approach*. *Applied Energy, 236*, 42-54. https://doi.org/10.1016/j.apenergy.2018.11.080

[9] Roberts, M. B., Bruce, A., & MacGill, I. (2019). *Impact of shared battery energy storage systems on photovoltaic self-consumption and electricity bills in apartment buildings*. *Applied Energy, 245*, 78-95. https://doi.org/10.1016/j.apenergy.2019.04.001

[10] Li, L., Cao, X., & Zhang, S. (2022). *Shared energy storage system for prosumers in a community: Investment decision, economic operation, and benefits allocation under a cost-effective way*. *Journal of Energy Storage, 50*, 104710. https://doi.org/10.1016/j.est.2022.104710

[11] Shapley, L. S. (1953). A value for n-person games. In H. W. Kuhn & A. W. Tucker (Eds.), *Contributions to the Theory of Games II* (pp. 307-317). Princeton University Press. https://doi.org/10.1515/9781400881970-018

[12] Pilling, R., Chang, S. C., & Luh, P. B. (2017). *Shapley value-based payment calculation for energy exchange between micro- and utility grids*. *Games, 8*(4), 45. https://doi.org/10.3390/g8040045

[13] Dabbagh, S. R., & Sheikh-El-Eslami, M. K. (2015). *Risk-based profit allocation to DERs integrated with a virtual power plant using cooperative game theory*. *Electric Power Systems Research, 121*, 368-378. https://doi.org/10.1016/j.epsr.2014.11.025

[14] Safdarian, A., Astero, P., Baranauskas, M., Keski-Koukkari, A., & Kulmala, A. (2021). *Coalitional game theory based value sharing in energy communities*. *IEEE Access, 9*, 78266-78275. https://doi.org/10.1109/ACCESS.2021.3081871

[15] Cremers, S., Robu, V., Zhang, P., Andoni, M., Norbu, S., & Flynn, D. (2023). *Efficient methods for approximating the Shapley value for asset sharing in energy communities*. *Applied Energy, 331*, 120328. https://doi.org/10.1016/j.apenergy.2022.120328

[16] Kennedy, J., & Eberhart, R. (1995). Particle swarm optimization. In *Proceedings of ICNN'95: International Conference on Neural Networks* (Vol. 4, pp. 1942-1948). IEEE. https://doi.org/10.1109/ICNN.1995.488968

[17] Shi, Y., & Eberhart, R. C. (1998). Parameter selection in particle swarm optimization. In V. W. Porto, N. Saravanan, D. Waagen, & A. E. Eiben (Eds.), *Evolutionary Programming VII* (pp. 591-600). Springer. https://doi.org/10.1007/BFb0040810

[18] Kerdphol, T., Qudaih, Y., & Mitani, Y. (2015). *Optimal battery energy storage size using particle swarm optimization for microgrid system*. *International Review of Electrical Engineering, 10*(2), 277-285. https://doi.org/10.15866/iree.v10i2.5350

[19] Kerdphol, T., Fuji, K., Mitani, Y., Watanabe, M., & Qudaih, Y. (2016). *Optimization of a battery energy storage system using particle swarm optimization for stand-alone microgrids*. *International Journal of Electrical Power & Energy Systems, 81*, 32-39. https://doi.org/10.1016/j.ijepes.2016.02.006

[20] National Energy System Operator. (n.d.). *Historic generation mix and carbon intensity:* https://www.neso.energy/data-portal/historic-generation-mix

[21] Cole, W., & Karmakar, A. (2023). Cost projections for utility-scale battery storage: 2023 update (NREL/TP-6A40-85332). National Renewable Energy Laboratory. https://docs.nrel.gov/docs/fy23osti/85332.pdf

ISBN: 978-605-80193-6-2

Frequency Regulation Using a Cascade T2F-PIDDN Controller for LFC of a 2-Area Multi-Source Power System

Mehmet Karayel[1] Turan Cakil[1]

[1]Karamanoglu Mehmetbey University, Kamil Özdag Faculty of Science, Department of Software Development, Karaman, Turkey

* Corresponding author e-mail: mehmetkarayel@kmu.edu.tr

***Abstract*— Energy demand has been rising markedly over the past few decades, driven by rapid population growth and accelerating industrial and technological advancement. To address this challenge, power engineers have increasingly turned to hybrid generation strategies that combine conventional plants with renewable sources within large-scale interconnected power systems (PSs). In such systems, frequency stands out as a key indicator of operational quality and reliability. Abrupt shifts in load demand or unexpected changes in system parameters can trigger frequency and inter-area power flow deviations that, if left unaddressed, threaten system integrity. Load-frequency control (LFC) serves as the principal mechanism for correcting these deviations in a fast and stable manner. For an LFC scheme to perform reliably under real operating conditions, it must be sufficiently robust to handle both external disturbances and internal parameter uncertainties. Motivated by these requirements, the present work introduces a cascade Type-2 Fuzzy Logic–PIDDN (T2F-PIDDN) controller aimed at improving LFC performance in interconnected PSs. The controller gains are determined through the Wild Horse Optimization Algorithm (WHOA), with the Integral of Time-weighted Absolute Error (ITAE) adopted as the cost function. Extensive comparative simulations confirm that the proposed controller delivers superior regulation performance across all examined operating scenarios.**

***Keywords*— Wild horse optimization algorithm, load-frequency control, type-2 fuzzy logic control, T2F-PIDDN controller**

I. INTRODUCTION

Driven by a steadily expanding global population and the ongoing march of industrialization, both energy demand and the structural complexity of power systems continue to rise. At its core, a power system is expected to deliver electricity to consumers in a manner that is uninterrupted, secure, and stable. Sustaining the stability, security, and reliability of the system ultimately hinges on supplying continuous, high-quality electrical power. Within power systems, the steadiness of frequency serves as a key marker of overall power quality. Load-frequency control (LFC) is central to preserving the balance between generation and demand, holding frequency steady, and safeguarding power quality as a whole. The chief aim of LFC extends beyond merely managing generation frequency and load shifts in interconnected systems; it also keeps the system frequency and tie-line power anchored at their nominal values, whether the system is operating normally or facing substantial load disturbances [1–4].

Alongside its role in damping load disturbances, an effective LFC controller must remain robust in the presence of uncertain system parameters and the inherently random character of load fluctuations. For this reason, achieving the targeted performance in today's intricate power systems calls for LFC controllers that are at once robust, resilient, and quick to respond. Over the past several years, a great deal of research attention has focused on devising LFC strategies capable of improving the dynamic and regulatory behavior of interconnected power systems. A survey of the literature reveals that the proportional-integral-derivative (PID) controller remains a widely favored structure in LFC applications. That said, the resulting control performance depends heavily not only on the chosen controller architecture but equally on how carefully its tuning parameters are selected [5], [6]. To tune these parameters in a systematic fashion, metaheuristic optimization techniques have gained broad acceptance within LFC research. The authors of [7], for example, tuned a PID controller via the Firefly Algorithm (FA) to regulate load frequency in combined-cycle gas turbine units.

In [8], quasi-oppositional Grey Wolf Optimization (QGWO) was applied to secure the optimal PID gains for LFC in a large-scale power system. Along similar lines, the study in [9] relied on Differential Evolution (DE) to fix suitable controller settings for automatic generation control (AGC) within an interconnected grid. A number of works have instead sought to reshape the classical PID architecture itself, aiming to overcome its built-in limitations and lift its LFC performance.

One such effort [10] merged fuzzy logic with a conventional PID to form an ICA-based fuzzy PID controller, whose effectiveness was shown on a two-area reheat thermal–PV system. Elsewhere, the work in [5] put forward a multistage PID structure driven by a hybrid of Stochastic Fractal Search (SFS) and Local Unimodal Sampling (LUS) to achieve better regulation. Against this backdrop, the present study develops a coordinated control-design framework that tackles the LFC problem in a power system drawing on both conventional and renewable generation.

Within this framework, a T2F-PIDDN controller is designed for a two-area interconnected system made up of reheat thermal, hydroelectric, and gas-turbine units. Its parameters are then optimized in a systematic way through the Wild Horse Optimization Algorithm (WHOA), yielding well-suited control settings and strengthening overall regulation.

To evaluate performance, a two-area multi-source LFC model is constructed and run within the MATLAB/Simulink environment. In addition, the WHOA is coded as a MATLAB M-file and linked with

ISBN: 978-605-80193-6-2

this power-system model, making it possible to pinpoint the optimal parameters of the proposed T2F-PIDDN controller

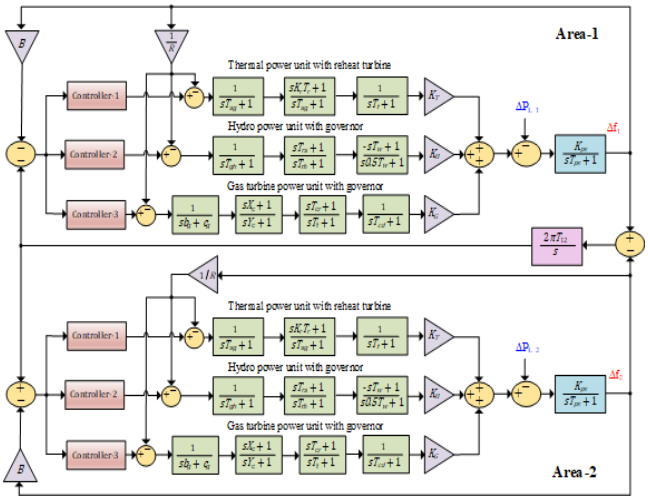


Fig. 1. Transfer-function model of the 2-area multi-source PS.

In Fig. 1, R denotes the governor regulation coefficient, B denotes the frequency bias coefficient, ΔP_{L_1} and ΔP_{L_2} represent the load variations, and Δf_1 and Δf_2 denote the area frequency deviations. The system parameters used in the study are: $T_{12} = 0.433$ pu MW/rad, $K_{ps} = 68.9566$ Hz/pu, $T_t = 0.3$ s, $B = 0.4312$ pu MW/Hz, $T_{gh} = 0.2$ s, $R = 2.4$ Hz/pu, $f_0 = 60$ Hz, $T_r = 10$ s, $T_{rs} = 5$ s, $T_{rh} = 28.75$ s, $T_{cr} = 0.01$ s, $T_{cd} = 0.5$ s, $T_f = 0.23$ s, $T_w = 1$ s, $b_g = 0.05$ s, $c_g = 1$, $X_c = 0.6$ s, $Y_c = 1$ s, $K_r = 0.3$ s, $K_T = 0.543478$ pu, $K_H = 0.326084$ pu, and $K_G = 0.130438$ pu, $T_{ps} = 11.49$ s, $T_{sg} = 0.08$ s.

II. CASCADE T2F-PIDDN CONTROLLER

Here, an augmented control structure built upon the Type-2 Fuzzy Logic–PID (T2F-PID) concept is introduced to capture the system's dynamic response while coping with the uncertainties that arise under varying operating conditions. The architecture is organized around two nested closed loops: the inner loop handles fine-grained performance regulation, whereas the outer loop refines the system's broader dynamic response and strengthens its reference-tracking capability. Positioned in the outer loop, the Type-2 Fuzzy Logic Controller (T2FLC) is tasked with suppressing the performance fluctuations that stem from parameter uncertainties, sensor noise, and external disturbances acting on the system model. Fuzzy-logic-based control approaches are widely used for nonlinear and uncertain dynamic systems whose mathematical modelling is difficult, since they allow expert control knowledge to be represented through a rule-based framework. In Type-1 fuzzy logic systems, however, membership functions are fixed by crisp boundaries, which restricts how well these systems can express uncertainty and often leads to degraded controller behavior once operating conditions shift. Type-2 Fuzzy Logic Systems (T2FLSs) address this shortcoming by introducing an extra dimension into the membership grades, which allows a wider spectrum of uncertainty to be folded into the control process. Thanks to this added flexibility, T2F control structures emerge as a dependable choice for high-precision applications that must contend with parameter drift, sensor measurement errors, shifting operating points, and external disturbances. The overall structure of the proposed T2FLC is presented in Fig. 2 and consists sequentially of fuzzification of the input variables, operation of the fuzzy inference mechanism and rule base, type-reduction of the resulting Type-2 fuzzy set, and defuzzification to generate the final crisp control signal.

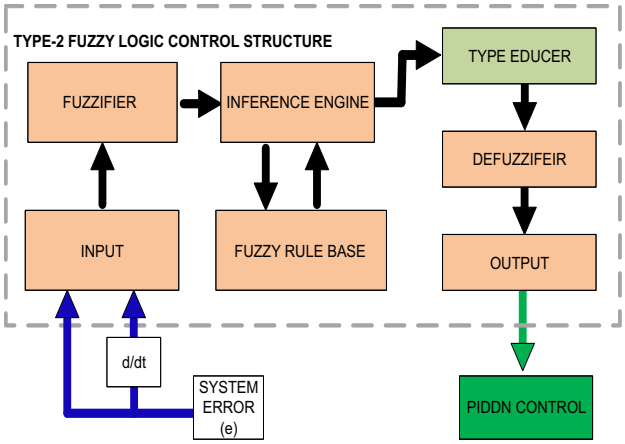


Fig. 2. Structure of the Type-2 fuzzy logic controller.

- Fuzzification: The system inputs are converted into fuzzy sets. In the Type-2 structure, the membership grades are represented by interval values.
- Rule Base: The controller uses rules in the form of IF–THEN statements derived from expert knowledge.
- Inference Mechanism: The firing strengths of the rules are calculated, and a fuzzy output is obtained.
- Type-Reduction: This is one of the main distinctions between Type-2 and Type-1 systems. The output fuzzy set is transformed into a Type-1 set. The Karnik–Mendel algorithm is among the most used methods.
- Defuzzification: The resulting Type-1 output is converted into a crisp numerical value.

The input membership functions of the Type-2 fuzzy logic controller are shown in Fig. 3. Triangular and trapezoidal membership functions are used to construct the input membership functions.

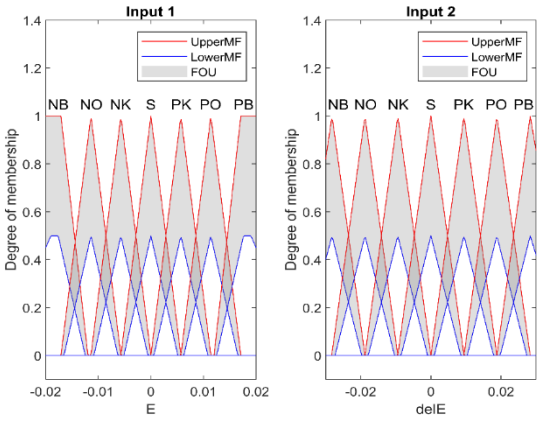


Fig. 3. Input functions of the Type-2 FLC.

PID controllers are among the most widely used control methods in industrial control systems. However, conventional PID structures may not always provide

ISBN: 978-605-80193-6-2

sufficient performance for systems characterized by complexity, nonlinearity, and uncertainty. Therefore, a T2F control structure is incorporated to enhance the dynamic performance of the PID controller. The configuration of the T2F-PIDDN controller is shown in Fig. 4.

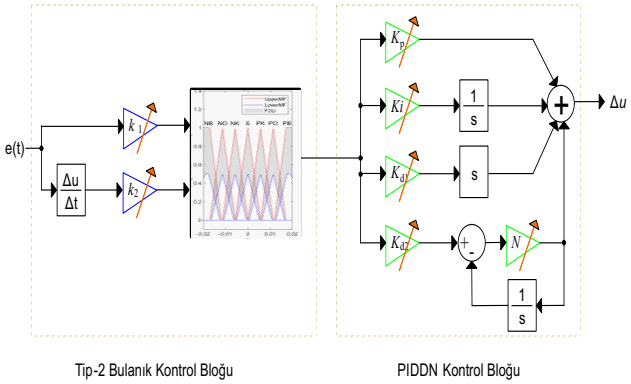


Fig. 4. T2F-PIDDN controller configuration.

III. WILD HORSE OPTIMIZATION ALGORITHM

WHOA was introduced by Naruei and Keynia in 2021 to solve challenging optimization problems with fast convergence and high accuracy. To achieve this, WHOA mathematically simulates the social behaviors of wild horses in nature and incorporates their mating, grazing, and leadership behaviors [13]. The main steps of WHOA are presented below to summarize its operation and facilitate understanding.

At the beginning of the optimization process, the initial population is partitioned into a predefined number of groups. In this formulation, N represents the total population size, PS denotes the number of groups, and N-G indicates the proportion of stallions within the overall population, which is treated as an algorithmic control parameter. Each group is assigned a leader, referred to as a stallion, while the remaining individuals consist of foals and mares. Based on the number of groups G, the stallions and the remaining N-G individuals are distributed as evenly as possible among the groups. At the initial stage, the group leaders are selected randomly. During the subsequent iterations, however, the stallion representing each group is determined according to the fitness values of its members.

$$G = N \times PS \tag{1}$$

The second stage of the algorithm corresponds to the grazing behaviour. In this phase, the stallion is defined as the reference center of the grazing region, and the members of the corresponding group perform a search in its surrounding neighbourhood. The mathematical representation of this grazing mechanism is presented in Eq. (2).

$$\bar{X}_{i,G}^{j} = 2Z\cos(2\pi RZ) \times (Stallion^{j} - X_{i,G}^{j}) + Stallion^{j} \tag{2}$$

Where, $\bar{X}_{i,G}^{j}$ denotes the new position of the foal and mare group members, $Stallion^{j}$ denotes the stallion position, R denotes a random number in the interval [-2, 2], and Z represents the adaptive mechanism calculated as given in Eq. (3).

$$P = \overrightarrow{R_1} < TDR;\quad IDX = (P == 0);\quad Z = R_2\theta IDX + \overrightarrow{R_3}\,\theta(\sim IDX) \tag{1}$$

Where, P represents a binary vector composed of zeros and ones, with its dimension corresponding to the number of decision variables in the optimization problem. $\overrightarrow{R_1}$ and $\overrightarrow{R_3}$ are random vectors generated from a uniform distribution over the interval [0,1], whereas R_2 denotes a uniformly distributed random scalar. The parameter TDR is an adaptive control factor initialized at 1 and progressively reduced according to Eq. (4) throughout the optimization process, ultimately approaching 0 at the final iteration.

$$TDR = 1 - iter \times \left(\frac{1}{iter_{mak}}\right) \tag{4}$$

where $iter$ denotes the current iteration number and $iter_{mak}$ represents the maximum number of iterations. One of the characteristic social behaviors of wild horses is the separation of foals from their original groups, followed by their integration into new groups and subsequent mating interactions. Prior to reaching maturity, male foals typically leave their family groups and join bachelor groups, whereas female foals move to other family groups, where they remain until maturity and establish mating relationships. In the mathematical model of WHOA, one foal is selected from the i-th group and transferred to a temporary group, while another foal from the j-th group is assigned to the same temporary group. These two foals are assumed to be a male and a female with no familial relationship, allowing mating to occur after they reach maturity. The offspring produced through this process subsequently leaves the temporary group and is transferred to another group, such as the k-th group. This cycle of group separation, mating, and offspring redistribution is repeatedly applied across the different horse groups. The mathematical formulation representing the separation and mating behavior of the horses is given in Eq. (5).

$$X_{G,K}^{P} = Crossover(X_{G,i}^{q}, X_{G,j}^{z}) \tag{5}$$

Another stage of WHOA is the group-leadership stage. Here, stallions (group leaders) lead the group toward a destination (e.g., a water source). The stallions compete for access to the water source so that the dominant group can use it first, followed by the other groups. The mathematical formulation of stallion leadership behavior is given in Eq. (6).

$$\overline{Stallion_{Gi}} = \begin{cases} 2Z\cos(2\pi RZ) \times (WH - Stallion_{Gi}) + WH & if\ R_3 > 0.5 \\ 2Z\cos(2\pi RZ) \times (WH - Stallion_{Gi}) + WH & if\ R_3 \le 0.5 \end{cases} \tag{6}$$

Where, $\overline{Stallion_{Gi}}$ denotes the next position of the leader, and WH denotes the location of the water source. The final stage is the replacement and selection of group leaders. The selection of the group leader is given in Eq. (7).

ISBN: 978-605-80193-6-2

$$\overline{Stallion_{Gi}} = \begin{cases} X_{G,i} \ \ if\ \beta(X_{G,i}) < \beta(Stallion_{Gi}) \\ Stallion_{Gi} \ \ if\ \beta(X_{G,i}) > \beta(Stallion_{Gi}) \end{cases} \quad (7)$$

Where, β denotes the objective function.

To obtain a better response for frequency regulation in an LFC system subject to load variations, the WHOA is used to search for the optimal values of the K_1, K_2, K_p, K_i, K_{d1}, K_{d2} and K_n parameters of the T2F-PIDDN controller, thereby improving the control performance of the considered system. Where t represents the time variable, T represents the total simulation period, ΔF_1 and ΔF_2 correspond to the frequency deviations in Areas 1 and 2, respectively, and ΔP_{tie} represents the tie-line power deviation. The optimization objective is to determine a controller configuration that minimizes the ITAE criterion while ensuring that the frequency and tie-line power deviations converge to 0 pu within the shortest attainable time. Accordingly, the design of the T2F-PIDDN controller is formulated as a constrained optimization problem, in which the ITAE objective is minimized subject to the inequality constraints specified in Eq. (8).

$$K_1^{min} \le K_1^* \le K_1^{max},\ K_2^{min} \le K_2^* \le K_2^{max},\ K_p^{min} \le K_p^* \le K_p^{max} \quad (8)$$

$$K_{d1}^{min} \le K_{d1}^* \le K_{d1}^{max},\ K_{d2}^{min} \le K_{d2}^* \le K_{d2}^{max}, K_i^{min} \le K_i^* \le K_i^{max}$$
$$K_n^{min} \le K_n^* \le K_n^{max}$$

In Eq. (8), the min and max superscripts denote the lower and upper bounds of the controller parameters, respectively. In this study, the lower and upper parameter bounds are set to 0 and 5, respectively. Following the optimization process, the values of K_1, K_2, K_p, K_i, K_{d1}, K_{d2} and K_n that yield the minimum ITAE objective-function value are selected as the optimal controller parameters and used in the simulation studies

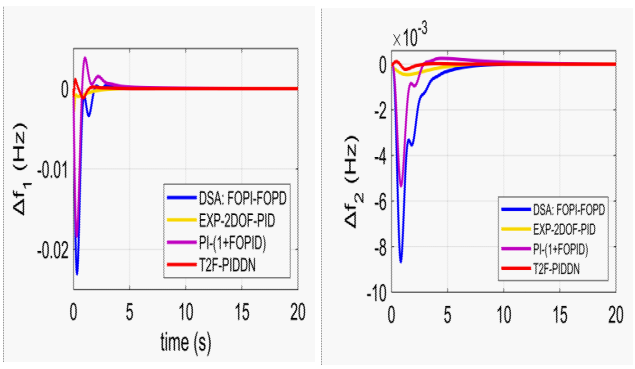
IV. SIMULATION RESULTS

Because the 2-area multi-source PS contains generation units with different structures, the controller parameters must be evaluated independently. Accordingly, the optimal values of a total of 21 parameters for the proposed T2F-PIDDN controller are determined for the two areas using the WHOA under identical conditions. Table 1 shows the numerical results of the transient performance response analysis of the proposed T2F-PIDDN controller and other compared approaches according to t_S, U_S, O_s and ITAE criteria in terms of frequencies and tie-line power. All results reported in Table 1 are obtained by applying a 0.02 pu load change to Area 1 at t_s = 0 s.

Table 1. Controller parameters and ITAE value.

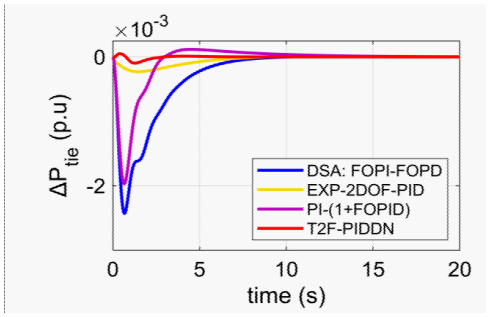
Cont.	$\|t_S\|$			$\|U_S\|$ Hz		$\|U_S\|$ pu MW	$\|O_S\|$ (Hz)		$\|O_S\|$ pu MW	
	Δf_1	Δf_2	ΔP_{tie}	Δf_1	Δf_2	ΔP_{tie}	Δf_1	Δf_2	ΔP_{tie}	J_{ITAE}
FOPI-FOPD [12]	4.3	4.5	3.9	0.023	0.0086	2.43e-3	4.7e-4	0	**0**	0.046
EXP-2DOF-PID [13]	2.1	3.4	1.9	0.002	**4.5e-4**	2.3e-4	**0**	3.7e-5	**0**	0.022
PI-(1+FOPID) [14]	3.4	2.2	1.9	0.017	0.005	1.71e-3	2.7e-3	**0**	0.6e-5	0.041
T2F-PIDDN	**1.1**	**0**	**0**	**0.001**	2.2e-4	**9.5e-5**	1.14e-3	1.1e-4	4.8e-5	**2.9e-3**

The system performance is evaluated in MATLAB/Simulink using the controller parameters obtained by the WHOA for the proposed study. The simulation responses of all controllers are obtained under identical conditions over the interval [0, 20] s for a 0.02 pu load change and are presented in Fig. 5. As observed in Fig. 5, the response of the WHOA-optimized T2F-PIDDN controller is faster, exhibits no significant oscillations, and provides more effective regulation than the other controllers.



(a) (b)

ISBN: 978-605-80193-6-2



(c)

Fig. 5. Comparative results for the 2-area multi-source PS.

An interconnected power system emerges when various load clusters and generation plants are tied together through transmission lines. As a result, the system experiences ongoing fluctuations in load throughout its operation. Faced with load changes of varying magnitude, the LFC controller is expected to react both swiftly and reliably, driving the power and frequency back to their nominal levels in the shortest possible time. To gauge the stability and robustness of the WHOA-tuned T2F-PIDDN controller, the system is accordingly subjected to randomly varying load profiles. For this purpose, the controller parameters given in Table 1 are kept fixed for the 2-Area PS and are not returned in response to changes in system conditions. During the simulation, load variations are applied to all areas as shown in Fig. 6. As illustrated, the load variations exhibit complex characteristics with continuous and instantaneous changes.

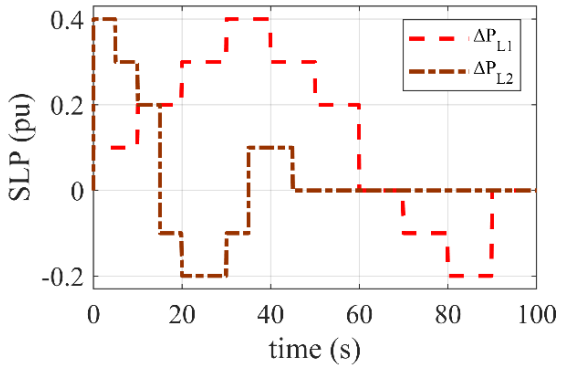
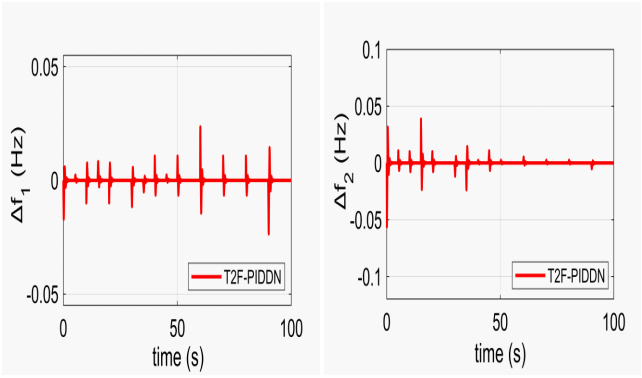
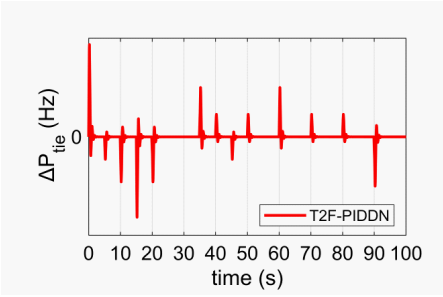


Fig. 6. Load variations applied to PS.

The response curves of the proposed controller obtained under the load variations applied at different times, as shown in Fig. 6, are presented in Fig. 7. The results indicate that the controller rapidly restores the frequency deviations and tie-line power deviations occurring in both areas to their reference values with low oscillation amplitudes, despite the connection and disconnection of loads with different magnitudes.



(a) (b)



(c)

Fig. 7. Respectively frequency and tie-line power deviations responses at change load.

V. CONCLUSIONS

In this study, a cascade T2F-PIDDN controller incorporating a filtered proportional-derivative structure is proposed for LFC in a 2-area interconnected multi-source PS comprising hydroelectric, thermal, and gas-turbine generating units. Unlike conventional trial-and-error parameter tuning, the controller gains are systematically optimized using WHOA. The effectiveness of the proposed control strategy is assessed under various load disturbance scenarios. The simulation results demonstrate that the proposed controller achieves satisfactory dynamic performance in terms of settling time, positive and negative maximum overshoot, oscillation attenuation, and ITAE for the operating conditions considered.

References

[1] R. Choudhary, J. N. Rai and Y. Arya, "Cascade FOPI-FOPTID controller with energy storage devices for AGC performance advancement of electric power systems", Sustain. Energy Technol. Assessments, vol. 53, 2022.

[2] S. Das, S. Datta, L. C. Saikia and S. K. Bhagat, "Effect of electric vehicles and TIDN-(1+PI) controller on LFC in hydro-thermal-archimedes wave energy-geothermal-wind generations based multiarea system", Proc. 4th Int. Conf. Energy Power Environ. (ICEPE), pp. 1-6, 2022.

[3] S. Oladipo, Y. Sun and Z. Wang, "Application of a new fusion of flower pollinated with pathfinder algorithm for AGC of multi-source interconnected power system", IEEE Access, vol. 9, pp. 94149-94168, 2021.

[4] 4. X. He, X. Liu and P. Li, "Coordinated false data injection attacks in AGC system and its

ISBN: 978-605-80193-6-2

countermeasure", IEEE Access, vol. 8, pp. 194640-194651, 2020.

[5] R. Sivalingam, S. Chinnamuthu and S. S. Dash, "A hybrid stochastic fractal search and local unimodal sampling based multistage PDF plus (1+PI) controller for automatic generation control of power systems", J. Franklin Inst., vol. 354, no. 12, pp. 4762-4783, 2017.

[6] M. Karayel, E. Çelik and İ. Yücedağ, "Fuzzy Logic-Based Single-Area Load-Frequency Control," III. International Scientific and Vocational Studies Congress-Engineering (BILMES EN 2019), 27-30 June, Nevşehir/Türkiye, (2019).

[7] L. C. Saikia and S. K. Sahu, "Automatic generation control of a combined cycle gas turbine plant with classical controllers using firefly algorithm", Int. J. Electr. Power Energy Syst., vol. 53, pp. 27-33, 2013.

[8] D. Guha, P. K. Roy and S. Banerjee, "Load frequency control of large-scale power system using quasi-oppositional grey wolf optimization algorithm", Eng. Sci. Technol. Int. J., vol. 19, no. 4, pp. 1693-1713, 2016.

[9] P. K. Hota and B. Mohanty, "Automatic generation control of multi-source power generation under deregulated environment", Int. J. Electr. Power Energy Syst., vol. 75, pp. 205-214, 2016.

[10] Y. Arya, "Automatic generation control of two-area electrical power systems via optimal fuzzy classical controller", J. Franklin Inst., vol. 355, no. 5, pp. 2662-2688, 2018.

[11] Naruei Iraj, Keynia Farshid, "Wild horse optimizer: a new meta-heuristic algorithm for solving engineering optimization problems", Eng. Comput. pp. 1-32, (2021).

[12] E. Çelik, "Design of new fractional order PI–fractional order PD cascade controller through dragonfly search algorithm for advanced load frequency control of power systems" Soft Computing, 25(2), 1193-1217, (2021).

[13] M. Karayel, "Design of Exponantial-2-DOF-PID Controller Using Stochastic Fractal Search Algorithm for Interconnected Electric Power Systems," International Journal of Smart Grid - ijSmartGrid, vol. 9, no. 2, pp. 71-82, 2025.

[14] J. Hussain, R. Zou, P. K. Pathak, A. Karni, and S. Akhtar, "Design of a novel cascade PI-(1+ FOPID) controller to enhance load frequency control performance in diverse electric power systems.", Electric Power Systems Research, vol. 243, 111488, 2025.

 ISBN: 978-605-80193-6-2

When Does Feature-Selection Stability Matter? Evidence Across Low- and High-Dimensional Classification Regimes

Ramazan EŞMELİ*
Van Yuzuncu Yil University, Van, Türkiye
* Corresponding author e-mail: ramazanesmeli@yyu.edu.tr

***Abstract*— Feature-selection stability is often treated as evidence that a selected subset is reliable, although reproducibility need not imply downstream predictive utility or accurate signal recovery. We evaluate this distinction across 25 classification tasks (eight real and seventeen synthetic), including eleven p > n problems, using ten stratified train/test resamples. Five selectors were compared: ANOVA F-score, mutual information, random-forest importance, L1 logistic selection, and a subsampled L1 stability-selection baseline under a primary 25% feature budget, with 10% and 50% sensitivity analyses in the high-dimensional regime. At 25%, stability rankings differed by dimensionality: FScore was most stable in p <= n tasks (mean corrected stability 0.869), whereas L1 was most stable in p > n tasks (0.860) and had the highest overall mean across 24 nondegenerate tasks (0.782). Predictive rankings did not mirror these stability rankings; all features had the highest mean LightGBM macro-F1 (0.729), whereas L1 had the highest logistic-regression macro-F1 (0.677). The overall Friedman test showed method heterogeneity (chi-square = 17.653, p = 0.0034), but corrected stability was not associated with predictive gain globally (Spearman rho = 0.050, p = 0.590). In p > n tasks, L1 yielded higher mean macro-F1 than all features for LightGBM (0.679 versus 0.660) and logistic regression (0.608 versus 0.580), while recovering all known signal variables at 25%. These findings support regime-aware evaluation of stability, prediction, signal recovery, computational cost, and feature budget as distinct criteria.**

***Keywords*— Feature selection, stability selection, chance-corrected stability, high-dimensional classification, LightGBM**

I. INTRODUCTION

Feature selection reduces the number of variables presented to a predictive model and can therefore improve computational efficiency, simplify interpretation, and reduce acquisition cost when variables correspond to sensors, laboratory measurements, or engineered descriptors [1]-[5]. The same reduction step can also introduce a reproducibility problem because small changes in the training sample may alter the selected subset even when downstream predictive performance remains similar, which makes stability relevant whenever the identity of the retained variables has scientific or engineering meaning [6], [7].

The central difficulty is that stability is not equivalent to predictive quality. A method can repeatedly select the same variables while excluding features that are useful to the downstream learner, whereas another method can alternate among correlated or redundant variables and still preserve prediction; this distinction is particularly important in high-dimensional settings, where many variables can act as substitutes and where sparse selection is often used as part of the modelling pipeline [6], [8], [9]. Raw overlap measures such as pairwise Jaccard similarity further complicate cross-dataset comparisons because chance agreement changes with dimensionality and selected-subset size, motivating chance-corrected estimators such as the stability measure proposed by Nogueira et al. [7].

This study provides a multi-regime empirical audit of whether feature-selection stability can serve as a proxy for downstream predictive utility or recovery of known signal. The benchmark contains 25 classification tasks, uses ten outer resamples, includes eleven p > n scenarios with up to 2,000 features, adopts chance-corrected stability as the primary reproducibility measure, and evaluates 10%, 25%, and 50% feature budgets in the high-dimensional regime. The contribution is not a new selector but a regime-aware benchmark that tests whether stability tracks predictive gain, whether selector stability rankings change when p exceeds n, and how feature budget alters reproducibility, signal recovery, and downstream performance.

II. RELATED WORK

Feature-selection methods are commonly organised as filters, wrappers, embedded procedures, and ensemble approaches [1]-[5], [10]. Univariate F statistics and mutual-information scores provide computationally inexpensive rankings [11], while random forests supply nonlinear embedded importance estimates [12], [13]; sparse linear models provide an alternative embedded mechanism in which zero coefficients define an explicit variable screen, and information-theoretic approaches such as minimum-redundancy maximum-relevance extend ranking beyond marginal relevance [14]-[16].

Stability research asks a different question by measuring whether a selection procedure returns comparable subsets after perturbations of the training data. Kalousis et al. examined this problem in high-dimensional spaces [6], Nogueira et al. developed a chance-corrected estimator with desirable statistical properties [7], and Meinshausen and Bühlmann introduced stability selection by combining sparse models with repeated subsampling and selection frequencies [8]. Ensemble feature selection has also been studied as a route to robust rankings [17], although improved reproducibility does not imply that the resulting variables will optimize a particular downstream learner.

Benchmark studies have shown that no single filter or selection mechanism dominates across heterogeneous datasets [9], while statistical comparison across multiple datasets requires attention to repeated rankings and

ISBN: 978-605-80193-6-2

multiplicity rather than reliance on one aggregate mean [18]. The present design connects these strands by comparing chance-corrected subset stability, predictive gain relative to all features, and known-signal recovery under a common feature budget, with separate analyses for low- or moderate-dimensional data and p > n problems.

III. MATERIALS AND METHODS

A. Datasets and Evaluation Design

The benchmark contains 25 classification tasks. Eight real datasets were obtained from scikit-learn and statsmodels: Breast Cancer Wisconsin, Wine, Digits, Iris, ANES96, a binary version of the Fair Affairs dataset, Spector, and ModeChoice [19], [20]. Seventeen synthetic tasks were generated with scikit-learn to control dimensionality, redundancy, class imbalance, label noise, and the identity of informative variables. Across the complete benchmark, sample size ranges from 32 to 6,366 and feature dimensionality ranges from 3 to 2,000; eleven synthetic tasks satisfy p > n, providing a dedicated high-dimensional regime rather than extrapolating conclusions from low-dimensional datasets.

TABLE I. BENCHMARK REGIMES AND EVALUATION BUDGETS

Regime	Tasks	n; p range	Budget
Real	8	32-6366; 3-61	25%
Synthetic p <= n	6	200-1000; 20-100	25%
Synthetic p > n	11	100-800; 500-2000	10/25/50%
Total	25	32-6366; 3-2000	25% primary

Each task was evaluated with ten stratified shuffle splits using a 75% training and 25% test partition, and every selection operation was fitted only on the training partition. For budget b and dimensionality p, the retained count was k = max(3, round(bp)), so the primary budget retained approximately 25% of the available variables with a minimum of three features. The eleven p > n tasks were additionally evaluated at 10% and 50% to assess sensitivity to the retained fraction. Synthetic variables were generated with shuffle disabled so that informative and redundant signal variables remained identifiable, allowing direct precision and recall calculations for selected subsets.

B. Feature Selectors

Five selection procedures were evaluated under the same top-k budget. FScore ranked variables with the univariate ANOVA F statistic, MI used mutual information with the class label, and RF used impurity importance from a balanced random forest containing 80 trees, square-root feature subsampling, and a minimum leaf size of two. L1 selection standardised the training data and fitted L1-regularised logistic regression over a fixed C grid of 0.005, 0.01, 0.03, 0.1, 0.3, 1, and 3, selecting the regularisation level whose number of non-zero coefficients was closest to the target budget and ranking variables by absolute coefficient magnitude.

StabilityL1 implements the central resampling mechanism of stability selection [8] while preserving the fixed-budget comparison required by this benchmark. For each outer training partition, the selected L1 regularisation level was applied to 25 stratified half-samples, variable selection frequencies were estimated from non-zero coefficients, and features were ranked by frequency, with full-data coefficient magnitude used only to break ties. Because the final output is forced to a common top-k budget, the procedure is used as a stability-selection baseline rather than as a claim of the original probability-of-false-selection error bound.

C. Stability Measures

Subset reproducibility was measured across the ten outer resamples. The primary measure was the chance-corrected estimator of Nogueira et al. [7], which adjusts observed feature-wise selection variance by the disagreement expected under the average selected fraction and therefore supports comparisons across tasks with different dimensionalities and feature budgets. If M denotes the number of resamples, p the number of available variables, q the mean selected fraction, and the feature-specific sigma squared term the unbiased selection variance, the estimator is written as follows.

$$\Phi = 1 - \frac{p^{-1} \Sigma \sigma_j{}^2}{q(1-q)} \quad (1)$$

The feature-wise variance term is calculated from the empirical selection probability pi-hat for each variable, which yields the following finite-sample correction across M resamples.

$$\sigma_j{}^2 = \frac{M}{M-1} \pi_j(1 - \pi_j) \quad (2)$$

Pairwise Jaccard similarity was retained as a secondary descriptive metric so that the corrected estimator could be compared with the overlap measure frequently used in earlier feature-selection studies; Jaccard was not used as the primary cross-dataset stability statistic because its expected value changes with the feature budget. The Spector task has p = 3, and the minimum retained subset is also k = 3; therefore q = 1 and the denominator of the chance-corrected estimator is zero. Spector was retained in all predictive comparisons but excluded from corrected-stability summaries and stability-gain correlations.

D. Downstream Models, Recovery Metrics, and Statistical Analysis

Every selected subset was evaluated with class-balanced logistic regression and LightGBM, while an all-features condition provided the predictive baseline. Logistic regression used standardised inputs and the L-BFGS solver, whereas LightGBM used 80 trees, 31 leaves, a learning rate of 0.05, one CPU thread, and fixed seeds [21]; macro-F1 was the primary predictive metric because the benchmark includes balanced, imbalanced, binary, and multiclass tasks. Synthetic problems were additionally scored by precision and recall against the known informative-plus-redundant signal set, and selection time was recorded to quantify the computational cost of resampling-based selection.

Method differences were first tested with a Friedman omnibus test on dataset-level macro-F1 values averaged across the two downstream learners, with Kendall's W reported as the omnibus effect size. Classifier-specific Friedman tests were followed by pairwise Wilcoxon signed-rank comparisons with Holm correction [18]. Spearman correlation quantified the association between stability and predictive gain relative to all features

ISBN: 978-605-80193-6-2

across dataset-selector cells, and the association was recomputed separately for p <= n and p > n regimes because a global correlation can conceal dimensionality-specific patterns.

IV. RESULTS

A. Overall Stability and Predictive Performance

At the primary 25% budget, L1 had the highest overall mean chance-corrected stability at 0.782 across the 24 nondegenerate tasks, followed by FScore at 0.663. The stability ordering changed markedly across dimensional regimes. Among p <= n tasks with defined corrected stability, FScore was highest at 0.869, followed by RF at 0.794 and L1 at 0.717; within the eleven p > n tasks, L1 was highest at 0.860, followed by FScore at 0.419, MI at 0.307, StabilityL1 at 0.263, and RF at 0.099. Predictive rankings did not mirror either stability ordering. All features was numerically highest for LightGBM (0.729), with RF also rounding to 0.729 and L1 reaching 0.728, whereas L1 was highest for logistic regression (0.677 versus 0.665 for all features).

TABLE II. OVERALL PRIMARY-BUDGET PERFORMANCE AND STABILITY

Method	Stability	LGBM F1	LR F1
All features	-	0.729	0.665
FScore	0.663	0.721	0.661
MI	0.478	0.712	0.658
RF	0.476	0.729	0.665
L1	0.782	0.728	0.677
StabilityL1	0.498	0.710	0.660

The overall Friedman test detected method-level heterogeneity across the 25 tasks (chi-square = 17.653, p = 0.0034; Kendall's W = 0.141). Classifier-specific tests were also significant for LightGBM (chi-square = 23.205, p = 0.00031; W = 0.186) and logistic regression (chi-square = 11.951, p = 0.035; W = 0.096), although the corresponding rank-effect sizes were small. After Holm correction, the learner-averaged L1 versus StabilityL1 comparison remained significant (adjusted p = 0.028). For LightGBM, all features versus MI (adjusted p = 0.0043) and L1 versus StabilityL1 (adjusted p = 0.023) remained significant, whereas no logistic-regression pair survived multiplicity correction. The results therefore support heterogeneous method behaviour without establishing a universal selector ranking across datasets and learners.

B. Stability Is Not a Global Proxy for Predictive Gain

Across the 120 valid dataset-selector cells at the primary budget, chance-corrected stability was essentially unrelated to mean predictive gain over the all-features baseline (Spearman rho = 0.050, p = 0.590), and raw Jaccard led to the same substantive conclusion (rho = 0.037, p = 0.685). Stratification by dimensionality did not produce a significant association: p <= n tasks had rho = 0.023 (p = 0.856), while the eleven p > n tasks showed a positive but non-significant trend (rho = 0.231, p = 0.090). The expanded benchmark therefore does not support using subset stability as a statistically reliable proxy for predictive gain at the primary feature budget.

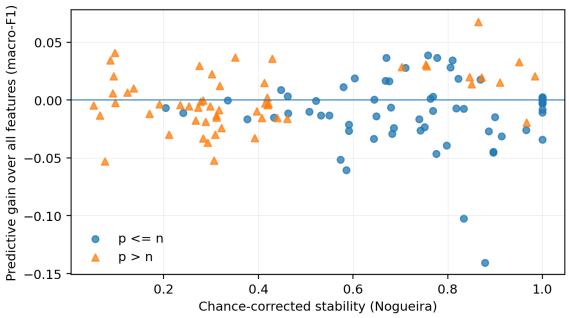


Fig. 1. Chance-corrected stability versus predictive gain across valid dataset-selector cells at the 25% budget. The global association is near zero; the p > n regime shows a positive but non-significant trend.

C. High-Dimensional Regime and Signal Recovery

The p > n regime showed the clearest descriptive advantage for feature selection. At the 25% budget, L1 yielded mean macro-F1 of 0.679 with LightGBM compared with 0.660 for all features, and 0.608 with logistic regression compared with 0.580 for all features. RF was only slightly above the all-features mean in this regime (0.662 for LightGBM and 0.585 for logistic regression). L1 also achieved the strongest recovery of known signal, with mean precision 0.167 and recall 1.000, indicating that the quarter-feature budget retained the complete informative-plus-redundant signal set while also selecting non-signal variables because the true signal set was sparse relative to p.

TABLE III. HIGH-DIMENSIONAL RESULTS AT THE 25% FEATURE BUDGET

Method	LGBM F1	LR F1	Signal R
All features	0.660	0.580	-
FScore	0.655	0.577	0.620
MI	0.643	0.568	0.408
RF	0.662	0.585	0.573
L1	0.679	0.608	1.000
StabilityL1	0.652	0.580	0.434

The comparison also separates resampling-based stability from fixed-budget predictive utility. StabilityL1 was the slowest selector at the primary budget, averaging 0.460 s per selection, compared with 0.306 s for MI, 0.182 s for RF, 0.013 s for L1, and less than 0.001 s for FScore. Despite this additional cost, StabilityL1 had neither the highest corrected stability nor the highest downstream macro-F1, while ordinary L1 had higher overall corrected stability and higher mean macro-F1 in the p > n regime. This comparison is specific to the common top-k design and does not test the original threshold-based false-selection guarantee of classical stability selection.

D. Feature-Budget Sensitivity

Budget sensitivity in the p > n tasks showed that L1 remained above the all-features mean for both learners at 10%, 25%, and 50%. LightGBM performance was nearly identical at 10% and 25% (0.679 and 0.679), while logistic regression was highest at 25% (0.608). Signal recall rose from 0.794 at the 10% budget to 1.000 at both 25% and 50%, and L1 chance-corrected stability increased from 0.659 to 0.860 and 0.928. The 50% budget therefore maximised reproducibility and signal coverage, whereas the 10-25% budgets retained fewer

ISBN: 978-605-80193-6-2

variables and produced slightly higher mean predictive performance.

TABLE IV. L1 BUDGET SENSITIVITY IN THE p > n REGIME

Budget	LGBM F1	LR F1	Signal R
10%	0.679	0.608	0.794
25%	0.679	0.608	1.000
50%	0.668	0.597	1.000
All features	0.660	0.580	-

V. DISCUSSION

The expanded benchmark supports a more cautious interpretation of feature-selection stability. Corrected stability was not a universal proxy for predictive gain, and the positive p > n stability-gain trend at the primary budget was not statistically significant. At the same time, the observed stability profile changed markedly by dimensional regime: FScore was most stable among p <= n tasks, whereas L1 was most stable in p > n tasks. The regime dependence observed here therefore concerns selector reproducibility patterns across dimensional settings; the data do not establish a significant regime-specific stability-utility association.

The most consistent selector-level pattern concerned ordinary L1. It had the highest corrected stability overall, the highest logistic-regression macro-F1 across all 25 tasks, and the highest mean macro-F1 for both downstream learners in the p > n regime. L1 also recovered the complete known signal set in the high-dimensional simulations at the 25% budget. StabilityL1, by contrast, used 25 half-samples per outer split, was slower, and did not yield higher fixed-budget stability or predictive performance than ordinary L1. This comparison applies only to the common top-k benchmark and should not be interpreted as a test of the theoretical error-control guarantees of threshold-based stability selection.

Budget sensitivity reinforces the need to separate feature-selection objectives. A 10% L1 budget retained fewer variables and recovered about 79% of known signal, while 25% recovered the full signal set and produced the highest logistic-regression mean; 50% further increased corrected stability but slightly reduced mean predictive performance. Feature-selection policy should therefore be tied to the analytical or application objective: parsimony, reproducibility, signal coverage, or predictive utility, rather than reduced to a universal preference for the most stable method or the largest retained subset.

VI. LIMITATIONS

The benchmark contains 25 tasks and eleven p > n scenarios, but only eight tasks are real datasets; the synthetic component expands dimensional regimes and provides known-signal ground truth rather than replacing broader real-data validation. The Spector dataset has only three predictors, so the minimum three-feature budget selects the full set and makes the Nogueira denominator degenerate; it contributes to predictive analyses but not to corrected-stability inference. The StabilityL1 baseline follows half-sampling and frequency ranking but uses a fixed top-k output, so the original probability-of-false-selection guarantee is not claimed. Regime-specific correlations are stratified analyses rather than a formal interaction test, and differences between p <= n and p > n should not be interpreted as statistically confirmed effect modification. The downstream evaluation is limited to logistic regression and LightGBM, and the L1 regularisation grid approximates the target feature budget rather than being tuned by nested predictive validation. Future work should extend the protocol to larger public benchmark suites, additional learners, correlated-block simulations, and threshold-based stability-selection variants.

VII. CONCLUSION

Across 25 classification tasks and ten repeated train/test resamples, chance-corrected feature-selection stability did not serve as a universal proxy for predictive gain. Stability rankings nevertheless changed across dimensional regimes, with FScore most stable among p <= n tasks and L1 most stable in p > n tasks. In the high-dimensional regime, L1 also yielded higher mean macro-F1 than the all-features baseline for both downstream learners while recovering the complete known signal set at the 25% budget. These findings support a regime-aware evaluation in which corrected stability, downstream prediction, signal recovery, computational cost, and feature budget are reported as separate criteria rather than assuming that the most reproducible subset is automatically the most useful.

References

[1] I. Guyon and A. Elisseeff, An introduction to variable and feature selection, Journal of Machine Learning Research, vol. 3, pp. 1157-1182, 2003.

[2] R. Kohavi and G. H. John, Wrappers for feature subset selection, Artificial Intelligence, vol. 97, no. 1-2, pp. 273-324, 1997, doi: 10.1016/S0004-3702(97)00043-X.

[3] Y. Saeys, I. Inza, and P. Larrañaga, A review of feature selection techniques in bioinformatics, Bioinformatics, vol. 23, no. 19, pp. 2507-2517, 2007, doi: 10.1093/bioinformatics/btm344.

[4] G. Chandrashekar and F. Sahin, A survey on feature selection methods, Computers & Electrical Engineering, vol. 40, no. 1, pp. 16-28, 2014, doi: 10.1016/j.compeleceng.2013.11.024.

[5] J. Li, K. Cheng, S. Wang, F. Morstatter, R. P. Trevino, J. Tang, and H. Liu, Feature selection: A data perspective, ACM Computing Surveys, vol. 50, no. 6, art. 94, 2017, doi: 10.1145/3136625.

[6] A. Kalousis, J. Prados, and M. Hilario, Stability of feature selection algorithms: a study on high-dimensional spaces, Knowledge and Information Systems, vol. 12, no. 1, pp. 95-116, 2007, doi: 10.1007/s10115-006-0040-8.

[7] S. Nogueira, K. Sechidis, and G. Brown, On the stability of feature selection algorithms, Journal of Machine Learning Research, vol. 18, no. 174, pp. 1-54, 2018.

[8] N. Meinshausen and P. Bühlmann, Stability selection, Journal of the Royal Statistical Society:

 ISBN: 978-605-80193-6-2

Series B, vol. 72, no. 4, pp. 417-473, 2010, doi: 10.1111/j.1467-9868.2010.00740.x.

[9] A. Bommert, X. Sun, B. Bischl, J. Rahnenführer, and M. Lang, Benchmark for filter methods for feature selection in high-dimensional classification data, Computational Statistics & Data Analysis, vol. 143, art. 106839, 2020, doi: 10.1016/j.csda.2019.106839.

[10] L. Yu and H. Liu, Efficient feature selection via analysis of relevance and redundancy, Journal of Machine Learning Research, vol. 5, pp. 1205-1224, 2004.

[11] R. Battiti, Using mutual information for selecting features in supervised neural net learning, IEEE Transactions on Neural Networks, vol. 5, no. 4, pp. 537-550, 1994, doi: 10.1109/72.298224.

[12] L. Breiman, Random forests, Machine Learning, vol. 45, pp. 5-32, 2001, doi: 10.1023/A:1010933404324.

[13] R. Genuer, J.-M. Poggi, and C. Tuleau-Malot, Variable selection using random forests, Pattern Recognition Letters, vol. 31, no. 14, pp. 2225-2236, 2010, doi: 10.1016/j.patrec.2010.03.014.

[14] C. Ding and H. Peng, Minimum redundancy feature selection from microarray gene expression data, Journal of Bioinformatics and Computational Biology, vol. 3, no. 2, pp. 185-205, 2005, doi: 10.1142/S0219720005001004.

[15] G. Brown, A. Pocock, M.-J. Zhao, and M. Luján, Conditional likelihood maximisation: A unifying framework for information theoretic feature selection, Journal of Machine Learning Research, vol. 13, pp. 27-66, 2012.

[16] H. Peng, F. Long, and C. Ding, Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 8, pp. 1226-1238, 2005, doi: 10.1109/TPAMI.2005.159.

[17] V. Bolón-Canedo and A. Alonso-Betanzos, Ensembles for feature selection: A review and future trends, Information Fusion, vol. 52, pp. 1-12, 2019, doi: 10.1016/j.inffus.2018.11.008.

[18] J. Demšar, Statistical comparisons of classifiers over multiple data sets, Journal of Machine Learning Research, vol. 7, pp. 1-30, 2006.

[19] F. Pedregosa et al., Scikit-learn: machine learning in Python, Journal of Machine Learning Research, vol. 12, pp. 2825-2830, 2011.

[20] S. Seabold and J. Perktold, Statsmodels: Econometric and statistical modeling with Python, in Proceedings of the 9th Python in Science Conference, pp. 92-96, 2010, doi: 10.25080/MAJORA-92BF1922-011.

[21] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, LightGBM: A highly efficient gradient boosting decision tree, in Advances in Neural Information Processing Systems 30, pp. 3146-3154, 2017.

ISBN: 978-605-80193-6-2

Classification of Coffee Beans Under Non-Ideal Conditions

Ali BİRCAN[1*], Nurdan AKHAN BAYKAN[2]
[1]Computer Engineering Department, Karamanoğlu Mehmetbey University, Karaman, Türkiye
[2]Computer Engineering Department, Konya Technical University, Konya, Türkiye
* Corresponding author e-mail: alibircan@kmu.edu.tr

***Abstract*— Computer-aided (based) classification aims to separate data with digitally-controlled systems. Classification algorithms are studied in many different areas and play important roles in solving different problems. In this study, a coffee beans database under non-ideal conditions is created and with this database success rate of different classification models on images is compared under non-ideal conditions, and results are presented.**

***Keywords*— Image processing, machine learning, classification, coffee beans.**

I. INTRODUCTION

In recent years, studies on the quality and categorical classification of agricultural products have increased. Some of these studies are on the determination of raw, unprocessed product quality. This surge in research is primarily driven by the growing global demand for food safety, stringent international export standards, and the escalating need to minimize post-harvest food waste. Traditional manual inspection techniques are inherently labor-intensive, time-consuming, prone to human error, and subjective. Consequently, transitioning toward rapid, non-destructive, and objective automated evaluation systems has become an imperative priority for both producers and supply chain stakeholders.

The coffee plant is produced and consumed all over the world. Since ancient times, coffee beans have been transported all over the world. Coffee beans are classified according to various characteristics and are grown in different regions. However, the coffee beans of these plants grown in various regions and continents also have different characteristics.

The quality of green coffee beans is examined with various aspects as variety, soil, climate, processing methods, and more. Coffee bean quality assessments fall into three categories as green coffee assessment, sensory assessment, and chemical assessment [1].

Coffea arabica, known commercially as Arabica, was first cultivated and traded on the Arabian Peninsula, from which it derives its name. Coffea canephora, known commercially as Robusta, established in the late nineteenth century was adopted in South and South-East Asia Arabica is cultivated across a wider range of regions than Robusta because it is grown at higher altitudes and is traded as a premium product.

Image acquisition is the first step in determining coffee bean characteristics. A study by [2] proposed a method for managing the quality of coffee beans using digital images. The Robusta coffee beans images were taken using a Sony DCS-800 20.1 Megapixel digital camera. Reference [3] preferred to use a black background instead of white. Other studies have also preferred different devices such as flatbed scanners [4] and smartphones [5].

Reference [6] introduced a computer vision system that takes green coffee beans into the CIE L*a*b* (Commission Internationale d'Eclairage) color space and classifies them according to their colors. Artificial neural networks (ANN) were used as the transformation model and the Bayesian classifier was used to classify coffee beans into four groups: whitish, reed green, green and bluish green. ANNs recorded generalization error of 1.15%, and Bayesian classifier was able to classify all beans according to their expected class (100% accuracy).

Reference [7] compared the performance of different data mining techniques in the classification of green coffee beans. A library called Classification Learner was used in the study, and the highest accuracy among 22 classifiers in this library was obtained by Coarse Tree Algorithm with 94.1%. The lowest classification percentage was obtained by the Boosted Tree Algorithm with an accuracy of 28.2%. Of the 22 algorithms, only 4 had a classification accuracy of less than 90%, while 18 algorithms achieved an accuracy of over 90%.

A study of real-time classification of coffee beans by Huang, Chou and Lee [8] tried to solve the problem of time and nature of the manual selection of coffee beans for specialty coffee products. First, image processing and data augmentation technologies were used. Then, deep learning was used together with convolutional neural networks to analyse the image information. The false positive rate was 0.1007 and the overall true recognition rate was 93%.

Classification of coffee beans was made not only with visible light but also using Near Infrared Spectroscopy (NIRS) [9]. Okubo and Kurata's study focused on determining green coffee beans' production area using LKS and classification analysis. Soft independent modelling of class analogy (SIMCA) was applied as a classification method.

In recent years, deep learning methods have also been used in the classification of coffee beans. CNN was used for classifying green coffee beans according to their defects by Pinto et al. [10]. The aim of the study is, as a first step, to develop an image processing system that classifies images of green coffee beans according to each type of defect. As a result, coffee beans were classified according to defect types with varying success from 72.4% to 98.7%. Then, a method using CNN in mobile systems has been suggested by Janandi and Cenggoro [11] to determine the quality of coffee beans. In the study, a mobile application supported by a deep learning-based model has been developed which will be used for automatic classification with a mobile phone camera. The study used ResNet-152 and VGG16 models and the highest success was 73.3%.

ISBN: 978-605-80193-6-2

In 2021, Yang, Jhong and Hsia[12] proposed a deep learning and transfer learning-based method for the classification of green coffee beans. In this study, a CNN-based method with local feature extraction performance is presented. Experimental results show that the proposed method can be used effectively to perform the recognition of green coffee beans and achieves an F1 score of about 97%.

In [13], a study has been carried out with transfer learning and deep convolutional neural networks, presented methods of classifying coffee beans according to their types. The goal of this study is to classify images of 3 different coffee beans with transfer learning method using 4 different CNN based models (SqueezeNet, Inception V3, VGG16, and VGG19). The dataset used for training of the models was solely created specifically for the study and 1554 coffee bean images of Espresso, Kenya and Starbucks Pike Place were collected. Cross-validation was used to test the models and to assess their robustness. Classification success, precision, recall, F-1 Score and ROC curves were used for detailed analysis of performance benchmarking. Average classification success of the models was obtained as 87.3% for SqueezeNet, 81.4% for Inception V3, 78.2% for VGG16, and 72.5% for VGG19.

Studies have been carried out by taking images in ideal conditions for the classification of coffee beans. In this study, the effect of environment conditions on the success of the classifier was investigated. For this purpose, images were taken under non-ideal conditions, and classification processes were carried out with different machine learning algorithms.

II. MATERIAL AND METHOD

A. Classifiers

In statistics and machine learning, classification is a process of determining the category of a new test object based on the knowledge obtained from training set [14]. Different methods are used to determine which class the new observation belongs to. For this purpose, the new observation and the samples in the training set are compared with different methods, and the ones that show the most similarity are assigned to the same category. In this study, the performance of basic machine learning algorithms and the CNN model was compared in the classification of coffee beans.

1) Support Vector Machine (SVM): SVM is a supervised learning model, with associated learning algorithms that examine data for classification, recognition and regression analysis tasks [15], [16], [17]. Similar to other supervised machine learning methods, SVM takes a set of train set with membership label. Then SVM training algorithm creates boundary (or boundaries) for assigning new examples to one category or the other. With only two output labels, SVM acts as nonprobabilistic binary linear classifier. SVM maps samples in training datasets to points in the training space and tries to maximize the width of the gap between the categories (Fig. 1). SVM can works with different kernels. The three kernels considered in this study are linear, the polynomial and the RBF kernel.

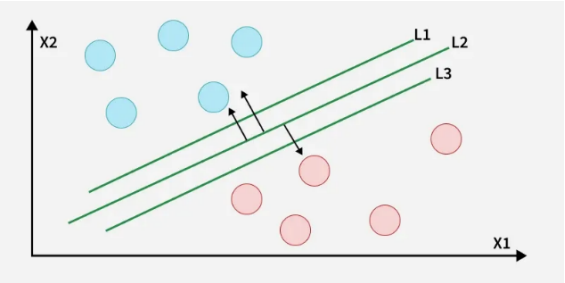


Fig. 1. Support Vector Machine Classification [18] .

2) k-Nearest Neighboor (k-NN): k-NN was developed by Fix and Hodges. [19], [20]. It is a non-parametric supervised learning method: no fixed functional form is estimated during training, the training samples are retained and the neighbourhood parameter k is a hyper-parameter that has to be selected. Because the method is non-parametric, its computational cost at prediction time grows with the size of the training set. Generally it is used for classification, clustering and regression tasks. In most tasks the input consists of the k closest training examples in a dataset. Because the algorithm relies on distance for classification, preprocessing of the dataset (such as normalizing the training dataset) can improve its accuracy significantly (in the cases that features represent different unit representations or in very different scale proportions). (Fig. 2).

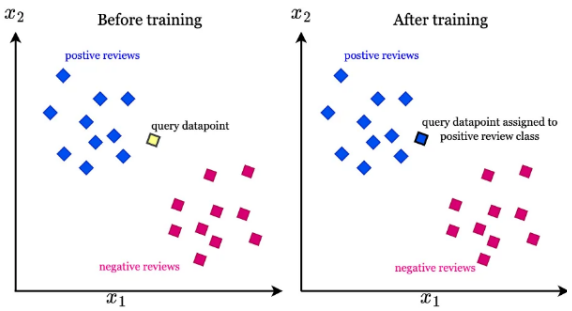


Fig. 2. k-NN classification [21] .

3) Decision Tree (DT): A decision tree creates nodes that act as decisions, and these nodes are used to split the dataset into subsets. Leaf nodes are used as the final decision nodes, and each node asks a different question, with given answer helping navigation to the leaf nodes. In decision trees training, for finding the best class separation loss (therefore splitting the dataset more accurately) Classification and Regression Tree (CART) algorithm is employed. Different methods are used to measure the quality of the separation: the Gini index, information gain, or mean squared error (MSE). The best-known disadvantages of decision trees are their bias and their tendency to memorize the training data [22] (Fig. 3).

ISBN: 978-605-80193-6-2

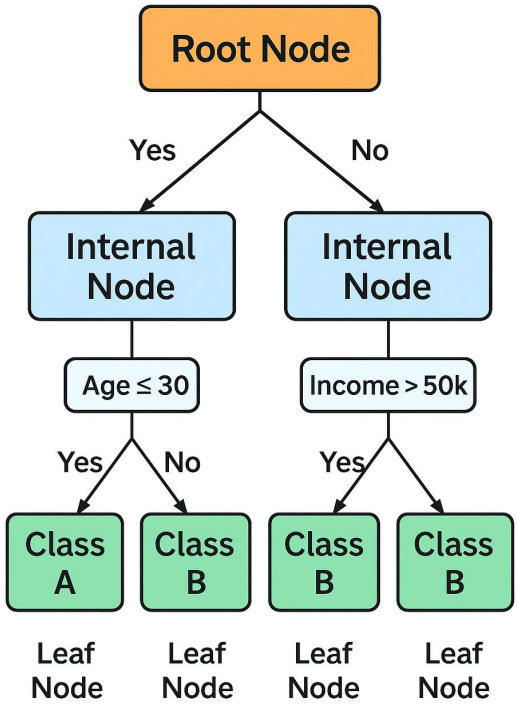


Fig. 3. Decision Trees [23] .

4) Random Forest (RF): Random forest algorithms consist of unrelated decision trees generated with using of bagging and feature randomness mechanisms [24]. A random forest consists of decorrelated decision trees. Usually feature aggregation methods are employed. That process creates decision trees that have low correlation between them. The main difference between decision trees and random forests is that while decision trees evaluate all possible feature separations, random forests operate by selecting only a subset of those features (Fig. 4).

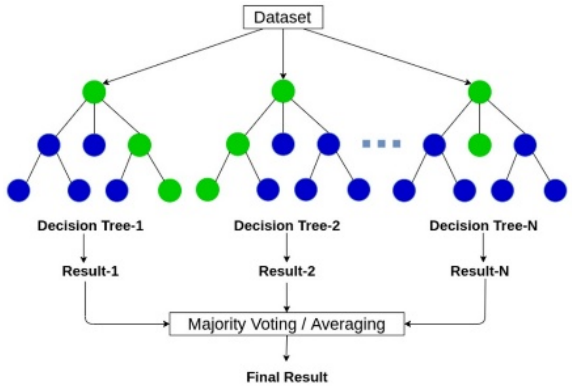


Fig. 4. Random Forest [25].

5) Naïve Bayes (NB): Naive Bayes method is a supervised learning algorithm. The mechanism of NB is the application of Bayes' theorem under the assumption that all input features contributes independently to the target-class probability [26]. Naive Bayes classifiers have different approaches for assumption of distribution. Generally, this type of classifier uses a small subset of training dataset to make estimation for the required parameters. Usually, NB classifiers can be tuned to be extremely fast compared with more complex methods (Fig. 5).

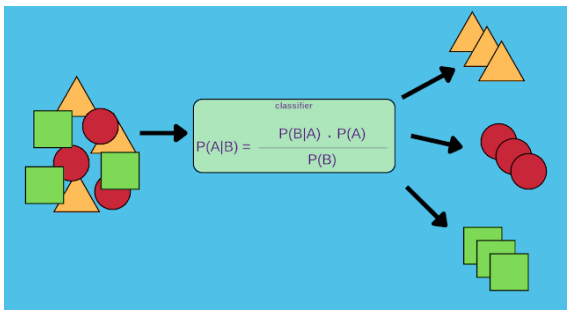


Fig. 5. Naïve Bayes [27].

6) Convolutional Neural Networks (CNN): Artificial neural networks (ANNs) are mostly multilayered neural networks used in machine learning to solve nonlinear problems of different areas. ANNs are models created based on the functioning of the human brain. Neural networks typically have input and output layers and contain a hidden layer (or layers) designed to transform the input into a representation that is relayed to the output layer and that supports the decision-making process. The number of hidden layers can differ depending on the type and the complexity of the problem in type and count, and also used to find patterns or manipulate the data that are too complex for a human to make or detect [28].

Traditional Artificial Neural Networks (ANNs), such as Multi-Layer Perceptrons (MLPs), require high-dimensional image matrices to be flattened into one-dimensional vectors. This flattening process completely disrupts the intrinsic spatial and structural correlations among neighbouring pixels [29]. Furthermore, fully connected topologies suffer from the curse of dimensionality; the massive parameter volume exponentially increases the risk of overfitting and computational burden.

In contrast, Convolutional Neural Networks (CNNs) overcome these limitations through three foundational architectural mechanisms: local receptive fields, shared weights, and spatial subsampling (pooling). By convolving learnable parameter kernels across the input tensor, CNNs inherently preserve spatial-temporal hierarchies, extracting low-level features (e.g., edges, textures) in initial layers and abstract, high-level semantic representations (e.g., geometric forms, complex defect patterns) in deeper layers [30]. The weight-sharing paradigm drastically curbs parameter count compared to fully connected architectures, ensuring high computational efficiency and robust translation invariance across the field of view. Consequently, deep convolutional architectures serve as the backbone for end-to-end visual classification, directly mapping raw pixel arrays to categorical labels without reliance on handcrafted feature extraction.

B. Image Processing

In image processing, three basic processing stages, namely preprocessing, feature extraction, and result stage (classification, clustering, interpretation, etc.) are applied to the data received [31].

ISBN: 978-605-80193-6-2

In the preprocessing stage, cleaning of unwanted data such as distortions, and incorrect field information which complicates the subsequent processing stages on the received images; extracting edge information of objects in the image; improving the picture to facilitate the analysis of the picture to separate different objects in the image, various operations such as moving the image to other coordinate planes are carried out [32].

In image processing, Gaussian function can be applied to the source image for achieving blurring effect. This type of blur is known as Gaussian blur and is commonly used for general noise reduction and detail reduction in image processing. [33].

Image thresholding is a set of processes for binarizing an image based on pixel densities. The input of the process usually an image that contains pixels that have color range (depending of the image) and thresholding range (or ranges). The output is a binary image. Usually using simple threshold means you have to specify the threshold value manually and involves process trial and error. Otsu method is frequently used in the literature for automatic threshold determination [34].

The Otsu method processes the image histogram by segmenting objects, minimizing variance in each class. Generally, this technique yields favourable results for bimodal images.

III. EXPERIMENTAL STUDY

A. Non-Ideal Conditions

In this study, the term "non-ideal conditions" refers to situations that differ from a controlled studio environment—such as variations in lighting, background, camera distance, viewing angle, shadows/occlusion (object overlap), and the camera/acquisition setup. Backgrounds are wrinkled A4 paper and camera distances and viewing angles is varied. There is no controlled light, so images have shadows of beans. Last, images captured with IPhone SE2 camera, rather than a machine vision specialized camera.

B. Dataset Characteristics

The dataset contains six green-coffee-bean classes. Beans of the Indian Robusta class were used as the single Robusta class, and beans from El Salvador, Guatemala, Honduras, Costa Rica and Peru as the five Arabica classes. Extracted bean image counts and shares can be followed in Table I.

TABLE I. Used Beans' information

Class	Species	Extracted bean images	Share of dataset
Indian Robusta	Robusta	299	21.8 %
Costa Rica	Arabica	266	19.4 %
El Salvador	Arabica	233	17.0 %
Guatemala	Arabica	205	15.0 %
Peru	Arabica	201	14.7 %
Honduras	Arabica	165	12.1 %
Total	—	1,369	100 %

C. Application

Images were taken from the mobile phone camera (iPhone SE2). 2 images were taken for each coffee bean set. The images were first pre-processed, and then, since the dataset was not balanced, it was tried to be balanced using resampling. All coding processes were done with PyCharm IDE and Python(3.9) programming language. Various libraries were used, mainly OpenCV, TensorFlow, Keras, and sckit-learn.

After the images taken were separated into individual beans, they were saved in the database. For the process of separating into individual beans, images were first converted to greyscale, and contour finding was performed after thresholding and opening processes were performed. The contours found were used for the separation of beans from the original image.

After the images were taken, the following preprocessing steps were applied.

1) Original Image (Fig. 6-a) is copied and converted to a gray level (Fig. 6-b).

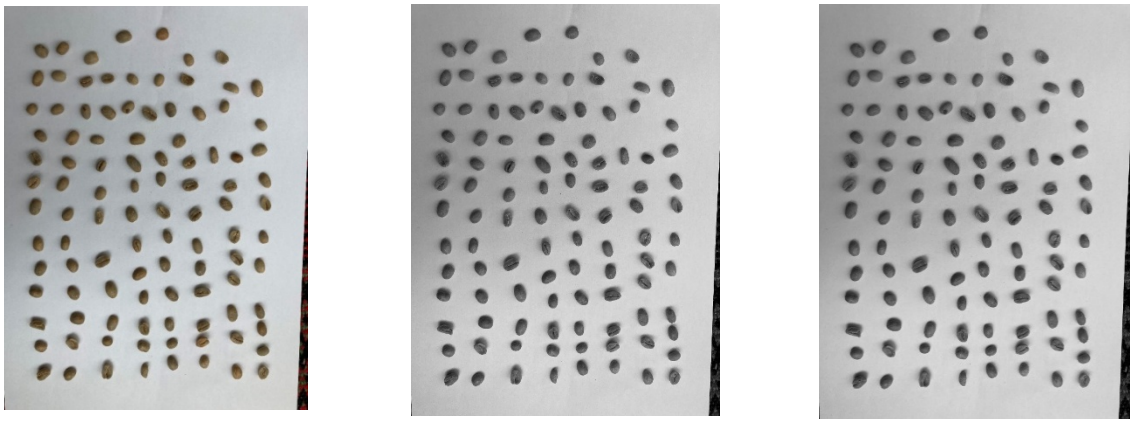
2) Gaussian blur was applied to gray image with a (5x5) kernel (Fig. 6-c).

3) Thresholding was done on the obtained image with the Otsu method (Fig. 6-d).

4) A 7x7 square structural element was used for the dilating operation (Fig. 6-e).

5) The contour-finding process was performed on the image. The reason for this process is to find the location and size of multiple beans on the same image. Found contours and their locations were determined on the original image and saved into separate files (Fig. 6-f).

ISBN: 978-605-80193-6-2



(a) Original image (b) Gray level image (c) Blurred image

(d) Thresholded image (e) Image after dilating (f) Detected beans in image

Fig. 6. Steps of preprocessing

6) In addition, especially for CNN, histogram equalization was performed on color images to increase the number of images (hence feeding the system with more data) while the data set is being loaded into the system. The first image color space changed to Luma Chroma color space. Then, histogram equalization was made on the Y channel (Light) of this image (to avoid light-induced irregularities) and it was returned to RGB color space. These images are also included in the dataset (Fig. 7).

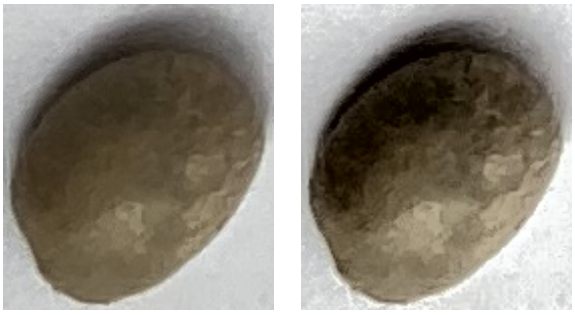


Fig. 7. Images with same bean (origin: Peru-Arabica). In left extracted bean from larger image, in right histogram equalized version.

7) The beans of each class were then saved into their respective class directory.

After the necessary operations were made on the images, the classification stage was started. At this stage, more than one classification method was used. Images have been adjusted to 120x120 px dimensions to have a fixed input, and RGB color space has continued to be used.

After the resizing process, there was the the class imbalance in the unbalanced data set (the case that the number of beans belonging to different classes is not the same) (Table I). To prevent the majority classes from dominating the objective, the training partition alone was rebalanced; the test partition was never resampled and is reported exactly as extracted, so that every test score is computed on real bean images. For the classical classifiers, SMOTE (Synthetic Minority Over-sampling Technique) was applied to the training partition only, after the split and before fitting, generating synthetic minority-class samples by interpolating between each minority sample and its k nearest minority neighbors. For the convolutional network, SMOTETomek was applied to the dataset as same pipeline in classical methods.

D. CNN Architecture and Training Settings

The convolutional network used in this study is compact six-class architecture built with the Keras functional API and is specified in Fig. 8.

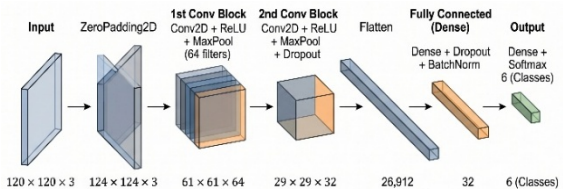


Fig. 8. CNN Architecture.

ISBN: 978-605-80193-6-2

The training and validation curves (of one the seeds) are given in Fig. 9.

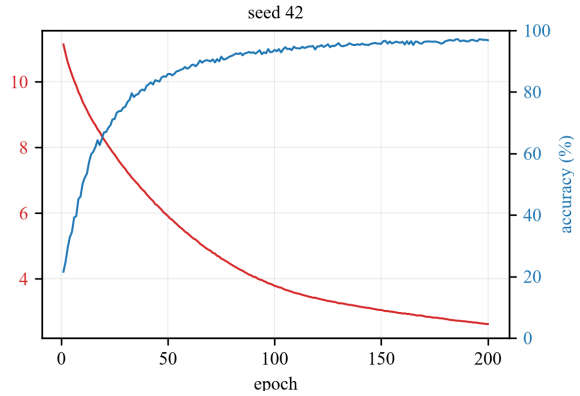


Fig. 9. Loss and accuracy curve of one the CNN runs (seed 42)

IV. RESULTS AND DISCUSSION

The performance of the thirteen configurations evaluated in this study is given in Table II.

TABLE II. Performance of the Configurations

Model	Accuracy (%)	Precision (macro, %)	Recall (macro, %)	F1-score (macro, %)
SVM (linear)	55.47	54.77	55.0	54.59
SVM (polynomial)	57.3	57.17	57.17	56.79
SVM (RBF)	51.82	50.19	51.13	50.16
RF (gini)	55.84	54.87	54.64	54.37
RF (entropy)	59.85	58.6	58.28	58.07
RF (log-loss)	59.85	58.6	58.28	58.07
k-NN (k = 3)	41.97	45.94	43.26	42.45
k-NN (k = 5)	42.34	46.37	43.85	42.3
k-NN (k = 7)	42.34	47.2	43.73	42.58
k-NN (k = 9)	40.88	45.53	42.03	40.9
DT	36.13	34.9	35.29	35.04
NB (Gaussian)	37.96	39.48	39.01	34.63
CNN	**87.96**	**88.48**	**87.81**	**87.91**

The confusion matrix of the best configuration with the convolutional network is shown in Fig. 9.

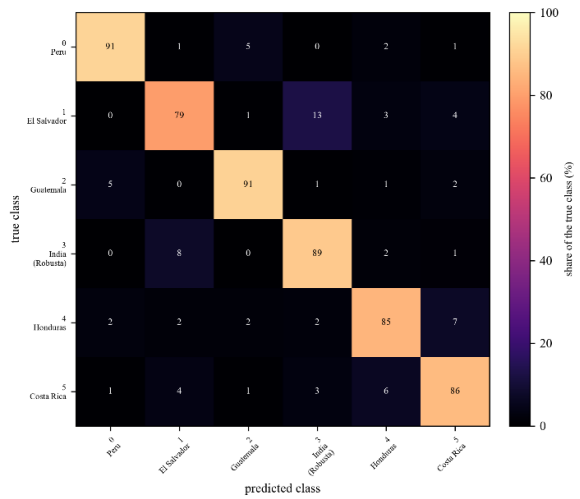


Fig. 9. Confusion matrix for CNN run.

The convolutional neural network delivers the best classification performance because its data-driven feature extraction jointly models color distributions, boundary shapes, and surface textures, providing essential robustness against variable illumination that pure color statistics fail to handle. The linear SVM emerges as the top-performing classical model, closely trailing the CNN. In contrast, Euclidean distances in k-NN are skewed by background noise and global intensity shifts, causing performance to drop as the neighbourhood expands. While Random Forests performed poorly; Decision Tree and Gaussian Naïve Bayes fail to generalize.

References

[1] J. N. Wintgens and others, Coffee: growing, processing, sustainable production. A guidebook for growers, processors, traders, and researchers. WILEY-VCH Verlag GmbH & Co. KGaA, 2004.

[2] Y. Hendrawan, S. Widyaningtyas, and S. Sucipto, "Computer vision for purity, phenol, and pH detection of Luwak Coffee green bean," TELKOMNIKA (Telecommunication Computing Electronics and Control), vol. 17, no. 6, pp. 3073–3085, 2019.

[3] F. F. L. dos Santos, J. T. F. Rosas, R. N. Martins, G. de M. Araújo, L. de A. Viana, and J. de P. Gonçalves, "Quality assessment of coffee beans through computer vision and machine learning algorithms," 2020.

[4] J. N. C. Sarino, M. M. Bayas, E. R. Arboleda, E. C. Guevarra, and R. M. Dellosa, "Classification of coffee bean degree of roast using image processing and neural network," International Journal of Scientific and Technology Research, vol. 8, no. 10, pp. 3231–3233, 2019.

[5] E. M. de Oliveira, D. S. Leme, B. H. G. Barbosa, M. P. Rodarte, and R. G. F. A. Pereira, "A computer vision system for coffee beans classification based on computational intelligence techniques," J. Food Eng., vol. 171, pp. 22–27, 2016.

[6] E. R. Arboleda, "Comparing performances of data mining algorithms for classification of green coffee beans," Int. J. Eng. Adv. Technol, vol. 8, no. 5, pp. 1563–1567, 2019.

[7] N.-F. Huang, D.-L. Chou, and C.-A. Lee, "Real-time classification of green coffee beans by using a convolutional neural network," in 2019 3rd International Conference on Imaging, Signal Processing and Communication (ICISPC), 2019, pp. 107–111.

[8] N. Okubo and Y. Kurata, "Nondestructive classification analysis of green coffee beans by using near-infrared spectroscopy," Foods, vol. 8, no. 2, p. 82, 2019.

[9] C. Pinto, J. Furukawa, H. Fukai, and S. Tamura, "Classification of Green coffee bean images basec on defect types using convolutional neural network (CNN)," in 2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA), 2017, pp. 1–5.

ISBN: 978-605-80193-6-2

[10] R. Janandi and T. W. Cenggoro, "An Implementation of Convolutional Neural Network for Coffee Beans Quality Classification in a Mobile Information System," in 2020 International Conference on Information Management and Technology (ICIMTech), 2020, pp. 218–222.

[11] P.-Y. Yang, S.-Y. Jhong, and C.-H. Hsia, "Green coffee beans classification using attention-based features and knowledge transfer," in 2021 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW), 2021, pp. 1–2.

[12] Y. Unal, Y. S. Taspinar, I. Cinar, R. Kursun, and M. Koklu, "Application of pre-trained deep convolutional neural networks for coffee beans species detection," Food Anal. Methods, vol. 15, no. 12, pp. 3232–3243, 2022.

[13] E. Alpaydin, Introduction to Machine Learning. in Adaptive Computation and Machine Learning series. MIT Press, 2014. [Online]. Available: https://books.google.com.tr/books?id=7f5bBAAAQBAJ

[14] B. E. Boser, I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," in Proceedings of the fifth annual workshop on Computational learning theory, 1992, pp. 144–152.

[15] C. Cortes and V. Vapnik, "Support-vector networks," Mach. Learn., vol. 20, pp. 273–297, 1995.

[16] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," Mach. Learn., vol. 46, pp. 389–422, 2002.

[17] GeeksforGeeks, "Support Vector Machine (SVM) Algorithm," 2024. [Online]. Available: https://www.geeksforgeeks.org/machine-learning/support-vector-machine-algorithm

[18] E. Fix and J. L. Hodges Jr, "Discriminatory analysis-nonparametric discrimination: Small sample performance," 1952.

[19] T. Cover and P. Hart, "Nearest neighbor pattern classification," IEEE Trans. Inf. Theory, vol. 13, no. 1, pp. 21–27, 1967.

[20] Intuitive Tutorials, "K-Nearest Neighbors Algorithm," Apr. 2023. [Online]. Available: https://intuitivetutorial.com/2023/04/07/k-nearest-neighbors-algorithm/

[21] M. Somvanshi, P. Chavan, S. Tambade, and S. V Shinde, "A review of machine learning techniques using decision tree and support vector machine," in 2016 international conference on computing communication control and automation (ICCUBEA), 2016, pp. 1–7.

[22] DigitalOcean, "Decision Trees Made Simple: Machine Learning Explained," Jun. 2025. [Online]. Available: https://www.digitalocean.com/community/tutorials/decision-trees-machine-learning-explained

[23] Y. Liu, Y. Wang, and J. Zhang, "New machine learning algorithm: Random forest," in Information Computing and Applications: Third International Conference, ICICA 2012, Chengde, China, September 14-16, 2012. Proceedings 3, 2012, pp. 246–252.

[24] Consuledge, "A Very Short Introduction of Random Forests," Dec. 2024. [Online]. Available: https://consuledge.com.au/blog/mastering-random-forests-a-beginners-guide-to-smarter-machine-learning/

[25] Y. Huang and L. Li, "Naive Bayes classification algorithm based on small sample set," in 2011 IEEE International conference on cloud computing and intelligence systems, 2011, pp. 34–39.

[26] P. Bhandari, "Naive Bayes Classifier," Apr. 2024. [Online]. Available: https://medium.com/@palakvb02/naive-bayes-classifier-211afcb07bff

[27] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, 2015.

[28] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.

[29] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," Adv. Neural Inf. Process. Syst., vol. 25, 2012.

[30] B. Kang, "A Review on Image & Video Processing," International Journal of Multimedia and Ubiquitous Engineering, vol. 2, 2007.

[31] S. P. Kuruvilla, J., Sukumaran D., Sankar, A., Joy, "A review on image processing and image segmentation," in International Conference on Data Mining and Advanced Computing, 2016, pp. 198–203.

[32] C. Solomon and T. Breckon, Fundamentals of Digital Image Processing: A practical approach with examples in Matlab. John Wiley & Sons, 2011.

[33] N. Otsu, "A threshold selection method from gray-level histograms," IEEE Trans. Syst. Man Cybern., vol. 9, no. 1, pp. 62–66, 1979.

 ISBN: 978-605-80193-6-2

A Model-Independent Upper Bound on Prediction Accuracy for M- and NM-Polynomial Topological Indices

Talha KIRCA[1*], Merve GÖRGÜLÜ[1], Hasan YILDIRIM[1]

[1]Karamanoğlu Mehmetbey University, Kamil Özdağ Faculty of Science, Department of Data Science and Analytics, Karaman, Türkiye

* Corresponding author e-mail: kircatalha1@gmail.com

***Abstract*—Topological indices derived from the M-polynomial and the neighborhood M-polynomial (NM-polynomial) are widely used to predict molecular properties, usually on about twenty compounds and without checking how many of them the indices can tell apart. We computed both index families automatically from SMILES for two public benchmarks: ESOL (1128 compounds, aqueous solubility) and FreeSolv (642 compounds, hydration free energy). On ESOL, 638 molecules (56.6 percent) fall into 138 groups whose members share an identical set of 20 index values; on FreeSolv, 479 molecules (74.6 percent) belong to such groups. Molecules with identical inputs must receive identical predictions, so this degeneracy sets an upper bound that no model can exceed. Computed from the data alone, the bound limits the coefficient of determination to 0.833 on ESOL and 0.565 on FreeSolv. Adding the M indices to the NM indices did not separate any additional molecules on ESOL. The best model we fitted reached 0.460 on ESOL under group-aware cross-validation, about half of the bound, while six simple descriptors, some of which reflect atom type, exceeded it. The cause is that the indices use only vertex degrees on a graph without hydrogen atoms, so they ignore atom type, bond order and hydrogen count: an alkane, an alcohol, a ketone and a urea can have the same index values. Y-randomization and applicability domain analysis support these results, and the bound is robust to exact grouping and to merging repeated records. The bound needs no model and applies to any prediction task with a fixed set of discrete features.**

***Keywords*—Topological indices, M-polynomial, NM-polynomial, structural degeneracy, QSPR, upper bound, model evaluation**

I. INTRODUCTION

Topological indices are numbers computed from the molecular graph with the hydrogen atoms removed. They are cheap to compute, deterministic and easy to interpret. For this reason they have been in use since the Wiener index [1] and the Randić connectivity index [2]. More recently, Deutsch and Klavžar introduced the M-polynomial [3]. It records how the edges of the graph are distributed by the degrees of their end vertices, and many standard indices can be obtained from it with simple operators. The idea was later extended to neighborhood degree sums, giving the neighborhood M-polynomial or NM-polynomial [4], [5], and to the complement graph, giving the CoM- and CoNM-polynomials [6], [7].

Many applications follow the same pattern. A study selects one drug class, computes one or two polynomial families for a handful to about twenty compounds, fits a few physicochemical properties to the indices, and reports the fit [7], [8]. These studies rarely ask a basic question: *how many of the molecules can the descriptors tell apart?* The discriminating power of topological indices has been studied on exhaustively generated graphs [9], [10], but not against measured properties. Limits on achievable accuracy have been derived from experimental error in the target values [11], and modelability indices estimate from nearest neighbors how well a data set can be modeled [12]. Neither gives a strict bound from the resolution of the descriptors. In statistical learning, the error caused by identical inputs with different outputs is part of the irreducible error [13]; here we compute it exactly from the degeneracy of the descriptors.

The answer matters because it sets a limit for every model. If two molecules have identical descriptor vectors, every model must give them the same prediction, whatever the algorithm and however much training data it has. If their measured values differ, part of the error can never be removed. The size of this error depends on the descriptors and the data, not on the model, and it can be computed exactly.

We study the polynomial families because they are how this literature produces its descriptors. The problem itself, however, lies in the representation. The indices are computed from vertex degrees on a graph without hydrogen atoms. The choice of polynomial affects how many molecules can be separated, but the loss of atom type and bond order comes from the graph itself.

With about twenty similar compounds, identical vectors are rare. The problem becomes visible on structurally diverse data sets, which have rarely been used with polynomial indices.

The same reasoning applies outside chemistry. A model built on a fixed set of discrete or coarsely binned features, such as counts, categorical process settings, binned sensor readings or soil classes, must give the same prediction to observations with identical features. Its accuracy is therefore limited by how finely the features divide the data. This limit is rarely reported. Molecular graph descriptors show it clearly, since more than half of a standard benchmark shares feature vectors. The procedure takes a few lines of code and can be used for any tabular data set whose features can coincide exactly. For continuous features, the tolerance-based version described in Section VI-B is more suitable.

The contributions of this paper are as follows.

- We computed the M- and NM-polynomial indices automatically from SMILES (simplified molecular-input line-entry system) strings for two public benchmarks. These data sets are 32 and 56 times larger than a typical set of twenty compounds.

ISBN: 978-605-80193-6-2

- We measured how often molecules share descriptor vectors and turned this into a model-independent upper bound on R^2 and a lower bound on the root mean square error (RMSE).
- We compared this bound with the accuracy of fitted models, with a published four-descriptor equation, and under two cross-validation schemes.

II. DESCRIPTORS

Let G be the molecular graph without hydrogen atoms. For a vertex u, d_u denotes its degree and $\delta_u = \sum_{v \in N(u)} d_v$ the sum of the degrees of its neighbors. For a vertex weight w_u, both polynomials have the form

$$P(G; x, y) = \sum_{i \le j} m_{ij}\, x^i y^j, \tag{1}$$

where m_{ij} is the number of edges uv of G with $\{w_u, w_v\} = \{i, j\}$. The M-polynomial uses $w_u = d_u$, and the NM-polynomial uses $w_u = \delta_u$.

From each polynomial we computed the same ten degree-based indices: the first and second Zagreb, modified second Zagreb, Randić, reciprocal Randić, symmetric division degree, harmonic, inverse sum indeg, augmented Zagreb and forgotten indices. This gives 2 × 10 = 20 descriptors per molecule. Each index is a sum over the edge partition, for example $M_1 = \sum m_{ij}(i + j)$ and $M_2 = \sum m_{ij} ij$.

The augmented Zagreb index contains the factor $(ij/(i + j - 2))^3$, which is undefined when $i = j = 1$. This happens only for graphs with two atoms, such as ethane, methanol and bromomethane. There are six such molecules in ESOL and eleven in FreeSolv. For these molecules the undefined term was skipped, and the cases were counted. Molecules with a single heavy atom (one in ESOL, three in FreeSolv) have no edges, so all their indices are zero.

III. DATA

ESOL [14] contains measured aqueous solubility values (log mol/L) for 1128 organic compounds. It is a standard regression benchmark in MoleculeNet [15]. The set is diverse. It ranges from methane to digoxin (1 to 55 heavy atoms) and includes simple organic compounds, drugs, pesticides and environmental contaminants. The data file also contains six simple descriptors and the predictions of Delaney's four-descriptor equation. We used both as reference points. FreeSolv [16] provides experimental hydration free energies (kcal/mol) for 642 small molecules.

We did not curate the data beyond parsing. Neither set contains multi-component SMILES. The 1128 ESOL records correspond to 1117 distinct structures. Repeated records always share a group. Merging them into one record with the mean target value changes the results only slightly: the combined panel then has 628 groups, 55.7% degenerate molecules, an upper bound of 0.832 (0.833 with the repeats) and a best group-aware R^2 of 0.451 (0.460).

IV. METHOD

A. Degeneracy Groups

Two molecules were placed in the same degeneracy group when their descriptor vectors, rounded to six decimals, were identical. The indices are fixed functions of integer edge counts, so molecules with the same edge partitions receive the same values up to floating-point error, which the rounding removes. Grouping by exact equality of the edge partitions, without any rounding, gives exactly the same groups, and so does rounding to 3, 4, 8 or 10 decimals. We call a molecule *degenerate* if it shares its vector with at least one other molecule. Groups were formed separately for three panels: M (10 indices), NM (10 indices) and *Combined*, which joins the two (20 indices). The number of groups directly measures how many molecules a panel can tell apart.

B. The Upper Bound

Let $C_1, \ldots, C_K$ be the groups formed by a panel and y_i the measured target. Any model gives the same prediction to all molecules in a group, so the best possible prediction for group C_k is its mean, $\bar{y}_{Ck}$. The residual sum of squares of any model is therefore at least the within-group sum of squares, which gives

$$R^2_{\max} = 1 - \frac{\sum_k \sum_{i \in C_k} \left(y_i - \bar{y}_{C_k}\right)^2}{\sum_i (y_i - \bar{y})^2}, \tag{2}$$

$$\mathrm{RMSE}_{\min} = \sqrt{\frac{1}{n} \sum_k \sum_{i \in C_k} \left(y_i - \bar{y}_{C_k}\right)^2}. \tag{3}$$

Here $\bar{y}$ is the overall mean and n is the number of molecules. Both quantities depend only on the descriptors and the measured values. No model is fitted and no parameter is tuned. Equation (2) is the correlation ratio η^2, the share of variance explained by the groups in a one-way analysis of variance. The formula itself is standard. Our contribution is to use it as a check on a descriptor set.

The bound holds strictly for any single model applied to the whole data set. It also holds strictly for the pooled out-of-fold predictions of the combined panel under the group-aware scheme described below, because all members of a group fall into the same test fold and are predicted by the same model. Under random folds, members of one group can be predicted by different models, and fold-averaged R^2 uses a different variance in each fold. In that case the upper bound serves as a reference value rather than a strict limit. The bound is also optimistic, since it counts only exact matches. Molecules with similar but not identical vectors do not lower the bound, but they make it harder for a model trained on limited data to reach it.

C. Models and Protocols

We used three regression models: ridge regression, random forest (RF) [17] and gradient boosting (XGBoost, [18]). Table I lists their settings and the software, including RDKit [19] and scikit-learn [20].

Each panel was evaluated with two five-fold schemes. The first split the molecules at random, as is common practice. The second used the degeneracy groups of the combined panel as the unit of splitting for every panel, so all panels were evaluated on the same

ISBN: 978-605-80193-6-2

folds and molecules with the same combined vector were never divided between the training and test sets. Groups were assigned to folds deterministically, from the largest to the smallest; with this assignment, all results were identical in two environments with different Python, NumPy and scikit-learn versions. R^2 was computed from the pooled out-of-fold predictions. The difference between the two schemes shows how much shared feature vectors inflate the usual estimates.

TABLE I. Model settings, validation protocol and software. All runs used the random seed 42; settings not listed are library defaults.

Item	Setting
Ridge	Standardized features; α from 13 log-spaced values in $[10^{-3}, 10^{3}]$ by internal cross-validation
RF	300 trees
XGBoost	400 trees, maximum depth 4, learning rate 0.05
Folds	Five; random and group-aware (combined-panel groups, assigned by size)
Software	Python 3.13.16, RDKit 2026.09.1, scikit-learn 1.6.1, XGBoost 3.4.1, NumPy 2.1.3, SciPy 1.16.3, scikit-posthocs 0.17.1

D. Statistical Comparison and Validation

We compared panels and algorithms with the two-step non-parametric procedure commonly used in benchmark studies [21]. Fold-level R^2 values from the group-aware scheme were arranged into 30 blocks. For the panel comparison, a block was one combination of data set, algorithm and fold; for the algorithm comparison, one combination of data set, panel and fold. When the Friedman test was significant, it was followed by the post-hoc Nemenyi test [22]. All tests used a significance level of 0.05.

We also carried out two checks based on the validation principles of the Organisation for Economic Co-operation and Development (OECD) [23]. Chance correlation was tested by *y*-randomization. The target values were shuffled 100 times and the whole group-aware procedure was repeated each time, so the smallest possible *p*-value is 1/101 = 0.0099. The applicability domain was assessed with Williams plots. The leverage of molecule i is $h_i = \mathbf{x}_i^{\mathrm{T}}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{+}\mathbf{x}_i$, where $\mathbf{x}_i$ is its descriptor vector and $\mathbf{X}$ is the descriptor matrix. The leverage threshold is $h^{*} = 3(p + 1)/n$, where p is the number of descriptors, and the limits for standardized residuals are ±3.

V. RESULTS

A. How Much the Descriptors Cannot Separate

Table II gives, for each panel, the number of groups, the number of degenerate molecules and the resulting bounds.

TABLE II. Number of groups and resulting bounds for each descriptor panel. A group is a set of molecules with an identical feature vector. $R^2_{\max}$ and $\mathrm{RMSE}_{\min}$ follow from (2) and hold for any model. RMSE is in log mol/L for ESOL and kcal/mol for FreeSolv. Combined joins the M and NM indices (20 indices).

Panel	Groups	Degenerate	$R^2_{\max}$	$\mathrm{RMSE}_{\min}$
ESOL (n = 1128)				
M	558	739 (65.5%)	0.811	0.912
NM	628	638 (56.6%)	0.833	0.856
Combined	628	638 (56.6%)	0.833	0.856
Delaney 6	919	328 (29.1%)	0.995	0.150
FreeSolv (n = 642)				
M	216	503 (78.3%)	0.552	2.574
NM	235	480 (74.8%)	0.563	2.542
Combined	236	479 (74.6%)	0.565	2.536

On ESOL, the combined panel divides the 1128 molecules into 628 groups, and 638 molecules (56.6%) are degenerate. FreeSolv contains smaller molecules. There the panel forms only 236 groups, and 74.6% of the molecules are degenerate.

Fig. 1 shows why. The four compounds in the figure have identical values for all twenty indices, yet their solubilities differ by 5.7 log units. They are an alkane, an alcohol, a ketone and a urea, so they differ in atom type, bond order and hydrogen count. A degree-based description of the graph without hydrogens does not see any of these differences. The group with the second-largest spread in solubility shows the same effect. Lindane, pentachlorophenol, hexamethylbenzene and *d*-inositol each consist of a six-membered ring with six substituents. They share one vector, but their solubilities range from −5.23 to +0.35. These cases are not positional isomers; they follow from the degree-based description itself.

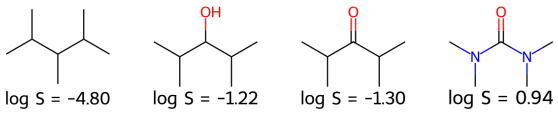


Fig. 1. Four ESOL compounds with identical values for all 20 descriptors (from left: 2,3,4-trimethylpentane, 2,4-dimethyl-3-pentanol, 2,4-dimethyl-3-pentanone and tetramethylurea). Their measured solubilities differ by 5.7 log units, yet any model must give them the same prediction.

B. What the Choice of Polynomial Does and Does Not Change

The choice of polynomial makes a measurable difference. The NM family separates more molecules than the M family: 628 groups against 558 on ESOL, and 235 against 216 on FreeSolv. This confirms by direct counting a point that is usually argued in words: the neighborhood degree sum distinguishes vertices more finely than the degree alone.

The difference is small, however. Both families leave more than half of ESOL and three quarters of FreeSolv degenerate, and their upper bounds differ by at most 0.022. Combining the two families adds almost nothing to NM. On ESOL the combined panel forms 628 groups, the same number as the ten NM indices alone; on FreeSolv it forms 236 groups against 235. Once the NM indices are present, the M indices separate practically no additional molecules.

Separating more molecules does not translate into clearly better predictions. NM separates more molecules than M, yet its mean rank is slightly worse, and the combined panel ranks best; none of these differences is significant (Table IV). The number of groups and the achieved accuracy should therefore be reported separately.

C. Achieved Performance Against the Upper Bound

Table III and Fig. 2 compare the fitted models with the upper bound in (2).

TABLE III. Five-fold cross-validated R^2 for each panel (best of ridge, RF and XGBoost; pooled out-of-fold predictions). *Random*: molecules are split at random. *Group*: molecules with the same

ISBN: 978-605-80193-6-2

combined vector stay in the same fold. Δ is Random minus Group. The last column repeats the upper bound from Table II.

Panel	Random	Group	Δ	R^2_{max}
ESOL (n = 1128)				
M	0.486	0.434	+0.052	0.811
NM	0.465	0.444	+0.021	0.833
Combined	0.485	0.460	+0.025	0.833
Delaney 6	0.839	0.842	−0.003	0.995
FreeSolv (n = 642)				
M	0.164	0.157	+0.007	0.552
NM	0.143	0.122	+0.021	0.563
Combined	0.184	0.174	+0.010	0.565

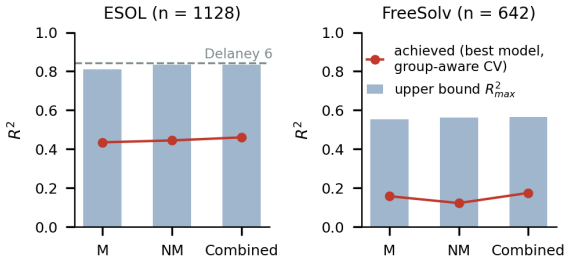


Fig. 2. Upper bound (bars) and achieved performance (markers; best of three models under group-aware cross-validation). The dashed line shows the Delaney 6 reference, which exceeds the upper bound of the combined panel.

On ESOL the combined panel reaches R^2 = 0.460, while its upper bound is 0.833. The descriptors are therefore limited in two ways. The upper bound is well below 1, and the models stay far below it. On FreeSolv the gap is even larger, 0.174 against 0.565.

Two reference points make the comparison clearer. The ESOL file includes six simple descriptors: minimum degree, molecular weight, hydrogen-bond donors, rings, rotatable bonds and polar surface area (*Delaney 6* in the tables). Molecular weight and polar surface area carry information about atom type. With the same group-aware scheme, these six descriptors give R^2 = 0.842. Delaney's original four-descriptor equation uses clogP (the calculated octanol–water partition coefficient), molecular weight, rotatable bonds and aromatic proportion. Without any fitting on our part, it gives R^2 = 0.811. The six descriptors thus exceed the upper bound of the combined panel, and the equation comes close to it. The six-descriptor panel also has far fewer degenerate molecules, with 919 groups and an upper bound of 0.995.

Fig. 3 shows the same result for individual molecules. Predictions from the combined panel form horizontal bands, in which molecules with different measured values receive the same prediction. The reference panel shows no such bands.

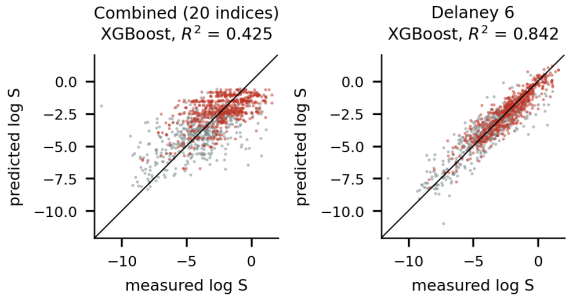


Fig. 3. Out-of-fold predictions on ESOL under group-aware cross-validation, XGBoost in both panels. Red: molecules that are degenerate under the combined panel; gray: unique molecules. The horizontal bands in the left panel are molecules with different measured values that receive the same prediction.

D. Statistical Comparison

At the panel level the Friedman test is not significant (χ^2_F = 4.27, p = 0.118; Table IV), so the three panels cannot be separated. With a different but equally valid assignment of groups to folds, the same test was significant (p = 0.003), which shows how fragile such differences are. At the algorithm level the test is significant (χ^2_F = 12.07, p = 0.0024): RF is better than ridge (p = 0.002) and XGBoost (p = 0.038), which cannot be separated from each other. Both kinds of difference are small compared with the gap to the upper bound. Because folds of the same data set are not independent, these p-values are descriptive.

TABLE IV. Friedman tests on fold-level R^2 at two levels (30 blocks: data set × algorithm × fold for panels; data set × panel × fold for algorithms). A lower mean rank is better. The values below the ranks are post-hoc Nemenyi p-values; bold values are significant at the 0.05 level. The panel-level test is not significant.

Panel level: χ^2_F = 4.27, p = 0.118			
	M	**NM**	**Combined**
Mean rank	2.00	2.27	1.73
M	—	0.556	0.556
NM	0.556	—	0.097
Combined	0.556	0.097	—
Algorithm level: χ^2_F = 12.07, p = 0.0024			
	Ridge	**RF**	**XGBoost**
Mean rank	2.37	1.50	2.13
Ridge	—	**0.002**	0.638
RF	**0.002**	—	**0.038**
XGBoost	0.638	**0.038**	—

E. Chance Correlation and Applicability Domain

For the best combined-panel model on each data set, all 100 shuffled runs scored below the real run. The empirical p-value is therefore at its minimum of 0.0099 for both benchmarks. The mean R^2 of the shuffled runs was −0.115 on ESOL and −0.109 on FreeSolv. The modest accuracy reported above is real and not due to chance.

Williams plots (Fig. 4) show that 95.5% of the molecules in both data sets lie inside the applicability domain. There are 44 and 20 structural outliers, respectively, and at most nine response outliers. The degenerate molecules are spread across the domain and are not concentrated at its edge. This suggests that their errors come from the descriptors rather than from extrapolation.

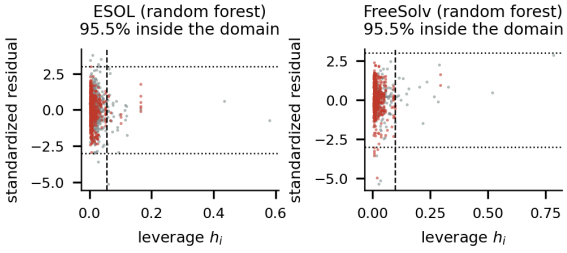


Fig. 4. Williams plots for the best combined-panel model. Dashed line: leverage threshold $h^* = 3(p + 1)/n$ (p: number of descriptors; n: number of molecules). Dotted lines: standardized residuals of ±3. Red: molecules that are degenerate under the combined panel; gray: unique molecules.

F. The Splitting Protocol

Switching from random folds to group-aware folds lowers the R^2 of the combined panel by 0.025 on ESOL and 0.010 on FreeSolv. On ESOL, the reference panel

ISBN: 978-605-80193-6-2

does not lose accuracy (−0.003). Random splitting does inflate the estimates when molecules share feature vectors, but only slightly. On these benchmarks the main problem is the limited information in the descriptors. Leakage between the training and test sets plays a much smaller role.

VI. DISCUSSION

The main result concerns what a family of descriptors can represent. Degree-based polynomial indices describe how atoms are connected and how the molecule branches. They do not describe which element sits at each vertex or how it is bonded. For properties that depend on polarity, hydrogen bonding or electrostatics, such as solubility and hydration free energy, this missing information is important. The six simple descriptors in the ESOL file, some of which reflect atom type, form 919 groups. All twenty connectivity indices together form only 628.

The choice of polynomial moves the upper bound only by hundredths. The bound itself is set earlier, when the molecule is represented as a graph without hydrogens whose vertices carry degrees instead of atom types. Our analysis therefore also applies to other indices built on this representation, including the classical Wiener and Randić indices.

These results have two practical consequences. First, reported performance should be compared with the upper bound of the descriptor set, not with the range of published R^2 values. A study that reports $R^2 = 0.95$ on twenty similar compounds is not necessarily overfitting. However, the result can only be interpreted once we know how many of those compounds the descriptors can tell apart. Equation (2) is cheap to compute and could be reported routinely with the fit.

Second, combining polynomial families should not be assumed to add information. On ESOL, the two families together separate exactly as many molecules as the NM family alone. Whether this holds for other chemical spaces is an open question that deserves attention.

A. Limitations

Two benchmarks and two physicochemical properties cannot settle the general case. Properties that depend mainly on size and shape should be easier for these descriptors, and we expect higher upper bounds for them. The bound counts only exact matches and is therefore optimistic. The Friedman and Nemenyi tests in Section V-D treat folds of the same data set as independent blocks, so their p-values are indicative only. Finally, we used a small set of models with standard settings and no further tuning. This does not affect the main argument, because no amount of tuning can raise the upper bound.

B. Future Work

We see three directions for future work. The first is to apply the same analysis to the complement-graph families, CoM and CoNM, which describe pairs of non-adjacent atoms and may separate different molecules. The second is to replace exact equality with a tolerance, so that the upper bound becomes a curve that depends on how close two vectors must be. The third, and the most important, is to develop polynomial indices that include atom type, for example by weighting vertices with atomic properties instead of degree alone.

VII. CONCLUSION

When computed automatically for structurally diverse data sets, the M- and NM-polynomial indices leave 56.6% of the ESOL molecules and 74.6% of the FreeSolv molecules degenerate. The resulting model-independent upper bounds are $R^2_{max} = 0.833$ and 0.565. The fitted models reach about half of the bound on ESOL and about a third of the bound on FreeSolv. A four-descriptor equation from 2004 already comes close to the bound of the combined panel. The limitation comes from the representation, because these descriptors do not see atom type or bond order. *Y*-randomization and the applicability domain analysis support these results. We suggest that quantitative structure–property relationship (QSPR) studies based on polynomial indices report the degeneracy statistics and the resulting upper bound as standard checks, in the same way that applicability domain and chance correlation tests are reported today.

The check needs one pass over the feature matrix and does not depend on the model, the tuning or the data split. Any engineering study that predicts from a fixed set of discrete features can report how many distinct observations the features can separate and compare its accuracy with this bound. If the upper bound itself is low, better models will not help, and a better representation is needed.

Data Availability

Both data sets are public: ESOL [14] through MoleculeNet [15] (https://moleculenet.org; file delaney-processed.csv in https://github.com/deepchem/deepchem) and FreeSolv [16] (https://github.com/MobleyLab/FreeSolv).

References

[1] Wiener, H. (1947). Structural determination of paraffin boiling points. *Journal of the American Chemical Society, 69*(1), 17–20. https://doi.org/10.1021/ja01193a005

[2] Randić, M. (1975). Characterization of molecular branching. *Journal of the American Chemical Society, 97*(23), 6609–6615. https://doi.org/10.1021/ja00856a001

[3] Deutsch, E., & Klavžar, S. (2015). M-polynomial and degree-based topological indices. *Iranian Journal of Mathematical Chemistry, 6*(2), 93–102. https://doi.org/10.22052/ijmc.2015.10106

[4] Verma, A., Mondal, S., De, N., & Pal, A. (2019). Topological properties of bismuth tri-iodide using neighborhood M-polynomial. *International Journal of Mathematics Trends and Technology, 67*, 83–90.

[5] Mondal, S., Siddiqui, M. K., De, N., & Pal, A. (2021). Neighborhood M-polynomial of crystallographic structures. *Biointerface Research*

 ISBN: 978-605-80193-6-2

in Applied Chemistry, 11(2), 9372–9381. https://doi.org/10.33263/BRIAC112.93729381

[6] Kirmani, S. A. K., & Ali, P. (2022). CoM-polynomial and topological coindices of hyaluronic acid conjugates. *Arabian Journal of Chemistry, 15*(7), Article 103911. https://doi.org/10.1016/j.arabjc.2022.103911

[7] Öztürk Sözen, E., & Eryaşar, E. (2024). QSPR analysis of some drug candidates investigated for COVID-19 via new topological coindices. *Polycyclic Aromatic Compounds, 44*(2), 1291–1308. https://doi.org/10.1080/10406638.2023.2191974

[8] Havare, Ö. Ç. (2021). Topological indices and QSPR modeling of some novel drugs used in the cancer treatment. *International Journal of Quantum Chemistry, 121*(24), Article e26813. https://doi.org/10.1002/qua.26813

[9] Dehmer, M., Grabner, M., & Varmuza, K. (2012). Information indices with high discriminative power for graphs. *PLoS ONE, 7*(2), Article e31214. https://doi.org/10.1371/journal.pone.0031214

[10] Dehmer, M., Grabner, M., & Furtula, B. (2012). Structural discrimination of networks by using distance, degree and eigenvalue-based measures. *PLoS ONE, 7*(7), Article e38564. https://doi.org/10.1371/journal.pone.0038564

[11] Brown, S. P., Muchmore, S. W., & Hajduk, P. J. (2009). Healthy skepticism: Assessing realistic model performance. *Drug Discovery Today, 14*(7–8), 420–427. https://doi.org/10.1016/j.drudis.2009.01.012

[12] Golbraikh, A., Muratov, E., Fourches, D., & Tropsha, A. (2014). Data set modelability by QSAR. *Journal of Chemical Information and Modeling, 54*(1), 1–4. https://doi.org/10.1021/ci400572x

[13] Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. https://doi.org/10.1007/978-0-387-84858-7

[14] Delaney, J. S. (2004). ESOL: Estimating aqueous solubility directly from molecular structure. *Journal of Chemical Information and Computer Sciences, 44*(3), 1000–1005. https://doi.org/10.1021/ci034243x

[15] Wu, Z., Ramsundar, B., Feinberg, E. N., Gomes, J., Geniesse, C., Pappu, A. S., Leswing, K., & Pande, V. (2018). MoleculeNet: A benchmark for molecular machine learning. *Chemical Science, 9*(2), 513–530. https://doi.org/10.1039/C7SC02664A

[16] Mobley, D. L., & Guthrie, J. P. (2014). FreeSolv: A database of experimental and calculated hydration free energies, with input files. *Journal of Computer-Aided Molecular Design, 28*(7), 711–720. https://doi.org/10.1007/s10822-014-9747-x

[17] Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32. https://doi.org/10.1023/A:1010933404324

[18] Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM. https://doi.org/10.1145/2939672.2939785

[19] Landrum, G., & RDKit Contributors. (2026). *RDKit: Open-source cheminformatics software* [Computer software]. https://www.rdkit.org

[20] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research, 12*, 2825–2830.

[21] Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research, 7*, 1–30.

[22] Nemenyi, P. B. (1963). *Distribution-free multiple comparisons* [Doctoral dissertation, Princeton University].

[23] OECD. (2007). *Guidance document on the validation of (quantitative) structure–activity relationship [(Q)SAR] models* (Series on Testing and Assessment No. 69).

ISBN: 978-605-80193-6-2

BI-RADS Risk Group Classification Using Four-View Mammogram Fusion and Deep Learning

Güliz Toz[1]*, Asiye Sözeri[2], Nisanur Erdal[1], Merve Yazıcı[1]

[1]Department of Computer Engineering, Karamanoğlu Mehmetbey University, Karaman, Türkiye
[2]Department of Radiology, Karabük University Research and Training Hospital, Karabük, Türkiye

* Corresponding author e-mail: guliztoz@kmu.edu.tr

***Abstract*—Early detection of breast cancer is important for timely treatment. Radiologists assess four standard mammographic views together to examine both breasts. Following this approach, this study evaluated a composite image formed by combining four mammograms from the same patient for Breast Imaging Reporting and Data System risk-group classification. The dataset was provided by the Turkish Ministry of Health for the 2023 TEKNOFEST Artificial Intelligence in Healthcare Competition. Cases with an incomplete assessment were excluded, and the task distinguished the low-risk group, categories one and two, from the high-risk group, categories four and five. Preprocessing included intensity inversion, thresholding, morphological operations, and masking to retain the breast region. The four views were then combined and resized to the model input dimensions. ResNet-50 and EfficientNet-B3 were trained using the same composite inputs and training hyperparameters. Performance was evaluated using accuracy, precision, recall, and the F1 score. Both models achieved 83 percent test accuracy and a macro F1 score of 0.83. ResNet-50 showed higher recall for the low-risk group and higher precision for the high-risk group. EfficientNet-B3 provided a more balanced precision and recall profile and higher recall for the high-risk group. These findings indicate that four-view composite images can be used for risk-group classification, with similar overall performance but different class-specific results across the two architectures.**

***Keywords*—Mammography, four-view fusion, BI-RADS, deep learning, ResNet-50, EfficientNet-B3.**

I. INTRODUCTION

Breast cancer is an important health concern among women, and early detection supports timely treatment [1]. Mammography is widely used for screening, although dense breast tissue can make suspicious findings difficult to identify [2].

The Breast Imaging Reporting and Data System (BI-RADS) standardizes the reporting of imaging findings and subsequent management. BI-RADS 1 and 2 indicate negative or benign findings, whereas BI-RADS 4 and 5 indicate suspicious or highly suspicious findings. These categories are imaging assessments rather than pathology-confirmed cancer labels [3], [4].

Mammographic classification has been studied using both handcrafted features and deep learning. Chougrada et al. evaluated deep convolutional neural networks for screening [5]. Mutlu et al. extracted lesion features and classified benign and malignant lesions using conventional classifiers [6]. Transfer learning has also been used for breast cancer detection [7].

Multi-view approaches extend this analysis beyond individual images. Petrini et al. trained an EfficientNet-based patch classifier, transferred its weights to a single-view classifier, and then trained a two-view model [8]. Karagöz et al. investigated a deep learning model that processed multiple mammographic views separately and evaluated them together [9].

Radiologists jointly examine left and right craniocaudal (CC) and mediolateral oblique (MLO) views. Following this approach, we combined four views from the same patient into one image and evaluated ResNet-50 and EfficientNet-B3 for BI-RADS 1–2 versus BI-RADS 4–5 classification. This representation presents bilateral and projection-specific information to a convolutional neural network (CNN) through a single input.

II. MATERIALS AND METHODS

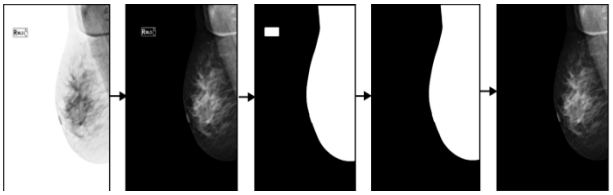
A. Dataset

The mammography dataset was prepared for the 2023 TEKNOFEST Artificial Intelligence in Healthcare Competition with support from the General Directorate of Health Information Systems of the Turkish Ministry of Health [10]. Images were provided in Digital Imaging and Communications in Medicine (DICOM) format and labeled as BI-RADS 0, 1–2, or 4–5. BI-RADS 0 cases were excluded because they required additional imaging or comparison with previous examinations. The remaining cases were grouped as low risk (1–2) and high risk (4–5).

B. Preprocessing and Four-View Fusion

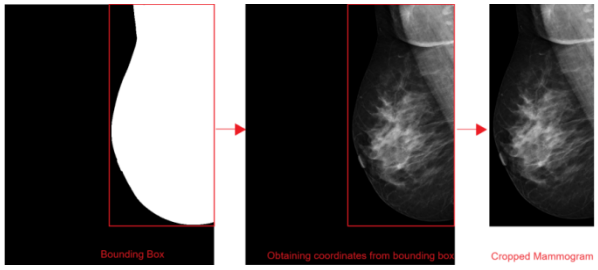
Image intensities were inverted and normalized to the 0–1 range. Binary masks were obtained using a threshold of 0.02, followed by morphological operations to fill gaps and remove small noisy regions. The largest connected component was retained as the breast region. Pixel-wise multiplication of the mask and mammogram suppressed labels and background. The preprocessing stages are shown in Fig. 1.

ISBN: 978-605-80193-6-2

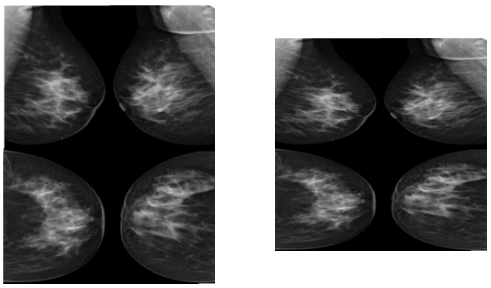


(a) (b) (c) (d) (e)

Fig. 1. Preprocessing: (a) original image, (b) intensity inversion, (c) thresholding, (d) area filtering, and (e) masking.



(a)



(b) (c)

Fig. 2. Four-view fusion: (a) cropping, (b) combination of left MLO, left CC, right MLO, and right CC views, and (c) resizing.

The breast regions were cropped, and left CC, right CC, left MLO, and right MLO images from the same examination were combined. The composite image was resized to 224 × 224 pixels for both models. Fig. 2 shows this process.

C. Data Augmentation

Horizontal flipping, small rotations, and minor scaling transformations were applied during training to improve generalization and reduce overfitting.

D. Deep Learning Models

ResNet-50 and EfficientNet-B3 were evaluated using the same four-view composite images. ResNet-50 uses residual connections to facilitate the training of deep networks and bottleneck blocks to reduce computational complexity [11]. Its classification head included global average pooling, dropout, and a fully connected layer. EfficientNet-B3 uses compound scaling to balance network depth, width, and input resolution and employs MBConv blocks [12]. Global average pooling and a fully connected layer were used for classification. Both models were trained with the same input images and hyperparameters to compare their classification performance.

III. EXPERIMENTAL SETUP

Both models were trained with Adamax and categorical cross-entropy using the settings in Table I. Validation results were monitored during training. Performance on a separate test set was assessed using accuracy, precision, recall, and F1 score.

TABLE I. Training Settings and Model Parameters

Hyper Parameters for Training		
Models	**ResNet-50**	**EfficientNet-B3**
Batch Size	20	20
Validation split	0.2	0.2
Epochs	100	100
Learning Rate	0.001	0.001
Loss Function	Categorical Cross-Entropy	Categorical Cross-Entropy
Optimizer	Adamax	Adamax
Total params	23,591,810	11,734,961
Trainable params	23,538,690	11,647,658
Non-trainable params	53,120	87,303

As shown in Table I, both models were trained with a batch size of 20 for 100 epochs, using a learning rate of 0.001 and a validation split of 0.2. EfficientNet-B3 had fewer total and trainable parameters than ResNet-50. Both architectures were evaluated using the same training settings.

IV. RESULTS AND DISCUSSION

Table II presents the test results. Both models achieved 83% accuracy and macro and weighted F1 scores of 0.83. ResNet-50 showed higher recall for BI-RADS 1–2 and higher precision for BI-RADS 4–5. EfficientNet-B3 had a more balanced class-wise profile and higher recall for BI-RADS 4–5.

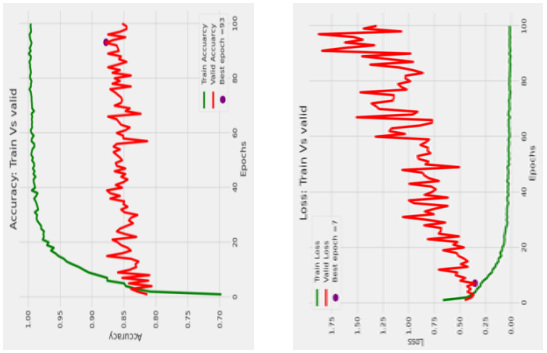
TABLE II. Classification Performance

Models	**ResNet-50**				**EfficientNet-B3**			
Evaluation Metrics	**Precision**	**Recall**	**F1-Score**	**Support Proportion**	**Precision**	**Recall**	**F1-Score**	**Support Proportion**
BI-RADS 1-2	0.79	**0.91**	**0.84**	0.4982	0.82	**0.84**	**0.83**	0.4982
BI-RADS 4-5	**0.88**	0.75	0.81	0.5018	**0.84**	0.81	0.82	0.5018
Macro Average	0.84	0.83	0.83		0.83	0.83	0.83	

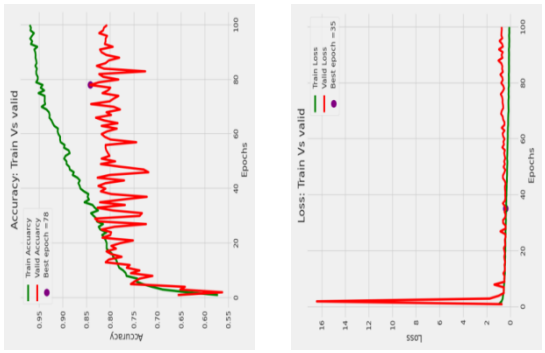
ISBN: 978-605-80193-6-2

Weighted Average	0.84	0.83	0.83		0.83	0.83	0.83	
Overall Accuracy	**0.83**				**0.83**			

Fig. 3 presents the training and validation accuracy and loss curves for both models. EfficientNet-B3 converged more quickly and maintained a more stable validation loss, whereas ResNet-50 showed greater fluctuations during training.

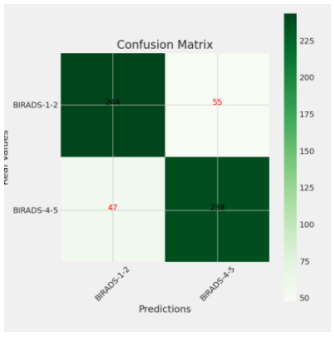


(a)

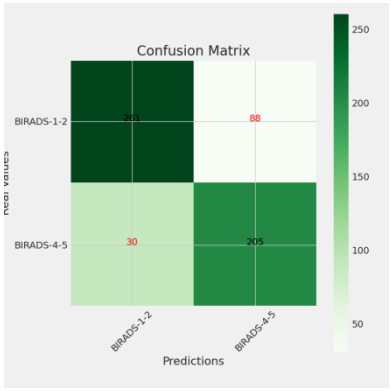


(b)

Fig. 3. Training and validation accuracy and loss curves for (a) EfficientNet-B3 and (b) ResNet-50.



(a)



(b)

Fig. 4. Confusion Matrix for (a) EfficientNet-B3 and (b) ResNet-50 models

Fig. 4 presents the confusion matrices for the two models. EfficientNet-B3 achieved higher recall for BI-RADS 4–5, correctly identifying a greater proportion of high-risk cases. ResNet-50 achieved higher recall for BI-RADS 1–2. These results are consistent with the class-wise metrics in Table II and show the different classification profiles of the two architectures.

Table III summarizes related studies. The relatively recent dataset and different classification targets make direct comparison difficult. Whereas earlier studies examined benign–malignant classification or individual BI-RADS categories, this study classified BI-RADS 1–2 and 4–5 using four views in one image. Despite these differences, the reported accuracy was close to that of some studies in the literature.

TABLE III. Comparison with Related Studies

Studies	Used Datasets	Methods	Accuracy (%)
Das et al. (2023)[13]	DDSM	CNN	80.4
Trang et al. (2023) [14]	Private	Xception+ GBM	84.5
Alruwaili et al. (2022)[7]	MIAS	ResNet-50	89.5
Das et al. (2024)[15]	Kaggle Database	ResNet-50	92.01
Rahman et al. (2023)[16]	INbreast	ResNet-50	93
Nasra et al. (2024) [17]	BreakHis	EfficientNet	94
Pillai et al. (2023)[18]	Breast Ultrasound Images Dataset	EfficientNet-B3	89.34
Proposed (2026)	MammosighTR	ResNet-50 EfficientNet-B3	83 83

Table III places the results alongside related studies using different datasets and methods. Both models achieved 83% accuracy with the four-view composite images, a value close to those reported in some previous studies. The proposed approach combines bilateral CC and MLO views in a single input for BI-RADS risk-group classification. Future work may investigate ultrasound or magnetic resonance imaging alongside mammography and explore self-supervised learning to improve classification and generalization.

ISBN: 978-605-80193-6-2

V. CONCLUSION

Four mammograms from the same patient were combined into a single input for BI-RADS 1–2 versus 4–5 classification. ResNet-50 and EfficientNet-B3 both achieved 83% accuracy and a macro F1 score of 0.83. ResNet-50 provided higher low-risk recall, whereas EfficientNet-B3 showed more balanced results across the risk groups. The findings support the use of composite mammograms for this classification task.

Acknowledgment

This work was supported by TÜBİTAK under the TÜBİTAK 2209 project 1919B012311128. The authors thank the Turkish Ministry of Health for providing the dataset.

References

[1] World Health Organization, "Breast cancer," July 3, 2026. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/breast-cancer. [Accessed: Oct. 6, 2026].

[2] National Cancer Institute, "Mammograms," updated Dec. 2, 2025. [Online]. Available: https://www.cancer.gov/types/breast/screening/mammograms. [Accessed: Oct. 6, 2026].

[3] J. H. R. Nascimento, V. D. Silva, and A. C. Maciel, "Accuracy of mammographic findings in breast cancer: correlation between BI-RADS classification and histological findings," Radiologia Brasileira, vol. 43, pp. 91–96, 2010.

[4] L. Luo et al., "Deep learning in breast cancer imaging: A decade of progress and future directions," IEEE Reviews in Biomedical Engineering, 2024.

[5] H. Chougrada, H. Zouakia, and O. Alheyane, "Deep convolutional neural networks for breast cancer screening," Computer Methods and Programs in Biomedicine, vol. 157, pp. 19–30, 2018.

[6] F. Mutlu, G. Çetinel, and S. Gül, "A fully-automated computer-aided breast lesion detection and classification system," Biomedical Signal Processing and Control, vol. 62, art. 102157, 2020.

[7] M. Alruwaili and W. Gouda, "Automated breast cancer detection models based on transfer learning," Sensors, vol. 22, no. 3, art. 876, 2022, doi: 10.3390/s22030876.

[8] D. G. P. Petrini et al., "Breast cancer diagnosis in two-view mammography using end-to-end trained efficientnet-based convolutional network," IEEE Access, vol. 10, pp. 77723–77731, 2022.

[9] M. Karagöz et al., "Deep learning-based breast cancer diagnosis with multiview mammography," Turkish Journal of Electrical Engineering and Computer Sciences, vol. 32, no. 3, pp. 382–402, 2024.

[10] U. Koç et al., "MammosighTR: Nationwide breast cancer screening mammogram dataset with BI-RADS annotations for artificial intelligence applications," Radiology: Artificial Intelligence, 2025.

[11] K. He et al., "Deep residual learning for image recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016.

[12] M. Tan and Q. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in Proceedings of the International Conference on Machine Learning, PMLR, 2019.

[13] H. S. Das, A. Das, A. Neog, S. Mallik, K. Bora, and Z. Zhao, "Breast cancer detection: Shallow convolutional neural network against deep convolutional neural networks based approach," Frontiers in Genetics, vol. 13, art. 1097207, 2023.

[14] N. T. H. Trang, K. Q. Long, P. L. An, and T. N. Dang, "Development of an artificial intelligence-based breast cancer detection model by combining mammograms and medical health records," Diagnostics, vol. 13, no. 3, art. 346, 2023.

[15] T. Das et al., "ResNet-50: The deep networks for automated breast cancer classification using MR images," in 2024 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC), IEEE, 2024.

[16] H. Rahman et al., "Efficient breast cancer diagnosis from complex mammographic images using deep convolutional neural network," Computational Intelligence and Neuroscience, vol. 2023, art. 7717712, 2023.

[17] P. Nasra, S. Gupta, and G. R. Kumar, "Leveraging EfficientNetB3 for accurate breast cancer classification: Insights and innovations," in 2024 9th International Conference on Communication and Electronics Systems (ICCES), pp. 1473–1478, 2024, doi: 10.1109/ICCES63552.2024.10859795.

[18] R. Pillai, N. Sharma, R. Chauhan, D. Rawat, and R. Gupta, "A deep learning approach for breast tumour diagnosis & classification from fine-tuned EfficientNetB3 model," in 2023 Global Conference on Information Technologies and Communications (GCITC), pp. 1–5, 2023, doi: 10.1109/GCITC60406.2023.10426155.

ISBN: 978-605-80193-6-2

Lightweight Convolutional Neural Networks for Wheat Yellow Rust Severity Classification: An Accuracy and Efficiency Benchmark

Hüseyin ELDEM[1] Alper DEMİR[1]

[1]Karamanoğlu Mehmetbey University, Kamil Özdağ Science Faculty, Software Development Department, Karaman, Turkey

* Corresponding author e-mail: heldem@kmu.edu.tr

***Abstract*—Yellow rust is a serious disease of wheat. Its severity on individual leaves is still mostly scored by trained experts by looking at the leaves. Deep learning can help automate this work, but many published models are large and are tested on augmented image sets. This can make them less practical for low-cost field devices and can also make the reported accuracy less reliable. In this study, five lightweight convolutional neural networks, MobileNetV3-Small, ShuffleNetV2, MobileNetV2, MobileNetV3-Large and EfficientNet-B0, were compared with ResNet-50 for six-level yellow rust severity classification. We used the raw images from the public Yellow-Rust-19 dataset. After duplicate files were removed, 5,417 leaf images were distributed among the training, validation, and test subsets through stratified sampling. Each model was initialized with ImageNet-pretrained weights and subsequently fine-tuned under an identical training protocol using three different random seeds. We compared accuracy, macro F1, quadratic weighted kappa, parameter count, computational cost and CPU time for one image. All models got between 70.8 and 72.7% accuracy and between 64 and 68% macro F1. The much larger ResNet-50 did not perform better than the lightweight models. MobileNetV3-Small reached 67.4% macro F1 with 1.5 million parameters and a latency of 9.1 ms per image. This result was close to the best model, EfficientNet-B0, but MobileNetV3-Small was 3.7 times faster than EfficientNet-B0 and 12 times faster than ResNet-50. More than 90% of the errors were between neighbouring severity levels. Grad-CAM maps also showed that the models were mostly using lesion areas for their decisions. Overall, the results suggest that small networks can be enough for leaf-level yellow rust screening on mobile hardware. The harder problem is still telling apart severity levels that are very close to each other.**

***Keywords*—Yellow rust, disease severity classification, deep learning, lightweight convolutional neural networks, MobileNetV3**

I. INTRODUCTION

Wheat is one of the most important food crops in the world. Pathogens and pests are estimated to reduce the yield that could be achieved from major crops by about one fifth globally [1]. Stripe rust, also known as yellow rust, is considered one of the major diseases affecting wheat production. It is caused by Puccinia striiformis f. sp. tritici and can spread quickly by wind-borne spores when conditions are cool and humid. In susceptible wheat varieties, it can cause serious yield losses [2]. Monitoring the disease and developing resistant varieties both need reliable severity measurements. Usually, experts do this by looking at each leaf and giving it an infection type from resistant to susceptible. This takes time, needs trained people, and can be subjective, especially when two severity levels look very similar.

Deep convolutional neural networks (CNNs) have shown that they can recognise plant diseases from leaf images with high accuracy [3], [4]. Several studies have also looked at yellow rust. Hayıt et al. developed the Yellow-Rust-19 dataset, which comprises six disease-severity categories determined through expert assessment and reported 91% accuracy with an Xception-based model [5]. A later study using DenseNet-201 features together with texture features reported 92.4% [6]. However, using these results in real-world applications raises two important concerns. One is the size of the models. Most previous studies mainly focus on accuracy, but a system designed for farmers or field workers may need to work on low-cost smartphones or embedded devices. For this reason, factors such as model size and inference time should also be considered. Another concern is the way the dataset was prepared in some of the previous studies. The studies used augmented images, and if images from the same original leaf are included in both the training and test sets, the evaluation results may be higher than expected. In this situation, the model may be tested on images that are very similar to ones it has already seen during training rather than on a completely new leaf. This can lead to data leakage and may not accurately reflect the model's performance in a real-world setting. Such leakage between training and test data has been identified as a widespread source of over-optimistic results across many fields that apply machine learning [7].

To address these two concerns, this study focuses on a controlled comparison of lightweight CNN models. In addition to classification accuracy, the models are compared in terms of their size and inference time. The study also uses a data-splitting approach designed to avoid having augmented versions of the same leaf appear in both the training and test sets. Based on this setup, the main contributions of this study are as follows:

- Five lightweight CNNs were evaluated alongside ResNet-50 as a reference model. All models were trained and evaluated under the same experimental conditions. The experiments were conducted on the original Yellow-Rust-19 images after removing duplicate samples. The dataset was then divided

ISBN: 978-605-80193-6-2

using a stratified split, and each experiment was run three times with different random seeds.

- The models were evaluated using accuracy, macro F1-score, and quadratic weighted kappa. Their computational requirements were also considered by measuring the number of parameters, multiply-accumulate operations (MACs), model size, and CPU latency. These results allow the classification performance of the models to be considered together with their computational cost, which is particularly relevant for mobile and embedded applications.
- In addition to the numerical results, the classification errors were examined to better understand the behavior of the models. Grad-CAM visualizations were used to identify the regions of the leaves that contributed to the model predictions and to examine cases where the models produced incorrect classifications.

II. RELATED WORK

Mohanty et al. [3] and Ferentinos [4] investigated the use of CNN-based approaches for detecting diseases in crops from leaf images. Most of these studies focus on identifying which disease is present. Too et al. [8] fine-tuned several ImageNet-pretrained architectures under a common protocol and showed that the choice of architecture affects identification performance, but their comparison addressed disease identification rather than severity grading. Severity grading is a more detailed task and has received less attention. For wheat yellow rust, Shafi et al. [9] classified three infection types, healthy, resistant and susceptible, from smartphone images using texture features and gradient boosting, and reported 92.3% accuracy. Mi et al. [10] graded wheat stripe rust severity from leaf images captured with mobile devices, using a DenseNet-based network with an attention mechanism. Hayıt et al. [5] created the Yellow-Rust-19 dataset with six infection types and proposed Yellow-Rust-Xception. The same group later combined deep features and texture features for the six-level task [6]. These studies mainly focused on classification accuracy, while model size and inference speed were rarely reported.

Lightweight CNN models have been widely used for applications with limited computational resources. MobileNetV2 [11] is one of these models and uses depthwise separable convolutions and inverted residual blocks to reduce the computational cost. MobileNetV3 [12] is another lightweight architecture designed to improve efficiency on mobile and embedded devices. It includes squeeze-and-excitation modules and the hard-swish activation function, and is available in Small and Large versions. ShuffleNetV2 [13] uses channel splitting and channel shuffling to reduce the computational cost while maintaining good performance. EfficientNet [14] follows a different approach by scaling the depth, width, and input resolution of the network together. In this study, ResNet-50 [15] is also included as a larger model to provide a comparison with the lightweight CNN models. However, relatively few studies have examined these networks for yellow rust severity classification using a strict separation between training and test data while simultaneously evaluating both prediction performance and computational requirements within the same experimental setup.

III. MATERIALS AND METHODS

A. Dataset

The Yellow-Rust-19 dataset [5] consists of wheat leaf images collected from experimental plots at the Field Crops Central Research Institute in Ankara, Türkiye. The plots are used for cereal rust research. Each leaf was evaluated by experts and assigned to one of six infection categories: no disease (0), resistant (R), moderately resistant (MR), moderately resistant to moderately susceptible (MRMS), moderately susceptible (MS), or susceptible (S). Two versions of the dataset are available. One is the raw image set, and the other is a processed version of about 15,000 images that includes augmented copies. In this study, only the raw images were used so that each image represents a different leaf. The leaves are single, long and narrow, and they were photographed on a dark background.

B. Data Preparation and Splitting

We checked for exact duplicate files and removed four duplicates, leaving 5417 images. The dataset was partitioned into training (70%), validation (15%), and test (15%) subsets through stratified random sampling with a fixed seed. The same split was used for every model (Table I). The classes are not balanced. There are 8.7 times more MS images than healthy images. All images were resized to 224 × 224 pixels and normalised according to the ImageNet channel statistics. Augmentation was used only for the training batches and only after the data split. It included random horizontal and vertical flips, rotation up to ±15°, scaling by ±10%, translation by ±5%, and brightness and contrast changes of ±15%.

TABLE I. Class Distribution of Yellow-Rust-19 Images After Duplicate Removal and Stratified Splitting

Severity class	Train	Val.	Test	Total
0 (no disease)	143	31	31	205
R (resistant)	253	54	54	361
MR (moderately resistant)	394	85	84	563
MRMS (MR to MS)	794	170	170	1134
MS (moderately susceptible)	1255	269	269	1793
S (susceptible)	952	204	205	1361
Total	**3791**	**813**	**813**	**5417**

C. Models

Six architectures were compared using their torchvision implementations: MobileNetV3-Small (MNV3-S) [12], ShuffleNetV2 1.0× (SNV2) [13], MobileNetV2 (MNV2) [11], MobileNetV3-Large (MNV3-L) [12], EfficientNet-B0 (EffNet-B0) [14], and ResNet-50 [15]. All networks started with ImageNet [16] weights. The final fully connected layer was replaced with a layer that has six outputs.

D. Training Protocol

All models were trained in PyTorch [17] using the same settings. AdamW was used with a learning rate of 3×10^{-4}, a weight decay of 0.05, and a batch size of 64. A one-epoch linear warm-up was followed by cosine learning-rate decay. Models were trained for up to 30 epochs, with early stopping after seven consecutive

ISBN: 978-605-80193-6-2

epochs without improvement in validation macro-F1. The best validation checkpoint was used for testing.

To address class imbalance, class-weighted cross-entropy loss was used. Mixed-precision training was performed on an NVIDIA Tesla T4 GPU. Each model was trained with three random seeds (42, 123, and 2026), and the results are reported as mean ± standard deviation.

E. Evaluation Metrics

The test set was evaluated using overall accuracy, macro-F1, and quadratic weighted kappa (QWK). Macro-F1 gives equal importance to each class, which makes it useful for imbalanced data. QWK considers the order of the severity levels and gives a higher penalty to larger prediction errors.

Model efficiency was evaluated based on the number of parameters, multiply-accumulate operations (MACs) for a 224 × 224 input, model size in 32-bit format, and CPU latency. Latency was calculated as the median of 100 forward passes with a batch size of one on a single Intel Xeon 2.0 GHz CPU thread, after 10 warm-up passes. Grad-CAM was used to visualize the image regions that influenced the model predictions.

IV. RESULTS

A. Classification Performance

Table II gives the test results. The six networks had quite similar results. Accuracy was between 70.8% and 72.7%, macro F1 was between 64.1% and 67.9%, and QWK was between 0.898 and 0.910. EfficientNet-B0 had the highest macro F1 at 67.9 ± 1.4%. MobileNetV3-Small was close at 67.4 ± 0.5%, and it had the highest accuracy as well as the smallest variation between the three seeds. Despite having approximately 15 times more parameters than MobileNetV3-Small, ResNet-50 placed in the middle of the ranking and did not outperform the best lightweight models on any metric except a marginally higher QWK. Since the differences between most models are similar to the variation between seeds, we do not think one architecture can clearly be called the best one.

TABLE II. Test Performance

Model	Params (M)	MACs (G)	Latency (ms)	Accuracy (%)	Macro F1 (%)	QWK
MNV3-S	1.52	**0.055**	**9.1**	**72.7±1.0**	67.4±0.5	0.906
SNV2	**1.26**	0.144	17.2	70.8±1.4	65.8±1.3	0.904
MNV3-L	4.21	0.215	18.9	71.5±1.2	65.3±1.4	0.903
MNV2	2.23	0.300	22.2	71.3±1.6	64.1±2.1	0.898
EffNet-B0	4.02	0.385	33.5	72.6±1.3	**67.9±1.4**	0.908
ResNet-50	23.52	4.087	111.8	71.8±1.4	66.1±0.6	**0.910**

The class-level results in Table III show a similar pattern for all the architectures. The susceptible class was the easiest to recognise, with F1 scores from 82% to 84%. MS and MRMS came next, at around 70% to 74%. The resistant classes were harder. The F1 score for R was between 49.6% and 53% for all models, while MR was between 51.7% and 61%. The healthy class showed considerable variation across the seeds, with differences of up to ±8.4 percentage points. This variation may be related to the relatively small number of healthy leaves in the test set, which contained only 31 samples.

TABLE III. Per-Class Test F1

Model	0	R	MR	MRMS	MS	S
MNV3-S	64.0±2.4	51.0±5.5	**61.3±2.2**	70.7±2.8	73.4±2.0	**84.0±1.3**
SNV2	63.8±3.4	50.7±3.3	57.4±1.5	71.0±2.5	69.5±2.9	82.5±0.9
MNV3-L	59.7±5.8	50.2±0.8	53.8±3.0	**72.8±1.7**	73.2±1.0	82.3±3.1
MNV2	53.4±8.4	49.6±5.0	54.4±6.7	70.3±2.2	73.1±2.2	83.8±0.4
EffNet-B0	**68.7±5.8**	51.8±7.3	58.6±2.0	71.4±0.6	73.3±2.5	83.3±0.6
ResNet-50	64.3±1.2	**53.0±6.0**	51.7±2.1	69.7±5.2	**73.8±2.7**	83.9±2.7

B. Accuracy and Efficiency Trade-Off

Figure 1 compares macro F1 with latency. MNV3-S was the fastest model, with 9.1 ms per image, 55 million MACs and a 5.9 MB model file. ResNet-50 was much slower, taking 111.8 ms, with 4.09 billion MACs and a 90 MB model. MNV3-S and EffNet-B0 provide the best trade-offs between performance and computational cost. The other four models are both slower and less accurate than at least one of these two. Going from MNV3-S to EffNet-B0 improves macro F1 by only 0.5 points, but latency becomes 3.7 times higher. Parameter count was not a good way to predict speed. For example, SNV2 has the fewest parameters, 1.26 million, but it was 1.9 times slower than MNV3-S. Training time followed a similar pattern. MNV3-S took about 53 seconds per run on the T4 GPU, while ResNet-50 took about 262 seconds.

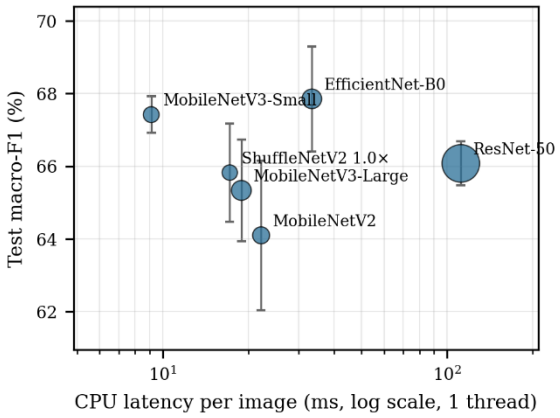


Fig. 1. Accuracy and efficiency trade-off

C. Error Analysis

Figure 2 shows the confusion matrix for EffNet-B0 using the seed with the median macro F1. There were 216 test errors, and 199 of them, or 92%, were between neighbouring severity levels. No healthy or resistant leaf was classified as MS or S. The most common confusions were healthy leaves classified as R (39% of healthy leaves), MR classified as R (29%), MRMS classified as MS (25%), and MS classified as S (18%). The differences between these categories are mainly related to lesion size and the extent of chlorosis. These categories can also be difficult to distinguish consistently, particularly in cases where the symptoms fall between two severity levels. The high QWK values, around 0.90 for all models, show that the predictions usually kept the correct order of severity even when the exact class was wrong.

ISBN: 978-605-80193-6-2

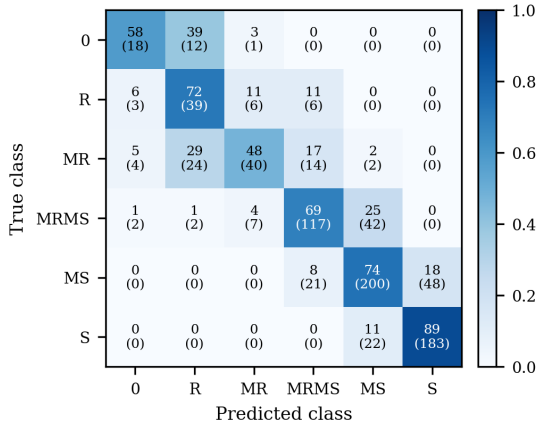


Fig. 2. Confusion matrix of EfficientNet-B0 on the test set

Figure 3 shows Grad-CAM maps from EffNet-B0 for one correctly classified test leaf from each class. For MR, MRMS, MS and S, the strongest activations were around the chlorotic stripes and groups of pustules. For S, the activated area covered most of the heavily infected leaf surface. For healthy leaves, the activation was spread across the leaf, which makes sense because the model was using the lack of lesions. For the R example, the strongest response was near the lower leaf edge and partly on the background. This suggests that the model may be using leaf-margin information for this class. However, this is based on only one example, so it needs to be checked on more images.

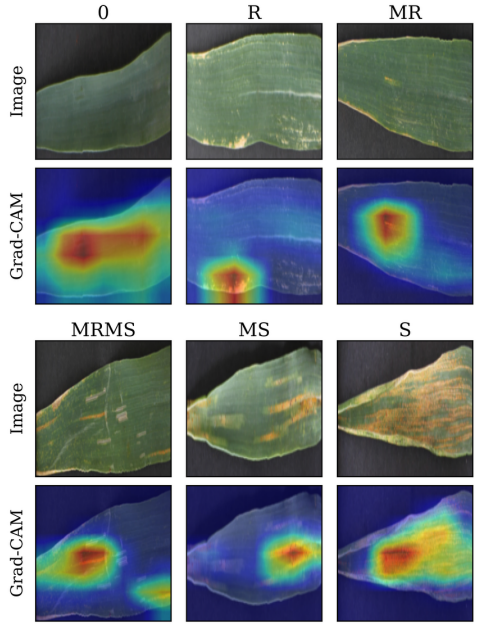


Fig. 3. Grad-CAM maps of EfficientNet-B0

V. DISCUSSION

The main practical result is that a bigger model did not give a clear advantage for leaf-level yellow rust severity classification on this dataset. The smallest network performed about as well as a model that is 15 times larger. MNV3-S took 9.1 ms per image on one server CPU thread, and its weights take less than 6 MB. The measured latency indicates that real-time inference may be feasible on mid-range smartphones. However, these results were obtained on the test system used in this study, so performance on the actual phones or embedded devices should be evaluated separately.

The accuracy obtained in this study was around 72%, which is lower than the 91–92% reported in previous studies using Yellow-Rust-19 [5], [6]. However, the results are not directly comparable because splitting procedures were different. In the earlier studies, augmented images were included in the dataset, which may have resulted in images from the same leaf appearing in both the training and test sets. This can lead to overly optimistic test accuracy because the model may encounter very similar samples during evaluation. In our experiments, we started with the raw images, removed duplicate samples, and performed augmentation only after separating the training and test sets. As a result, the evaluation provides a more realistic measure of performance on previously unseen leaves. We therefore recommend that future work using Yellow-Rust-19 clearly describe how the data are divided before reporting the results.

This study also has some limitations. The wheat leaves are relatively long and narrow, but they were resized to a square input format. This changes their original proportions and may make small lesions more difficult for the models to detect. Using the original aspect ratio or dividing the leaves into smaller image regions could be useful, particularly for samples with early-stage infection. Another limitation is that all images were collected at a single site and captured against a relatively uniform dark background. So, performance in real field conditions, where the background and lighting can change, still needs to be tested. The latency was measured on a server CPU and no quantisation was used. The healthy class is also small, and only three random seeds were used. Finally, the cross-entropy loss does not use the fact that the severity classes are ordered. Ordinal loss functions may help reduce mistakes between neighbouring levels.

VI. CONCLUSION

This study compared five lightweight CNNs with a ResNet-50 reference for six-level wheat yellow rust severity classification using the raw Yellow-Rust-19 images and a leakage-free protocol. All models had similar accuracy, and the larger network did not give a clear benefit. MobileNetV3-Small gave the best balance between accuracy and speed, with 67.4% macro F1 and 9.1 ms per image. Most mistakes were between neighbouring severity levels. In future work, quantized models suitable for deployment on smartphones will be developed, and their performance will be evaluated using real-world field images.

References

[1] S. Savary, L. Willocquet, S. J. Pethybridge, P. Esker, N. McRoberts, and A. Nelson, “The global burden of pathogens and pests on major food crops,” *Nat. Ecol. Evol.*, vol. 3, no. 3, pp. 430–439, Feb. 2019, doi: 10.1038/s41559-018-0793-y.

[2] W. Chen, C. Wellings, X. Chen, Z. Kang, and T. Liu, “Wheat stripe (yellow) rust caused by *P uccinia striiformis* f. sp. *tritici*,” *Mol. Plant Pathol.*, vol. 15, no. 5, pp. 433–446, Jun. 2014, doi: 10.1111/mpp.12116.

ISBN: 978-605-80193-6-2

[3] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using Deep Learning for Image-Based Plant Disease Detection," *Front. Plant Sci.*, vol. 7, p. 1419, Sep. 2016, doi: 10.3389/fpls.2016.01419.

[4] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Comput. Electron. Agric.*, vol. 145, pp. 311–318, Feb. 2018, doi: 10.1016/j.compag.2018.01.009.

[5] T. Hayit, H. Erbay, F. Varçın, F. Hayit, and N. Akci, "Determination of the severity level of yellow rust disease in wheat by using convolutional neural networks," *J. Plant Pathol.*, vol. 103, no. 3, pp. 923–934, Aug. 2021, doi: 10.1007/s42161-021-00886-2.

[6] T. Hayıt, H. Erbay, F. Varçın, F. Hayıt, and N. Akci, "The classification of wheat yellow rust disease based on a combination of textural and deep features," *Multimed. Tools Appl.*, vol. 82, no. 30, pp. 47405–47423, Dec. 2023, doi: 10.1007/s11042-023-15199-y.

[7] S. Kapoor and A. Narayanan, "Leakage and the reproducibility crisis in machine-learning-based science," *Patterns*, vol. 4, no. 9, p. 100804, Sep. 2023, doi: 10.1016/j.patter.2023.100804.

[8] E. C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification," *Comput. Electron. Agric.*, vol. 161, pp. 272–279, 2019.

[9] U. Shafi *et al.*, "Wheat Yellow Rust Disease Infection Type Classification Using Texture Features," *Sensors*, vol. 22, no. 1, p. 146, Dec. 2021, doi: 10.3390/s22010146.

[10] Z. Mi, X. Zhang, J. Su, D. Han, and B. Su, "Wheat Stripe Rust Grading by Deep Learning With Attention Mechanism and Images From Mobile Devices," *Front. Plant Sci.*, vol. 11, p. 558126, Sep. 2020, doi: 10.3389/fpls.2020.558126.

[11] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT: IEEE, Jun. 2018, pp. 4510–4520. doi: 10.1109/CVPR.2018.00474.

[12] A. Howard *et al.*, "Searching for MobileNetV3," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Korea (South): IEEE, Oct. 2019, pp. 1314–1324. doi: 10.1109/ICCV.2019.00140.

[13] N. Ma, X. Zhang, H.-T. Zheng, and J. Sun, "ShuffleNet V2: Practical Guidelines for Efficient CNN Architecture Design," in *Computer Vision – ECCV 2018*, vol. 11218, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds., in Lecture Notes in Computer Science, vol. 11218. , Cham: Springer International Publishing, 2018, pp. 122–138. doi: 10.1007/978-3-030-01264-9_8.

[14] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*, PmLR, 2019, pp. 6105–6114.

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 770–778. doi: 10.1109/CVPR.2016.90.

[16] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL: IEEE, Jun. 2009, pp. 248–255. doi: 10.1109/CVPR.2009.5206848.

[17] A. Paszke *et al.*, "Pytorch: An imperative style, high-performance deep learning library," *Adv. Neural Inf. Process. Syst.*, vol. 32, 2019.

ISBN: 978-605-80193-6-2

Seasonal Characteristics of Long-Term Precipitation Trends in Douala, Cameroon, Based on ERA5 Reanalysis

Abdel Karim Kouotou Ndam[1] Djiguiba Abdallah Kone Gogbe[1] Soner Çağatay Bağçacı[1]
[1]Karamanoglu Mehmetbey University

Article note: This paper is derived from the first author's graduation thesis.

* Corresponding author e-mail: scbagcaci@kmu.edu.tr

***Abstract*—Understanding the long-term precipitation variability is essential for climate-risk assessment, particularly in rapidly urbanizing coastal regions exposed to rainfall-related hazards. This study investigates precipitation trends over Cameroon, with a particular emphasis on Douala, using ERA5 daily precipitation data for the period 1976–2025. Daily data were quality controlled and aggregated into monthly, seasonal, and annual totals. Trend detection was performed using the Mann–Kendall test, the Seasonal Mann–Kendall test, and Sen's slope estimator. The analysis focused on aggregated time scales, as raw daily precipitation series are strongly affected by intermittency and seasonality. Results indicate a statistically significant decreasing trend in monthly precipitation when seasonality is accounted for, with a Seasonal Mann–Kendall statistic of S = −1346 (p = 0.001163) and a Sen's slope of approximately −0.573 mm year−1. Calendar-month analysis shows that this decrease is concentrated mainly in December–April. Seasonal analysis reveals significant decreases during DJF and MAM, while no significant trends are detected for JJA, SON, or annual precipitation totals. These findings suggest that precipitation change in Douala during 1976–2025 is characterized primarily by seasonal redistribution rather than a uniform decline in annual rainfall, with potential implications for flood risk, water management, and urban planning.**

***Keywords*—Douala, Cameroon, ERA5, Mann–Kendall test, Sen's slope**

I. INTRODUCTION

Climate change influences precipitation through complex interactions between thermodynamic and dynamic processes. While the Clausius–Clapeyron relationship suggests an approximate 7% increase in atmospheric moisture-holding capacity per 1°C of warming, this does not directly translate into increased precipitation, as rainfall is strongly modulated by atmospheric circulation, regional climate characteristics, and land–sea interactions [1], [2].

Cameroon exhibits pronounced climatic diversity due to its latitudinal extent, complex topography, and proximity to the Gulf of Guinea. The southern part of the country is characterized by a humid tropical climate with two rainy seasons—a main rainy season from mid-March to June and a shorter secondary rainy season from September to November—whereas the northern regions experience a semi-arid climate with a single rainy season generally extending from June to September [3]. In western Cameroon, locations such as Debundscha near Mount Cameroon receive some of the highest annual rainfall totals in Africa, exceeding 10,000 mm [4].

Previous studies have reported strong spatiotemporal variability in rainfall across Cameroon. Ebodé [5] documented substantial regional differences in precipitation variability, while Kaaria and Guindo [3] reported decreasing rainfall tendencies in northern and far northern regions. Coastal regions, including the Littoral Region where Douala is located, may respond differently due to their direct exposure to maritime moisture transport from the Gulf of Guinea.

Douala is the largest city and a major economic hub of Cameroon. Its low elevation, coastal location, and rapid urban expansion increase its vulnerability to rainfall-related hazards, particularly flooding. Recent studies have highlighted the growing risks posed by rainfall variability to infrastructure, population, and economic activities in the city [6]. Despite its importance, long-term precipitation trend analyses in Cameroon have largely focused on inland regions, with comparatively limited attention given to Douala and the surrounding coastal zone.

This study aims to examine the spatial and temporal characteristics of precipitation variability over Cameroon, with particular emphasis on Douala, for the period 1976–2025 using ERA5 reanalysis data. ERA5 provides spatially and temporally consistent atmospheric fields and is particularly valuable in regions where long-term station observations may be sparse or discontinuous [7]. Precipitation trends are assessed using non-parametric methods that are widely applied in hydroclimatological studies, namely the Mann–Kendall test, the Seasonal Mann–Kendall test, and Sen's slope estimator [8]–[10].

II. DATA AND METHODS

A. Study Area

The analysis focuses on Cameroon, with particular emphasis on Douala, located in the Littoral Region along the Gulf of Guinea (Fig. 1). Douala experiences a humid tropical climate and is directly influenced by maritime moisture fluxes. Its low elevation and dense urban development make it particularly sensitive to rainfall variability and flood-related impacts, providing a suitable case study for examining long-term precipitation changes and their seasonal characteristics.

ISBN: 978-605-80193-6-2

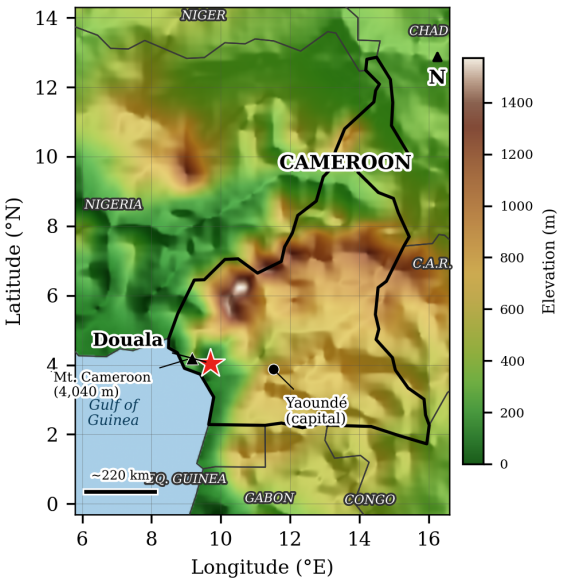


Fig. 1. Location of Douala within Cameroon, on the Gulf of Guinea coast, with country borders and shaded-relief topography (ERA5 orography, 0.25°). Mt. Cameroon and the capital, Yaoundé, are shown for reference.

B. ERA5 Precipitation Data

Daily precipitation data were obtained from the ERA5 reanalysis for the period 1 January 1976 to 31 December 2025. ERA5 provides globally consistent reanalysis fields at high spatial and temporal resolution and has been widely used for climate variability and trend analyses [7]. The dataset comprises 18,263 daily observations for Douala.

Prior to analysis, the dataset was screened for missing dates, duplicate records, and missing precipitation values. No missing or duplicate records were identified. Daily precipitation values range from 0.0 mm to 255.64 mm, with a mean of 8.84 mm day−1 and a median of 6.36 mm day−1.

C. Temporal Aggregation

Trend analysis was conducted primarily at monthly, seasonal, and annual time scales. Daily precipitation series are characterized by strong intermittency, a large proportion of zero-rainfall days, and pronounced seasonality. Applying trend tests directly to raw daily data can therefore yield statistically significant results that are difficult to interpret climatologically. Consequently, daily data were used only as the basis for aggregation.

Monthly precipitation totals were calculated by summing daily precipitation within each calendar month. Seasonal precipitation totals were computed for four climatological seasons: DJF (December–February), MAM (March–May), JJA (June–August), and SON (September–November). For DJF, December was assigned to the subsequent year to ensure seasonal continuity. Annual precipitation totals were obtained by summing daily precipitation within each calendar year.

D. Trend Detection Methods

Monotonic trends were evaluated using the Mann–Kendall test [8], [9]. This non-parametric test is based on data ranks and does not assume normality, making it suitable for precipitation data. Statistical significance was assessed at the 5% level.

To account for the strong seasonal cycle of monthly precipitation, the Seasonal Mann–Kendall test was applied using the 12 calendar months as seasons. This approach evaluates trends within comparable months across years and avoids mixing months with different climatological characteristics. In addition, separate Mann–Kendall tests were conducted for each calendar month to identify the months contributing most strongly to the overall trend.

Trend magnitudes were quantified using Sen's slope estimator, which provides a robust estimate of the median rate of change and is less sensitive to outliers than ordinary least squares regression [10]. Sen's slope values are reported in mm year−1.

E. Manual Verification of the Mann–Kendall Statistic

Before the algorithmic implementation was applied to the full 1976–2025 series, the pairwise sign comparisons underlying the Mann–Kendall S statistic were additionally computed by hand, following the definition in Section II-D. Table I lists the manual comparisons for the first 10 daily observations of the record (1–10 January 1976); the resulting sum, S = 17, exactly reproduces the value obtained from the automated algorithm for the same 10 values. The same manual procedure was then extended to the complete first year of record (1 January–31 December 1976, n = 366), for which the cumulative hand calculation again matched the automated result exactly (S = 754). This agreement, both for a small illustrative subset and for a full annual series, was used to confirm the correctness of the algorithm before it was applied to the complete 50-year dataset.

TABLE I. MANUAL VERIFICATION OF THE MANN–KENDALL S STATISTIC FOR THE FIRST 10 DAILY OBSERVATIONS (1–10 JANUARY 1976)

i	Date	xi (mm)	Comparisons	Row sum Si
1	1976-01-01	0.93	9	9
2	1976-01-02	1.51	8	4
3	1976-01-03	1.91	7	−1
4	1976-01-04	1.52	6	2
5	1976-01-05	2.44	5	−3
6	1976-01-06	1.24	4	4
7	1976-01-07	1.62	3	1
8	1976-01-08	2.26	2	0
9	1976-01-09	1.39	1	1
—	Total (n = 10)	—	—	17

Si = (number of later, larger values) − (number of later, smaller values) for observation i; ΣSi = S. The n = 366 check (full year 1976) reproduced S = 754, matching the automated computation for the same subset.

III. RESULTS

Table II summarizes the Mann–Kendall and Sen's slope statistics obtained for every temporal aggregation examined, from the raw daily series to annual totals, over the complete 1976–2025 record; these results are discussed by time scale in the following subsections.

ISBN: 978-605-80193-6-2

TABLE II. SUMMARY OF MANN–KENDALL TREND STATISTICS ACROSS TEMPORAL SCALES, DOUALA, 1976–2025

Scale	n	S	Z	p	Sen's slope (mm yr⁻¹)	Trend
Daily	18,263	−10,655,059	−12.95	<0.001	—	Dec.
Monthly (MK)	600	−4,730	−0.96	0.335	−0.027	n.s.
Monthly (Seasonal MK)	600	−1,346	−3.25	0.001	−0.573	Dec.
DJF	51	−587	−4.76	<0.001	−4.51	Dec.
MAM	50	−309	−2.58	0.010	−3.59	Dec.
JJA	50	65	0.54	0.592	+1.55	n.s.
SON	50	185	1.54	0.124	+1.79	n.s.
Annual	50	−111	−0.92	0.358	−3.69	n.s.

Dec. = statistically significant decreasing trend ($p < 0.05$); n.s. = not significant. Sen's slope for the raw daily series was not computed exactly, owing to the computational cost of the pairwise-slope calculation at $n = 18{,}263$; its ordinary-least-squares trend is approximately −0.005 mm yr⁻1, i.e., statistically significant but climatologically small at the daily scale (see the monthly and seasonal aggregates for the physically meaningful trend magnitudes).

A. Monthly Trends

The Seasonal Mann–Kendall test applied to monthly precipitation totals indicates a statistically significant decreasing trend over the 1976–2025 period (S = −1346, Z = −3.2478, p = 0.001163). The corresponding Sen's slope is approximately −0.573 mm year−1, indicating a gradual but persistent decline in monthly precipitation when seasonality is explicitly considered.

Calendar-month analysis reveals that the decreasing signal is not uniform throughout the year. Statistically significant decreases are detected in January, February, March, April, and December, whereas no significant trends are found for most months between May and November. This pattern indicates that the monthly-scale decrease is concentrated primarily in the early part of the year and at the end of the calendar year.

B. Seasonal Trends

Seasonal precipitation analysis shows that significant decreases are confined to specific seasons. DJF exhibits a strong decreasing trend (S = −587, p < 0.001) with a Sen's slope of approximately −4.51 mm year−1. MAM also shows a statistically significant decrease (S = −309, p = 0.00998) with a Sen's slope of approximately −3.59 mm year−1.

In contrast, no statistically significant trends are detected for JJA or SON. These results indicate a clear seasonal asymmetry in precipitation change, with declining rainfall during DJF and MAM and relatively stable conditions during the core rainy seasons (Fig. 2).

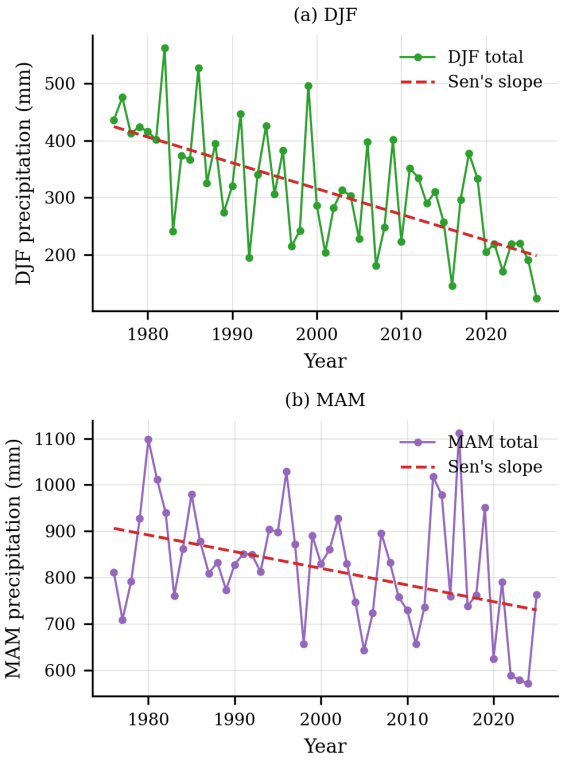


Fig. 2. Seasonal precipitation totals with statistically significant decreasing trends and Sen's slope trend lines: (a) DJF (December–February); (b) MAM (March–May).

C. Annual Trends

Annual precipitation totals do not exhibit a statistically significant trend over the study period (S = −111, p = 0.3575). Although the estimated Sen's slope is negative, the trend is not statistically distinguishable from zero, indicating that total annual rainfall in Douala has not changed significantly during the last five decades (Fig. 3).

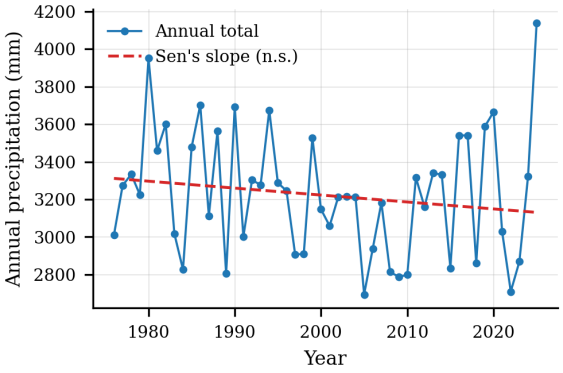


Fig. 3. Annual precipitation totals for Douala, 1976–2025, with the Sen's slope trend line (trend not statistically significant, p = 0.358).

D. Synthesis

Overall, the results demonstrate that precipitation changes in Douala are expressed primarily through seasonal redistribution rather than changes in annual totals. Decreasing rainfall during DJF and MAM is offset by the absence of significant trends during JJA and SON, resulting in no statistically significant change in annual precipitation.

ISBN: 978-605-80193-6-2

IV. CONCLUSION

Using ERA5 reanalysis data for 1976–2025, this study examined long-term precipitation variability over Cameroon, with particular emphasis on Douala. While annual precipitation totals do not show a statistically significant trend, significant decreases are detected at monthly and seasonal scales when seasonality is explicitly considered. The decline is concentrated mainly in December–April and during the DJF and MAM seasons.

These findings suggest that recent precipitation changes in Douala are characterized by a redistribution of rainfall within the year rather than a uniform long-term decrease in annual totals. Such seasonal shifts may have important implications for flood risk management, urban drainage design, and climate adaptation strategies in this rapidly growing coastal city.

References

[1] Berg, P., Christensen, O. B., Klehmet, K., Lenderink, G., Olsson, J., Teichmann, C., & Yang, W. (2019). Summertime precipitation extremes in a EURO-CORDEX 0.11° ensemble at an hourly resolution. *Natural Hazards and Earth System Sciences, 19*(4), 957–971. https://doi.org/10.5194/nhess-19-957-2019

[2] Duzenli, E., Yucel, I., Pilatin, H., & Yilmaz, M. T. (2021). Evaluating the performance of a WRF initial and physics ensemble over Eastern Black Sea and Mediterranean regions in Turkey. *Atmospheric Research, 248*, 105184. https://doi.org/10.1016/j.atmosres.2020.105184

[3] Kaaria, W., & Guindo, A. (2024). *Cameroun Analyse Climatique Août 2024.*

[4] Nkiaka, E., & Tume, S. J. P. (2025). Changing rainfall patterns and their climatic drivers in one of the rainiest places on Earth, Debundscha, Gulf of Guinea. *International Journal of Climatology, 45*(11), e70009. https://doi.org/10.1002/joc.70009

[5] Ebodé, V. B. (2022). Analysis of the spatio-temporal rainfall variability in Cameroon over the period 1950 to 2019. *Atmosphere, 13*(11), 1769. https://doi.org/10.3390/atmos13111769

[6] Chounna Yemele, G., Minderhoud, P. S. J., Ohenhen, L. O., Seeger, K., Shirzaei, M., & Teatini, P. (2026). Land subsidence dynamics and their interplay with spatial and temporal land-use transitions in the Douala coastland, Cameroon. *Natural Hazards and Earth System Sciences, 26*, 551–570. https://doi.org/10.5194/nhess-26-551-2026

[7] Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., Nicolas, J., Peubey, C., Radu, R., Schepers, D., Simmons, A., Soci, C., Abdalla, S., Abellan, X., Balsamo, G., Bechtold, P., Biavati, G., Bidlot, J., Bonavita, M., & Thépaut, J.-N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society, 146*(730), 1999–2049. https://doi.org/10.1002/qj.3803

[8] Mann, H. B. (1945). Nonparametric tests against trend. *Econometrica, 13*(3), 245–259. https://doi.org/10.2307/1907187

[9] Kendall, M. G. (1975). *Rank Correlation Methods* (4th ed.). Charles Griffin.

[10] Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *Journal of the American Statistical Association, 63*(324), 1379–1389. https://doi.org/10.1080/01621459.1968.10480934

ISBN: 978-605-80193-6-2

Sand Column Test for Evaluating the Performance of Grout Injection in Historic Structures

Murat ERDEMİR[1] Prof. Dr. Sadık Alper YILDIZEL[2]

[1]Konya Technical University, Graduate Education Institute, Department of Architecture, Konya, Türkiye

[2]Karamanoglu Mehmetbey University, Faculty of Engineering, Department of Civil Engineering, Karaman, Türkiye

* Corresponding author e-mail: e228126001004@ktun.edu.tr

***Abstract*—From the past to the present, numerous studies have been conducted and various methods developed to reinforce historical and modern structures. One such method—the injection reinforcement technique—is widely used to address strength losses in structures caused by voids in the foundation and within walls. This method is also applied in crack repairs to prevent water leaks in structures such as dams. With advancing technology, injection applications can now be carried out in a much more comprehensive and precise manner. The injection process is fundamentally based on the principle of injecting a special liquid grout into cracks and voids in the structure using injection machines. For this reason, the prepared injection slurry must possess injectability properties suitable for the specific structure of the void or crack into which it will be applied. Testing the material's injectability prior to field application is of great importance. This helps prevent potential problems, saves time, and avoids possible economic losses.**

This study will focus on injection applications designed to fill voids within rubble walls of historic structures. The study will describe the sand column test setup and its operating principle, which were developed to test the penetration properties of the injection material. The aim is to create a model of the voids in the walls of historic structures and to evaluate the penetration properties of the injection material within these voids. The goal is for the injection material used in historic structures to make a tangible contribution to the structural reinforcement processes of these structures.

***Keywords*— Injection method, historic structure, reinforcement, restoration, sand column test**

I. INTRODUCTION

Techniques and methods for strengthening and repairing structures are constantly evolving. While some methods are reversible, others are not. Applying methods that are very difficult or impossible to reverse to a structure carries certain risks. Otherwise, instead of repairing or strengthening the structure or its materials, permanent damage may be caused. The grout injection method is also one of the irreversible methods. If this method is to be used in a structure, great care must be taken, and all necessary short- and long-term tests on the injection material must be conducted.

A review of the literature reveals that numerous studies have been conducted on the injectability of grouts [1–8]. These studies generally utilized two different types of binders: cement-based grouts [2, 9, 10] and lime-based grouts [11, 12]. While cement-based grouts are used in new structures and dams, lime-based grouts are used more frequently in historic structures.

In the studies conducted, various sand column setups were established to evaluate the penetration of the injection material. Some of these are as follows:

• Kalagri et al. [12] used plastic cylindrical containers 250 mm in diameter and 500 mm in height. A mixture of stone and mortar in a 20:50 ratio was prepared and poured into the containers to create a wall cavity model. The injection process was then carried out at a constant pressure of 0.75 bar.



Fig. 1. Cylindrical containers prepared for injection [12].

- Binda et al. [11] prepared cylindrical Plexiglas containers 120 mm in diameter and 300 mm in height and filled them with mortar, travertine, and carcarenite taken from the structure. The injection process was carried out at a pressure below 1 atm.

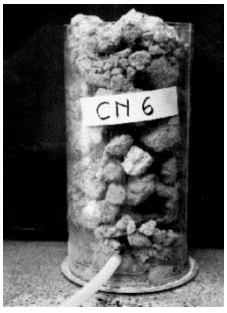


Fig. 2. Cylindrical Plexiglas containers made from materials salvaged from building rubble [11].

- Akyurt Takır [13] used a transparent plastic tube 360 mm long and 22.2 ± 0.3 mm in diameter. He conducted the sand column test by placing silica sand with a grain size range of 2.5–4 mm inside the tube.

ISBN: 978-605-80193-6-2



Fig. 3. An experimental setup prepared using silica sand in the 2.5–4 mm range [13].

Within the scope of this study, the methods and equipment used to test the penetration property—one of the fresh-state properties of lime-based injection materials used in the reinforcement of historic structures and crack repair—were examined. To test the material's penetration property, the sand column test specified in TS EN 1771 [14] was conducted. The sand column test is extremely important for understanding how the injection material will penetrate a structure and whether blockages will occur before the material is applied to the structure. This allows potential problems that may arise during the injection process to be identified in advance, thereby saving time and money. The scope of this study includes creating a model of the voids within the walls of a historic structure in a test setup, and examining and evaluating the penetration properties of the injection material in this setup. Additionally, one of the study's objectives is to understand the logic behind the test setup and to develop a guide for its construction and use in future studies.

II. METHOD

Quantitative research methods were used in this study. A sand column test was conducted to determine the penetration properties of a natural hydraulic lime-based injection material used in the reinforcement and restoration of historic structures. This test setup consists of various components. These components and their purposes are as follows:

- Pressure vessel: A metal vessel into which the injection syrup is placed and which uses the pressure generated inside to pump the syrup.
- Air source: A device that continuously supplies air to the tank containing the syrup to generate pressure.
- Air regulator: A device that ensures the air from the air source is delivered to the pressure vessel at the desired intervals and at a constant pressure.
- Air pressure gauge: A calibrated gauge located on the air regulator and on the pressure vessel that indicates the pressure level inside the vessel.
- Transparent plastic air hose: A transparent hose with an 8 mm diameter that conveys air from the air source to the pressure chamber.
- Transparent plastic liquid hose: A transparent hose that conveys the injection material from the pressure vessel to the sand-filled transparent tube.
- Air on/off valve: A valve located at the air inlet of the pressure vessel, used to open or close the air supply as needed.
- Transparent glass tube: A transparent cylindrical container 36 cm long and 22.2±0.3 mm in diameter, filled with sand.
- Sand: Washed sand with a grain size of 2–4 mm, used to simulate the environment in which the injection will take place.

Material discharge container: A graduated container through which the injection material flows past the sand-filled transparent container.



Fig. 4. Sand column test setup developed in the study



Fig. 5. Filling the interior of a transparent glass tube with 2–4 mm silica-based river sand.

The sand column test was conducted in accordance with TS EN 1771 [14]. The test utilized a transparent glass tube 36 cm long and 22.2 ± 0.3 mm in diameter, filled with silica-based river sand with grain sizes ranging from 2 to 4 mm [15]. Here, the time (T36) it took for the injection material to travel from the bottom to the top of the tube under a pressure of 0.075 ± 0.0025 MPa, as well as the amount of material accumulated in the graduated container, were measured.

III. CONCLUSIONS

Since the injection method used in the reinforcement of structures is irreversible, the material to be used must be compatible with the structure and meet the necessary requirements. If the injection material is to be used to reinforce or restore a historic structure, even greater care must be taken. Otherwise, these cultural heritage sites—which have managed to stand for centuries—may be damaged, resulting in permanent harm.

ISBN: 978-605-80193-6-2

Within the scope of this study, the focus was on the sand column test—used to evaluate one of the most important properties of the injection material, namely its penetration capability—and the components that make up the test setup. This test is crucial for assessing, in advance, the injection material's ability to penetrate the voids within the wall. In the study, the time (T36) it took for the natural hydraulic lime-based injection material to travel from the bottom to the top of the transparent glass tube was measured as 30 seconds. By the end of the 48th second, 20 ml of liquid had accumulated in the collection vessel. For the material to be considered acceptable, the slurry must travel from the bottom to the top of the tube in less than 50 seconds [16, 17]. Since the slurry accumulated 20 ml in the collection vessel in less than 50 seconds in this study, the material falls into the "easily applicable" category.

In conclusion, when producing injection materials, tests must be conducted in accordance with the necessary standards, and the material mixture best suited for the application must be identified. This study serves as an example for researchers and professionals in the field in terms of understanding how the sand column experimental setup is constructed and how it operates.

References

[1] Miltiadou-Fezans, A., & Tassios, T. P. (2013). Penetrability of hydraulic grouts. *Materials and Structures*, *46*, 1653-1671. https://doi.org/10.1617/s11527-012-0005-1

[2] Markou, I. N., Christodoulou, D. N., Petala, E. S., & Atmatzidis, D. K. (2018). Injectability of Microfine Cement Grouts into Limestone Sands with Different Gradations: Experimental Investigation and Prediction. *Geotechnical and Geological Engineering*, *36*(2), 959-981. https://doi.org/10.1007/s10706-017-0368-8

[3] Bras, A., & Henriques, F. M. A. (2012). Natural Hydraulic Lime Based Grouts–The Selection of Grout Injection Parameters for Masonry Consolidation. *Construction and Building Materials*, *26*(1), 135-144. https://doi.org/doi:10.1016/j.conbuildmat.2011.05.012

[4] Oktay, D., Yüzer, N., & Ulukaya, S. (2015). Rheological properties of hydraulic grouts used in consolidation of brick masonry walls. *REHAB2015, Porto, Portugal*, 22-24.

[5] Gökyiğit Arpacı, E. Y. (2016). Tarihi yığma yapıların onarımında kullanılan enjeksiyon malzemesinin (grout) performans değerlendirmesi ve 19. yüzyıl tuğla yığma yapılarda örneklenmesi. Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü, Mimarlık, Yüksek Lisans Tezi, İstanbul.

[6] Oktay, D. (2017). Tarihi Kireç Harçlarına Uygun Enjeksiyon Malzemelerinin Üretilmesi ve Model Duvarlarda Uygulamasının Değerlendirilmesi. Yıldız Teknik Üniversitesi, İnşaat Mühendisliği, Doktora Tezi, İstanbul.

[7] Gökyiğit-Arpaci, E. Y., Oktay, D., Yüzer, N., Ulukaya, S., & Ekşi-Akbulut, D. (2019). Performance evaluation of lime-based grout used for consolidation of brick masonry walls. *Journal of Materials in Civil Engineering*, *31*(6), 4019059-4019059. https://doi.org/10.1061/(ASCEMT.1943-5533.0002692

[8] Dinç Şengönül, B., Oktay, D., & Yüzer, N. (2020). Effect of temperature, resting time and brick dust (Horasan) on the rheological properties of hydraulic lime-based grouts. *Construction and Building Materials*, *265*, 120644. https://doi.org/10.1016/j.conbuildmat.2020.120644

[9] Axelsson, M., & Gustafson, G. (2010). The PenetraCone, a new robust field measurement device for determining the penetrability of cementitious grouts. *Tunnelling and Underground Space Technology*, *25*(1), 1-8. https://doi.org/10.1016/j.tust.2009.06.004

[10] Axelsson, M., Gustafson, G., & Fransson, Å. (2009). Stop mechanism for cementitious grouts at different water-to-cement ratios. *Tunnelling and Underground Space Technology*, *24*(4), 390-397. https://doi.org/10.1016/j.tust.2008.11.001

[11] Binda, L., Baronio, G., Tiraboschi, C., & Tedeschi, C. (2003). Experimental research for the choice of adequate materials for the reconstruction of the Cathedral of Noto. *Construction and Building Materials*, *17*(8), 629-639. https://doi.org/10.1016/S0950-0618(03)00059-X

[12] Kalagri, A., Miltiadou-Fezans, A., & Vintzileou, E. (2010). Design and Evaluation Of Hydraulic Lime Grouts For The Strengthening Of Stone Masonry Historic Structures. *Materials and Structures*, *43*(8), 1135-1146. https://doi.org/10.1617/s11527-009-9572-1

[13] Akyurt Takır, D. (2022). Tarihi Yığma Yapıların Onarımında Kullanılan Enjeksiyon Yöntemi ve Hidrolik Kireç Esaslı Enjeksiyon Malzemelerinin (Grout) Performans Değerlendirilmesi. İstanbul Teknik Üniversitesi, Mimarlık, Yüksek Lisans Tezi, İstanbul.

[14] TS EN 1771. (2004). Beton yapılar - Koruma ve tamir için mamul ve sistemler-Deney metotları - Enjekte edilebilirlik tayini ve yarma deneyi.

[15] Dinç Şengönül, B. (2024). Kapadokya Atık Volkanik Tüfü İle Üretilen Kireç Esaslı Enjeksiyon Malzemesinin Performansının Değerlendirilmesi. Yıldız Teknik Üniversitesi, İnşaat Mühendisliği, Doktora Tezi, İstanbul.

[16] Miltiadou-Fezans, A., & Tassios, T. P. (2012). Fluidity of Hydraulic Grouts for Masonry Strengthening. *Materials and Structures*, *45*(12), 1817-1828. https://doi.org/10.1617/s11527-012-9872-8

[17] Oktay, D., Uğuryol, M., Güner, M. B., Hazar Yoruç, A. B., & Yüzer, N. (2020). Tarihi binaların sağlamlaştırılması için killi kireçtaşından üretilen enjeksiyon malzemesinin karakterizasyonu. *Journal of the Faculty of Engineering and Architecture of Gazi University* *35*(3), 1387-1397. https://doi.org/10.17341/gazimmfd.468457

ISBN: 978-605-80193-6-2

Continuous 3D-Printed Polymer Reinforcement in Foam Concrete: A Review of Materials, Structures, Interface Behavior, and Research

Özge Zor Avcan[1], Sadık Alper Yıldızel[2], Ülkü Sultan Keskin[1]
[1]Department of Civil Engineering, Konya Technical University, Konya, Türkiye
[2]Department of Civil Engineering, Karamanoğlu Mehmetbey University, Karaman, Türkiye
* Corresponding author e-mail: ozor@ktun.edu.tr

***Abstract*—Foam concrete is a lightweight cementitious material characterized by a highly porous cellular structure, low density, and low thermal conductivity. These characteristics make it attractive for lightweight construction, filling, insulation, and other applications where reduced self-weight is advantageous. However, the same porous structure results in limited tensile and flexural capacity, brittle fracture, and low post-cracking resistance. Various reinforcement strategies, particularly discrete fiber reinforcement, have been investigated to overcome these limitations. Although fibers can improve crack control and post-cracking behaviour, their random distribution and orientation limit direct control over reinforcement architecture. Recent advances in additive manufacturing have enabled continuous polymeric meshes and lattice structures to be fabricated using fused deposition modelling (FDM), providing control over reinforcement geometry, orientation, material quantity, and spatial arrangement. This review critically examines the potential of continuous 3D-printed polymer reinforcement for cementitious composites, with particular emphasis on foam concrete and polyethylene terephthalate glycol (PET-G) systems. The review considers the relationship between foam-concrete porosity and mechanical performance, the advantages and limitations of discrete fiber reinforcement, architected polymer reinforcement, FDM-induced material anisotropy, polymer-cement interfacial behaviour, and durability. Existing studies have demonstrated that 3D-printed polymer lattices, grids, and other architected structures can improve ductility, crack bridging, energy absorption, and post-cracking behaviour in cementitious materials. However, most available studies have focused on conventional concrete, mortar, or ultra-high-performance concrete rather than highly porous foam concrete. Therefore, the reported mechanisms cannot be transferred directly to foam concrete without considering differences in pore structure, matrix continuity, confinement, and interfacial conditions. In particular, the combined effects of reinforcement quantity, mesh geometry, spatial placement, pore structure, and polymer-cement bond remain insufficiently understood. Future studies should independently control reinforcement amount and placement, characterize the polymer-cement interface quantitatively, evaluate density-normalized mechanical performance, and consider thermal and long-term durability characteristics simultaneously. Continuous 3D-printed polymer reinforcement therefore represents a promising approach for developing lightweight cementitious composites with controlled reinforcement architectures, but further experimental and mechanistic studies are required before reliable design principles can be established for foam concrete.**

Keywords—Foam concrete, 3D-printed reinforcement, polyethylene terephthalate glycol, fused deposition modelling, polymer mesh, lattice reinforcement, cementitious composites, polymer-cement interface.

I. INTRODUCTION

Foam concrete is a lightweight cementitious material in which a large volume of air voids is introduced into a cementitious matrix. Its low density, flowability, and thermal-insulation capability have resulted in applications in lightweight construction, filling materials, insulation layers, and other civil engineering applications [1], [2]. Recent reviews emphasized that foam concrete could be formulated over a wide range of densities and compositions, allowing its properties to be tailored according to the intended application [1], [2].

The cellular structure of foam concrete is simultaneously responsible for its advantages and limitations. Increasing the volume of air voids generally reduces density and thermal conductivity but also decreases the amount of solid material available to transfer mechanical stresses. Consequently, density, pore size, pore distribution, pore connectivity, matrix composition, and curing conditions strongly influence mechanical performance [3],[5]. Investigations of foamed cement paste further showed that changes in matrix composition could modify phase composition and microstructure [3], while micropore characteristics influenced both stress distribution and heat transfer [4].

The relationship between pore structure and mechanical behaviour becomes particularly important at low densities, where the cementitious walls surrounding the pores become relatively thin. Falliano et al. identified density and mixture-related parameters as important factors controlling the compressive strength of foamed concrete [5]. Thus, the mechanical design of a reinforced foam-concrete system cannot be separated from the characteristics of its cellular matrix.

A major limitation of foam concrete is its low tensile and flexural capacity. Because cementitious matrices are inherently weak in tension, cracking can develop at relatively low tensile stresses. The porous structure may further promote local stress concentrations and non-uniform crack propagation. Fiber-reinforced foamed concrete has therefore attracted considerable attention as

ISBN: 978-605-80193-6-2

a means of improving tensile, flexural, toughness, and post-cracking behaviour [6].

Discrete fibers provide an established reinforcement strategy. However, their effectiveness depends on dispersion, orientation, embedment length, and fiber-matrix interaction. Polypropylene fiber studies in lightweight concrete demonstrated their influence on shrinkage cracking [7], while carbon-fiber-reinforced foamed concrete was reported to exhibit modified mechanical performance and failure mechanisms [8].

Discrete fibers therefore provide reinforcement throughout the matrix, but their spatial distribution is largely governed by mixing and casting. This limits direct control over reinforcement orientation and location. A continuous reinforcement system offers a fundamentally different approach by defining the reinforcement architecture before fabrication.

Recent advances in additive manufacturing have enabled polymeric meshes, lattices, trusses, grids, and other architected structures to be fabricated using FDM. The use of digitally controlled reinforcement geometries allows filament orientation, aperture size, topology, and spatial distribution to be deliberately designed [9]. Studies on 3D-printed polymer lattices have demonstrated that architected polymer reinforcement can substantially enhance the ductility of cementitious composites [10]. Similarly, 3D-printed gyroid structures have been investigated as a means of modifying the mechanical response of concrete [11].

More recent investigations have extended this concept to different reinforcement architectures and embedded polymer systems. Namakiaraghi et al. demonstrated the potential of architected 3D-printed polymers in cement-based composites and highlighted the role of reinforcement architecture and surface characteristics [12]. Other studies have investigated flexible fiber-reinforced polymer grids and in-process embedded polymer reinforcement, demonstrating the broader potential of digitally fabricated continuous reinforcement systems [13], [14].

However, an important distinction must be made between dense cementitious materials and foam concrete. Most studies on architected polymer reinforcement have considered conventional concrete, mortar, or ultra-high-performance concrete [10],[14]. In contrast, foam concrete contains a substantially larger volume of pores and has a more heterogeneous load-bearing skeleton. Therefore, reinforcement mechanisms identified in dense cementitious matrices cannot automatically be assumed to operate in the same manner in foam concrete.

The suitability of a polymer reinforcement is also influenced by its manufacturing process. PET-G is a potentially attractive polymer for FDM-based reinforcement because it can be processed into controlled geometries. Nevertheless, the mechanical properties of FDM-produced PET-G components depend on layer height, infill strategy, printing direction, temperature, and other processing parameters [15],[18]. Consequently, the behaviour of printed reinforcement should be considered separately from the nominal properties of the raw filament.

The polymer-cement interface represents another critical component of the composite system. A printed mesh can contribute to load transfer only if stresses can be transferred effectively between the cementitious matrix and polymer reinforcement. Recent studies have demonstrated that FDM-induced surface morphology can influence mechanical interlocking and bond behaviour. Bol et al. reported that the morphology of FDM-printed polymer edges can increase bond strength through interlocking effects, with a reported increase of approximately 56% for the investigated configuration [19]. The importance of surface modification has also been demonstrated experimentally, showing that improved polymer-matrix bonding can lead to improved ductility and mechanical performance [20].

Durability represents an additional unresolved issue. Accelerated alkaline exposure studies have examined the stability of polymers used for 3D printing in cementitious or geopolymer-related environments [21]. In parallel, freeze-thaw deterioration has been investigated extensively for foam concrete, where repeated freezing and thawing can result in strength degradation [22].

Although several reviews have recently discussed 3D-printed reinforcement using FDM [23],[24], a specific synthesis connecting architected polymer reinforcement with the cellular structure and performance requirements of foam concrete remains limited. Moreover, recent developments in lightweight cementitious composites and polymer-based reinforcement indicate that the subject is evolving rapidly [25], [26].

Accordingly, this review critically evaluates the current literature on continuous 3D-printed polymer reinforcement, with particular emphasis on foam concrete and PET-G-based mesh systems. The review focuses on five interconnected issues: foam-concrete structure and mechanical behaviour, conventional fiber reinforcement, architected polymer reinforcement, FDM and PET-G characteristics, and polymer-cement interfacial and durability considerations. Particular attention is given to the differences between evidence obtained from dense cementitious materials and the requirements of highly porous foam concrete.

II. REVIEW METHODOLOGY

This study adopted a structured critical-review approach rather than a systematic review protocol. The literature was organized around the interaction of four main research fields: (i) foam concrete and its mechanical and physical behaviour, (ii) fiber-reinforced foam concrete, (iii) 3D-printed polymer reinforcement for cementitious composites, and (iv) FDM-produced polymer materials and polymer-cement interface behaviour.

The literature selection was guided by combinations of the following concepts: “foam concrete,” “foamed concrete,” “fiber-reinforced foamed concrete,” “3D-printed polymer reinforcement,” “polymer lattice concrete,” “architected polymer reinforcement,” “PET-G,” “FDM reinforcement,” “polymer-cement bond,” and “3D-printed cementitious composites.”

ISBN: 978-605-80193-6-2

Peer-reviewed journal articles were prioritized, particularly recent studies addressing architected reinforcement, FDM-based polymer components, reinforcement-matrix bond, and lightweight cementitious composites. Earlier studies were retained where they provided foundational evidence concerning foam-concrete behaviour or fiber reinforcement.

The review evaluated the literature according to five criteria:

1. relevance to lightweight or cementitious composites;
2. relevance to continuous or architected reinforcement;
3. characterization of reinforcement geometry and material;
4. evidence concerning mechanical or interfacial behaviour; and
5. applicability of the reported mechanisms to foam concrete.

The review did not exclude studies conducted on dense concrete, mortar, or ultra-high-performance concrete because they provide important mechanistic evidence. However, their findings were treated as indirect evidence when discussing foam concrete. This distinction is essential because differences in pore structure, matrix continuity, confinement, and density can alter reinforcement efficiency.

The review therefore did not attempt to establish a single quantitative performance ranking. Instead, it synthesized the available evidence to identify which mechanisms were reasonably established, which findings remained material- or geometry-dependent, and which research questions were unresolved specifically for foam concrete.

III. FOAM CONCRETE: STRUCTURE AND PERFORMANCE

A. Cellular Structure and Mechanical Behaviour

Foam concrete derives its characteristic properties from its cellular pore structure. The pore system is generated by introducing preformed foam into a cementitious matrix, producing a material containing a high volume fraction of air voids [1], [2], [27].

The resulting pore structure controls several properties simultaneously. A higher void content generally decreases density and thermal conductivity but also reduces mechanical strength. However, density alone cannot fully describe mechanical behaviour because pore size, pore shape, pore distribution, and pore connectivity influence stress concentrations and crack propagation [3], [5].

Chung et al. demonstrated that changes in matrix composition can modify the phase composition and microstructure of foamed cement paste [3]. Bie et al. further investigated micropores and their influence on stress distribution and heat conduction [4]. These findings indicate that reinforcement should not be considered independently from the cellular matrix.

For continuous reinforcement, this issue is particularly important because a printed filament may be surrounded by relatively dense cement paste in one region and by thin pore walls or locally porous material in another. Consequently, the effective bond and stress-transfer mechanism may vary along the reinforcement length.

B. Density and Thermal Performance

Low density is one of the principal advantages of foam concrete. Reduced self-weight can decrease structural dead loads and facilitate transportation and handling [1], [2].

The cellular structure also provides low thermal conductivity. However, mechanical and thermal requirements can conflict. Increasing the solid fraction generally improves strength but may increase density and thermal conductivity.

This trade-off becomes particularly relevant when polymer reinforcement is introduced. A reinforcement system that improves strength by substantially increasing the solid fraction may reduce the lightweight and insulating advantages of foam concrete.

Recent work on lightweight and ultra-lightweight printable cementitious composites further emphasizes the importance of balancing mechanical performance with density and processability [25].

Continuous mesh reinforcement provides a potentially material-efficient alternative because reinforcement can be concentrated along predefined load-transfer paths while retaining open areas between filaments.

C. Mechanical Limitations

The low tensile capacity of foam concrete remains one of its principal limitations. The porous structure reduces the effective load-bearing cross-section and can promote localized cracking.

Amran et al. showed that fiber addition can improve tensile and flexural behaviour in foamed concrete [6], while Lei et al. investigated carbon-fiber-reinforced foamed concrete and demonstrated that reinforcement can modify mechanical performance and failure mechanisms [8].

These studies establish the value of reinforcement in foam concrete. However, they do not resolve whether a continuous architected reinforcement system would provide greater reinforcement efficiency than randomly distributed fibers.

IV. CONVENTIONAL REINFORCEMENT OF FOAM CONCRETE

A. Discrete Fiber Reinforcement

Discrete fibers bridge cracks after the cementitious matrix begins to fracture. Depending on the fiber type and interface, the resulting mechanisms can include fiber debonding, pull-out, stretching, and rupture.

Daneti et al. investigated polypropylene fibers in lightweight concrete and demonstrated their influence on shrinkage cracking [7]. Lei et al. examined carbon

ISBN: 978-605-80193-6-2

fibers in foamed concrete and reported changes in mechanical performance and failure mechanisms [8].

Fiber reinforcement therefore provides an established method for improving the tensile and flexural performance of lightweight cementitious materials.

B. Advantages and Limitations

The principal advantage of discrete fibers is their distributed nature. Fibers can bridge cracks at multiple locations without requiring predefined reinforcement placement.

However, their effectiveness depends on dispersion and orientation. Fiber orientation is influenced by mixing and casting, while the effective embedment length varies among individual fibers.

The interface between fiber and cementitious matrix is also critical. Recent work on fiber surface treatment demonstrates that modification of the reinforcement surface can improve the interaction between reinforcement and cementitious materials [26].

The main limitation of discrete reinforcement from an architectural perspective is therefore the limited ability to prescribe the exact reinforcement location and orientation.

C. Transition Toward Continuous Reinforcement

Continuous reinforcement changes the reinforcement problem from statistical distribution to geometric design.

A printed mesh can be positioned at predefined locations and oriented to cross expected crack planes. The reinforcement geometry can therefore become an explicit design variable.

This transition is particularly relevant to foam concrete, where reinforcement may need to be positioned strategically because the low-density matrix cannot provide the same confinement and load transfer as dense concrete.

V. 3D-PRINTED POLYMER REINFORCEMENT

A. Architected Reinforcement

Additive manufacturing enables polymer reinforcement to be fabricated in geometries that are difficult to produce using conventional manufacturing methods.

Salazar et al. demonstrated the use of 3D-printed polymer lattices in ultra-high-performance concrete and reported substantial ductility enhancement [10]. Their results showed that the architecture and orientation of printed filaments relative to expected tensile stresses affected the mechanical response. Salazar et al. demonstrated that 3D-printed polymer lattices can substantially enhance the ductility of cementitious composites [10], as illustrated in Fig. 1.

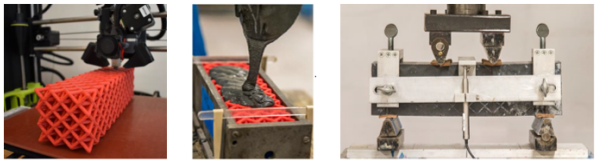


Fig. 1. Fabrication of lattice-reinforced concrete beams using 3D-printed polymeric lattices [10].

Katzer et al. investigated 3D-printed plastic gyroid reinforcement in concrete and demonstrated that reinforcement topology can be used to modify the mechanical response [11].

Namakiaraghi et al. further demonstrated that changing reinforcement micro-architecture and surface features can influence ductility and bond behaviour [12].

These studies establish the feasibility of architected reinforcement but do not directly establish its optimum design for foam concrete.

B. Mesh and Lattice Geometry

The geometry of printed reinforcement can be described by filament diameter, aperture size, open-area ratio, reinforcement volume, topology, and spatial arrangement.

A square mesh provides a relatively simple and reproducible architecture. More complex lattices, such as gyroid structures, provide three-dimensional load paths but introduce greater manufacturing and characterization complexity. Katzer et al. demonstrated that different 3D-printed gyroid reinforcement topologies can modify the mechanical response of concrete, as illustrated in Fig. 2 [11].



Fig. 2. 3D models of gyroid spatial elements with different reinforcement densities [11].

Recent reviews of FDM-based reinforcement emphasize that reinforcement topology is not merely a geometric detail but an important structural design variable [23], [24].

For foam concrete, the optimum geometry should not necessarily be the geometry producing the highest absolute strength. Instead, reinforcement efficiency should be evaluated relative to density, polymer consumption, and thermal performance.

C. Crack Bridging and Post-Cracking Behaviour

The principal contribution of continuous polymer reinforcement is expected after matrix cracking.

A properly anchored mesh may bridge cracks, redistribute tensile stresses, delay crack opening, and increase residual load capacity.

ISBN: 978-605-80193-6-2

However, the reinforcement contribution depends on the competition between polymer deformation, interface failure, matrix cracking, and reinforcement rupture. A polymer with high tensile capacity cannot provide its full capacity if premature pull-out occurs.

Therefore, reinforcement tensile strength alone is insufficient for assessing the effectiveness of a printed polymer mesh.

D. Reinforcement Quantity and Placement

Reinforcement quantity and placement should be treated as independent variables.

Increasing polymer volume may increase load-bearing capacity but also increase density and material consumption. Similarly, a mesh placed near a tensile surface may have a different flexural contribution from the same mesh positioned at mid-depth.

Consequently, experiments that simultaneously change reinforcement quantity and spatial placement cannot uniquely identify the contribution of either parameter.

This is one of the most important unresolved issues for foam-concrete applications.

VI. PET-G AND FDM-BASED REINFORCEMENT

A. PET-G as a Candidate Material

PET-G is a thermoplastic polymer suitable for FDM processing and can be fabricated into controlled mesh and lattice architectures.

However, its suitability as reinforcement should be considered in terms of the printed component rather than the raw filament. FDM processing introduces layer interfaces and manufacturing-induced anisotropy that may substantially influence mechanical behaviour.

Therefore, PET-G should be regarded as a promising candidate rather than a universally validated reinforcement material for foam concrete.

B. Influence of FDM Parameters

Ergene et al. investigated PET-G components produced by FDM and demonstrated the influence of layer height and test temperature on impact behaviour [15]. Lopes et al. examined the influence of microstructure pattern and infill density on PET-G properties [16]. Zisopol et al. investigated FDM parameters affecting the compressive behaviour of PET-G [17].

Bembenek et al. also demonstrated that processing parameters influence the tensile behaviour of FDM-produced PET-G and PLA [18].

These studies show that printing parameters are engineering variables rather than merely manufacturing details. Post-processing conditions can also modify the mechanical properties of FDM-produced rPET-G components, indicating that the final reinforcement behaviour may depend not only on printing parameters but also on subsequent treatment [28].

C. Printing-Induced Anisotropy

The layer-by-layer manufacturing process can produce anisotropic mechanical behaviour.

Strength parallel to deposited filaments may differ from strength perpendicular to layer interfaces. For a mesh, this anisotropy can influence tensile capacity, filament intersections, crack bridging, and failure mode.

The broader FDM literature confirms that process parameters and printing strategy can substantially influence the mechanical performance of printed polymer components [18], [29].

Accordingly, a printed PET-G reinforcement should be described using at least:

- filament diameter;
- printing temperature;
- layer height;
- printing orientation;
- infill strategy;
- raster direction; and
- relevant post-processing conditions.

D. Surface Morphology

FDM also generates a characteristic surface morphology. Individual deposited layers create geometric irregularities that may influence mechanical interlocking with the cementitious matrix.

This phenomenon is particularly important because a manufacturing feature that is unfavorable for the intrinsic tensile strength of a printed polymer may simultaneously improve its bond with cementitious material.

Bol et al. demonstrated that FDM-induced geometrical features can enhance reinforcement-matrix bond through interlocking [19]. Their study reported a bond-strength increase of approximately 56% for the investigated printing configuration.

Therefore, printing-induced morphology should not simply be treated as a defect; its structural function should also be considered.

VII. POLYMER-CEMENT INTERFACE

A. Bond Mechanisms

The performance of polymer reinforcement depends strongly on stress transfer across the interface.

Potential mechanisms include adhesion, friction, mechanical interlocking, surface roughness, geometric confinement, and anchorage.

The contribution of each mechanism depends on polymer type, surface morphology, matrix composition, reinforcement geometry, and loading direction.

B. Printing Direction

Printing direction can affect both the mechanical anisotropy of the polymer and the morphology of the interface.

ISBN: 978-605-80193-6-2

Bol et al. demonstrated that the directional morphology of FDM-produced polymer edges influences bond behaviour [19]. Their work specifically showed that the elliptical morphology generated at printed polymer edges can contribute to mechanical interlocking and affect bond strength. Bol et al. demonstrated that the morphology of FDM-printed polymer edges influences bond behaviour through mechanical interlocking, as shown in Fig. 3 [19].

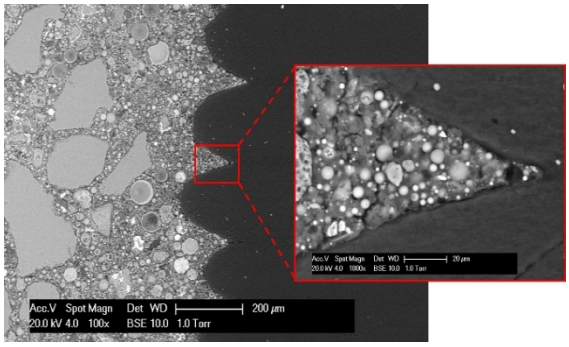


Fig. 3. SEM image of the interface between FDM 3D-printed polymer and cementitious mortar before failure [19].

This result has an important implication for PET-G mesh design: printing orientation should be treated as both a material-processing parameter and a structural-design parameter.

C. Pull-Out Behaviour

Pull-out testing provides a direct means of quantifying reinforcement-matrix interaction.

Important parameters include maximum pull-out force, bond stress, slip, post-peak response, embedded length, and failure mode.

For foam concrete, pull-out behaviour may differ from dense concrete because the reinforcement can be surrounded by thinner and more heterogeneous cementitious regions.

Future studies should therefore combine pull-out tests with surface characterization and microscopic examination.

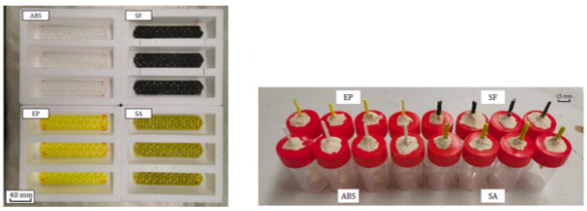
D. Reinforcement Geometry and Interface

Mesh aperture and filament spacing influence the amount of cementitious material surrounding the reinforcement and the degree of geometric confinement.

A smaller aperture may provide more crack-bridging elements per unit area but also increase polymer consumption and potentially interfere with matrix continuity. A larger aperture reduces polymer consumption but may provide fewer reinforcement crossings.

The optimum geometry should therefore be established by considering both mechanical performance and interfacial efficiency.

Surface modification represents another potential strategy. Xu et al. demonstrated that improving the bond between 3D-printed polymer reinforcement and a cementitious matrix can significantly enhance the mechanical response of the resulting composite [20]. Xu et al. demonstrated that surface modification of 3D-printed polymeric reinforcement can improve bond strength and mechanical performance, as illustrated in Fig. 4 [20].



(a) (b)

Fig. 4. Printed and surface-modified 3D-printed polymeric lattice reinforcements and bars used for pull-out testing [20].

In their study, surface treatments involving epoxy, sand, and steel fibers were used to increase bond strength.

VIII. DURABILITY AND LONG-TERM PERFORMANCE

A. Alkaline Exposure

Cementitious materials provide an alkaline environment that may influence polymers over long service periods.

Santana et al. investigated accelerated alkaline attack of polymers used for 3D printing and assessed their durability in geopolymer-based matrices [21].

Such studies provide useful evidence concerning polymer stability. However, accelerated alkaline exposure in a geopolymer-related environment should not automatically be interpreted as proof of long-term PET-G durability in Portland-cement-based foam concrete.

Long-term studies should investigate polymer behaviour under representative cementitious pore solutions and realistic moisture and temperature conditions.

B. Freeze-Thaw Behaviour

Freeze-thaw exposure is particularly relevant to porous materials because water stored in pores can contribute to internal damage during repeated freezing and thawing.

Chen et al. investigated the strength deterioration of foamed concrete under freeze-thaw cycles and demonstrated that repeated exposure can deteriorate mechanical performance [22].

Continuous polymer reinforcement may restrain crack development, but its effectiveness will depend on mesh architecture, bond quality, saturation condition, and reinforcement location.

Freeze-thaw resistance should therefore be treated separately from alkaline durability.

C. Long-Term Research Needs

The addition of polymer reinforcement introduces additional durability variables, including polymer degradation, interface deterioration, moisture interaction, and changes in mechanical properties over time.

ISBN: 978-605-80193-6-2

Future studies should evaluate alkaline exposure, wet-dry cycling, freeze-thaw cycling, thermal cycling, moisture effects, sustained loading, and retention of polymer-matrix bond.

IX. CRITICAL COMPARISON AND FUTURE RESEARCH DIRECTIONS

The most important contribution of the reviewed literature is not the demonstration that 3D-printed polymer reinforcement can improve cementitious composites; this has already been demonstrated in several dense cementitious systems. The more important question is how far these findings can be transferred to foam concrete.

TABLE I summarizes representative studies on 3D-printed polymer reinforcement and their relevance to foam concrete.

Study	Matrix / Reinforcement	Architectural variable	Key finding	Relevance to foam concrete
Salazar et al. [10]	Concrete / 3D-printed polymer lattice	Lattice geometry	Improved ductility	Requires validation in porous matrices
Katzer et al. [11]	Concrete / 3D-printed gyroid	Topology	Topology affected mechanical response	Limited transferability to foam concrete
Namakiaraghi et al. [12]	Cement-based / 3D-printed polymer	Reinforcement geometry	Architecture enabled mechanical control	Application to foam concrete remains limited
Bol et al. [19]	Cementitious / 3D-printed polymer	Printing direction	Interface response depended on print direction	Foam-concrete interface not addressed
Lei et al. [8]	Foamed concrete / carbon fibers	Fiber content	Reinforcement modified mechanical response	Discrete, not architected reinforcement
Daneti et al. [7]	Lightweight concrete / PP fibers	Fiber distribution	Influenced cracking behaviour	Less controllable than printed meshes
Santana et al. [21]	Geopolymer / 3D-printed polymers	Polymer type / exposure	Highlighted alkaline durability	Geopolymer conditions differ from foam concrete

The table reveals a clear literature asymmetry. Foam-concrete studies provide substantial evidence concerning the effects of pores and discrete reinforcement, whereas 3D-printed reinforcement studies provide stronger evidence concerning architectural control, bond, and crack bridging. Very few studies connect these two evidence bases directly.

A. Dense-Matrix Evidence Cannot Be Directly Transferred

The first major research gap concerns matrix type.

The ductility improvements reported by Salazar et al. were demonstrated in an ultra-high-performance concrete matrix [10]. Similarly, the architected reinforcement studies of Namakiaraghi et al. and Katzer et al. were conducted using relatively dense cementitious matrices [11], [12].

Such studies establish mechanisms but do not define the optimum reinforcement architecture for foam concrete.

In foam concrete, the reinforcement is embedded in a cellular matrix with local variations in pore-wall thickness and matrix continuity. The effective confinement of the reinforcement may therefore be lower, while local cracking may be governed by pore geometry rather than only by matrix fracture.

Accordingly, future studies should explicitly distinguish mechanism transfer from performance transfer.

B. Reinforcement Quantity and Placement

The second major gap is the independent influence of reinforcement amount and spatial position.

The number of meshes , polymer volume, mesh aperture, and mesh location can all affect mechanical response. If these variables are changed simultaneously, the resulting improvement cannot be attributed uniquely to one parameter.

A controlled experimental design should therefore isolate:

- reinforcement volume;
- mesh aperture;
- filament diameter;
- number of meshes; and
- vertical position.

This is particularly important in flexural applications, where the distance of the reinforcement from the tensile face strongly affects its contribution. Related work on lightweight concrete beams strengthened with ultra-lightweight ECC-FRP mesh composites also indicates the potential of continuous mesh-type reinforcement systems for lightweight cementitious materials [30]. However, such evidence remains indirect for FDM-produced PET-G reinforcement and does not resolve the specific effects of printed polymer architecture in foam concrete.

C. Mesh Geometry Optimization

The literature demonstrates that reinforcement topology influences performance. However, no universal optimum architecture has been established.

Square meshes are attractive because of their simple manufacturing and reproducibility. Gyroid and other three-dimensional lattices may provide more complex load paths but require greater material and manufacturing complexity.

ISBN: 978-605-80193-6-2

For foam concrete, geometry should be optimized according to:

- mechanical strength;
- crack bridging;
- polymer consumption;
- density;
- thermal conductivity;
- interface quality; and
- durability.

Recent state of the artwork on FDM reinforcement similarly emphasizes the importance of reinforcement architecture and topology in controlling structural performance [23]. The review by Momeni et al. also emphasizes the distinction between 3D-printed concrete and concrete reinforced using 3D-printed systems [24].

D. Density-Normalized Performance

Absolute strength is insufficient for evaluating reinforcement efficiency in lightweight composites.

A reinforcement system that increases strength by substantially increasing density may provide less benefit than a lighter system with a smaller absolute strength increase.

A useful parameter is the density-normalized strength:

$$\eta c = \frac{fc}{\rho}$$

where $f_c$is compressive strength and $\rho$is density.

Similarly, flexural efficiency may be expressed as:

$$\eta_f = \frac{f_f}{\rho}$$

where $f_f$is flexural strength.

These indicators should not replace absolute strength values, but they can provide a more meaningful basis for comparing lightweight reinforcement systems.

Future studies should therefore report both absolute and density-normalized properties.

E. Mechanical-Thermal Trade-Off

Foam concrete is frequently selected because of its low density and thermal-insulation capability.

Consequently, reinforcement should not be optimized using mechanical strength alone. Polymer volume and mesh architecture may influence thermal conductivity and density.

A mechanically optimal reinforcement configuration may therefore not be optimal when thermal performance is included.

This multi-property trade-off is currently insufficiently investigated.

F. Interface Characterization

The interface represents one of the most promising areas for future research.

The findings of Bol et al. demonstrate that FDM-induced morphology can affect bond behaviour [19], while Xu et al. demonstrate the potential of surface modification to enhance reinforcement-matrix interaction [20].

For foam concrete, however, the interface may be affected by local pore formation around the reinforcement.

Future studies should combine:

- pull-out testing;
- bond stress-slip analysis;
- digital image correlation;
- surface profilometry;
- microscopy; and
- numerical interface modelling.

G. Durability

The durability evidence remains fragmented.

Freeze-thaw behaviour is relatively well established for foam concrete [22], while alkaline exposure provides information concerning polymer stability [21]. However, the combined long-term behaviour of PET-G reinforcement and foam concrete under simultaneous moisture, thermal, and mechanical exposure remains largely unresolved.

H. Sustainability

The environmental implications of printed polymer reinforcement also require consideration.

Future research should compare:

- polymer consumption;
- printing energy;
- manufacturing waste;
- recyclability;
- service life;
- reinforcement efficiency; and
- alternatives such as discrete fibers and conventional reinforcement.

The sustainability advantage of lightweight polymer reinforcement should therefore be demonstrated through a life-cycle perspective rather than assumed from low density alone.

X. FUTURE RESEARCH FRAMEWORK

Based on the reviewed literature, future research should follow a multiscale framework.

At the material scale, the printed polymer should be characterized in terms of tensile behaviour, anisotropy, printing parameters, and surface morphology.

ISBN: 978-605-80193-6-2

At the interface scale, bond strength, slip, embedment length, and failure mode should be quantified.

At the matrix scale, foam-concrete density, pore size distribution, pore connectivity, and matrix strength should be characterized.

At the structural scale, mesh geometry, reinforcement quantity, spatial placement, and crack-bridging behaviour should be investigated.

Finally, at the performance scale, mechanical efficiency should be evaluated together with density, thermal conductivity, durability, and material consumption.

The framework is consistent with recent roadmaps for FDM-based reinforcement, which identify reinforcement architecture, material selection, manufacturing parameters, and structural performance as interconnected design variables [23].

TABLE II. Priority research gaps and recommended approaches

Research scale	Matrix	Reinforcement	Research gap	Recommended approach	Key outcome
Material	Foam concrete	FDM polymer	Polymer anisotropy and print variability	Test print orientation and parameters	Strength and anisotropy
Interface	Foam concrete	PET-G mesh	Bond and pull-out mechanisms	Vary embedment length and mesh	Bond strength and failure mode
Matrix	Foam concrete	PET-G mesh	Limited pore–reinforcement data	Characterize density and pore structure	Matrix–reinforcement interaction
Structural	Foam concrete	3D-printed mesh	Geometry, quantity and placement effects	Controlled parametric testing	Crack bridging and load capacity
Performance	Foam concrete	PET-G mesh	Limited multi-property assessment	Assess strength, density, thermal and	Overall material efficiency

This framework indicates that future research should move from proof-of-concept demonstrations toward controlled optimization.

XI. CONCLUSIONS

This review critically examined continuous 3D-printed polymer reinforcement in cementitious composites, with particular emphasis on its potential application to foam concrete.

First, foam concrete has well-established advantages in terms of low density and thermal performance, but its cellular structure limits tensile and flexural capacity and contributes to brittle cracking. Discrete fiber reinforcement can mitigate these limitations, although its spatial distribution and orientation are not directly controlled.

Second, studies on dense cementitious materials demonstrate that architected polymer reinforcement can improve ductility, crack bridging, energy absorption, and post-cracking behaviour. However, these results should be interpreted as mechanistic evidence rather than directly transferable performance data for foam concrete.

Third, FDM provides an important design opportunity because reinforcement geometry, orientation, aperture, and placement can be controlled digitally. PET-G is a promising candidate for such reinforcement, but the properties of the printed component depend on FDM parameters and should not be inferred solely from raw-filament data.

Fourth, the polymer-cement interface is a critical determinant of reinforcement efficiency. Recent evidence demonstrates that FDM-induced surface morphology can contribute to mechanical interlocking and substantially influence bond behaviour [19]. Surface modification can also improve bond strength and consequently enhance composite ductility [20].

Fifth, the most important unresolved issue is the interaction between continuous printed reinforcement and the cellular pore structure of foam concrete. Reinforcement quantity, mesh geometry, spatial placement, interface behaviour, density, thermal conductivity, and durability should be investigated as interconnected variables rather than independently.

The literature therefore supports a clear research opportunity: continuous 3D-printed polymer reinforcement may provide a controllable and potentially material-efficient reinforcement architecture for foam concrete, but its effectiveness must be established specifically within the porous matrix rather than inferred from dense concrete systems.

Future studies should therefore prioritize independent control of reinforcement quantity and placement, quantitative interface characterization, density-normalized performance, mesh-geometry optimization, and simultaneous evaluation of mechanical, thermal, durability, and sustainability performance.

ACKNOWLEDGMENT

The authors have no acknowledgments to declare.

References

[1] O. Gencel, T. Bilir, Z. Bademler, and T. Ozbakkaloglu, “A detailed review on foam concrete composites: Ingredients, properties, and microstructure,” *Applied Sciences*, vol. 12, no. 11, Art. no. 5752, 2022.

[2] Y. Fu, L. Wang, C. Wang, and H. Li, “Foam concrete: A state-of-the-art and state-of-the-practice review,” *Advances in Materials Science and Engineering*, Art. no. 6153602, 2020.

ISBN: 978-605-80193-6-2

[3] S.-Y. Chung, J.-S. Kim, C. Lehmann, D. Stephan, T.-S. Han, and M. Abd Elrahman, "Investigation of phase composition and microstructure of foamed cement paste with different supplementary cementing materials," *Cement and Concrete Composites*, vol. 109, Art. no. 103560, 2020.

[4] Y. Bie, S. Ba, and S. Chen, "Studies on foamed concrete micropores and their effects on stress distribution and heat conduction," *Journal of Building Engineering*, vol. 87, Art. no. 109152, 2024.

[5] D. Falliano, D. De Domenico, G. Ricciardi, and E. Gugliandolo, "Key factors affecting the compressive strength of foamed concrete," *IOP Conference Series: Materials Science and Engineering*, 2018.

[6] M. Amran, R. Fediuk, N. Vatin, Y. H. Lee, G. Murali, T. Ozbakkaloglu, S. Klyuev, and H. Alabduljabber, "Fiber-reinforced foamed concretes: A review," *Materials*, vol. 13, no. 19, Art. no. 4323, 2020.

[7] S. B. Daneti, T.-H. Wee, and T. T. Tso, "Effect of polypropylene fibers on the shrinkage cracking behaviour of lightweight concrete," *Magazine of Concrete Research*, vol. 63, no. 11, pp. 871-881, 2011.

[8] Z. Lei, Q. Li, Y. Zhou, and B. Yin, "Mechanical performance and reinforcing mechanisms of foamed concrete strengthened by carbon fibers," *Journal of Building Engineering*, vol. 97, Art. no. 110765, 2024.

[9] P. Zandifaez, Z. Shen, R. Sorgenfrei, Y. Li, D. Dias-da-Costa, and H. Zhou, "Pathways to formulate lightweight and ultra-lightweight 3D printable cementitious composites," *Construction and Building Materials*, vol. 424, Art. no. 135874, 2024.

[10] B. Salazar, P. Aghdasi, I. D. Williams, C. P. Ostertag, and H. K. Taylor, "Polymer lattice-reinforcement for enhancing ductility of concrete," *Materials & Design*, vol. 196, Art. no. 109184, 2020.

[11] J. Katzer, A. Skoratko, T. Szatkiewicz, and O. Sucharda, "Shaping mechanical properties of concrete elements by 3D-printed plastic gyroid reinforcing elements," *Structural Concrete*, 2025.

[12] P. Namakiaraghi, A. Sadighi, R. Spragg, A. R. Najafi, and Y. A. Farnam, "Towards development of cement-based composites reinforced with architected 3D-printed polymers," *Construction and Building Materials*, vol. 422, Art. no. 135838, 2024.

[13] T. Ding, H. Dong, P. Sikora, and G. Lin, "3D printed concrete reinforced with flexible fiber reinforced polymer strips or grids: Concept and bond tests," *Structures*, 2025.

[14] H.-Q. Sun, S.-S. Xie, J.-J. Zeng, and Y. Zhuge, "3D-printed concrete with in-process embedded fiber-reinforced polymer grid reinforcement," *Communications Engineering*, vol. 5, Art. no. 72, 2026.

[15] B. Ergene, H. Ispartalı, and U. Karakılınç, "Impact behaviour of PET-G parts produced by fused deposition modelling depending on layer height and test temperature," *Journal of the Faculty of Engineering and Architecture of Gazi University*, vol. 38, no. 3, pp. 1345-1359, 2023.

[16] L. Lopes, D. Reis, M. Paula Junior, and M. Almeida, "Influence of 3D microstructure pattern and infill density on the mechanical and thermal properties of PET-G filaments," *Polymers*, vol. 15, no. 10, Art. no. 2268, 2023.

[17] D. G. Zisopol, M. Minescu, and D. V. Iacob, "A study on the influence of FDM parameters on the compressive behaviour of PET-G parts," *Engineering, Technology & Applied Science Research*, vol. 14, no. 2, pp. 13592-13597, 2024.

[18] M. Bembenek, Ł. Kowalski, and A. Kosoń-Schab, "Research on the influence of processing parameters on the specific tensile strength of FDM additive manufactured PET-G and PLA materials," *Polymers*, vol. 14, no. 12, Art. no. 2446, 2022.

[19] R. J. M. Bol, Y. Xu, M. Luković, and B. Šavija, "Does printing direction influence the bond between 3D printed polymeric reinforcement and cementitious matrix?" *Engineering Failure Analysis*, vol. 174, Art. no. 109471, 2025, doi: 10.1016/j.engfailanal.2025.109471.

[20] Y. Xu, Z. Wan, and B. Šavija, "Elevating mechanical performance of cementitious composites with surface-modified 3D-printed polymeric reinforcements," *Developments in the Built Environment*, vol. 19, Art. no. 100522, 2024, doi: 10.1016/j.dibe.2024.100522.

[21] H. A. Santana, J. S. Andrade Neto, D. V. Ribeiro, M. S. Cilla, and C. M. Dias, "Accelerated alkaline attack of 3D printing polymers to assess their durability in geopolymer-based matrices," *Journal of Materials in Civil Engineering*, vol. 33, no. 11, Art. no. 04021327, 2021.

[22] J. Chen, B. Chen, S. Qiang, and S. Chen, "Study on strength deterioration mechanism of foamed concrete under freeze-thaw cycles: Experiment and numerical simulation," *Construction and Building Materials*, vol. 438, Art. no. 137083, 2024.

[23] M. Hematibahar, R. Fediuk, K. Momeni, M. Kharun, A. Bhowmik, and V. Romanovski, "Strategic roadmap for 3D-printed reinforcement using fused deposition modeling: A state-of-the-art review," *Engineering Reports*, vol. 7, no. 6, Art. no. e70232, 2025, doi: 10.1002/eng2.70232.

[24] K. Momeni, N. I. Vatin, M. Hematibahar, and T. H. G. Gebre, "Differences between 3D printed concrete and 3D printing reinforced concrete technologies: A review," *Frontiers in Built Environment*, vol. 10, Art. no. 1450628, 2025.

[25] A. A. Mohammed, H. O. Aziz, and H. F. Mahmood, "A comprehensive review of foam concrete and its applications in civil engineering," *Discover Concrete and Cement*, vol. 2, Art. no. 34, 2026.

[26] J. Viradiya and R. Gupta, "The role of fiber surface treatment on improving mechanical performance of cementitious composites," *Construction and Building Materials*, vol. 469, Art. no. 140452, 2025.

ISBN: 978-605-80193-6-2

[27] A. S. Abdulazeez and A. A. Sani, "Properties of foam concrete: A review," *Discover Concrete and Cement*, vol. 2, Art. no. 4, 2026.

[28] D. Lozrt, J. Votava, R. Henzl, V. Kumbár, P. Dostál, and J. Čupera, "Analysis of the changes in the mechanical properties of the 3D printed polymer rPET-G with different types of post-processing treatments," *Applied Sciences*, vol. 13, no. 16, Art. no. 9234, 2023.

[29] D. Popescu, A. Zapciu, C. Amza, F. Baciu, and R. Marinescu, "FDM process parameters influence over the mechanical properties of polymer specimens: A review," *Polymer Testing*, vol. 69, pp. 157-166, 2018.

[30] Z. Hu, M. Elchalakani, R. Hassanli, H. Ran, M. A. Sadakkathulla, and S. Nie, "Structural health monitoring of lightweight concrete beams strengthened with ultra-lightweight ECC-FRP mesh composite," *Engineering Structures*, vol. 322, Art. no. 119224, 2025.

ISBN: 978-605-80193-6-2

Bird Strike Investigation of Bio-Inspired Laminated Composites with Different End Conditions and Cutouts

Meryem Dilara KOP[1] Aman GARG[2a,2b] Mehmet AVCAR[3]

[1]Suleyman Demirel University, Graduate School of Natural and Applied Sciences, Isparta, Türkiye

[2a]Shantou University, Faculty, Department of Civil Engineering and Smart Cities, Shantou, Guangdong 515063 PR China

[2b]The NorthCap University, Faculty, Department of Multidisciplinary Engineering, Gurugram, Haryana 122017, India

[3]Suleyman Demirel University, Department of Civil Engineering, Isparta, Türkiye

Article note: This study is derived from the ongoing Ph.D. thesis of the first author in the Department of Civil Engineering at Suleyman Demirel University.

* Corresponding author e-mail: dilara.kop@hotmail.com

***Abstract*— Bio-inspired laminated composite materials are increasingly being used in various engineering fields. This study investigates the strain energy capacity and damage dissipation performance of bio-inspired helicoidal laminated composite plates in comparison with conventional quasi-isotropic composite layups, considering different boundary conditions and cutout diameters, using a hybrid Smoothed Particle Hydrodynamics–Finite Element Analysis method under bird strike. Three stacking sequences of bio-inspired helicoidal laminated composite plates are examined: linear helicoidal, recursive helicoidal, and semi-circular helicoidal. The main novelty of the study lies in the combined investigation of cutout diameter and boundary condition effects on the strain energy capacity and damage dissipation behavior of these bio-inspired helicoidal configurations, an interaction that has not been systematically addressed in previous studies on bio-inspired composite structures. Abaqus Explicit software is used for analysis. The impactor is modeled by smoothed particle hydrodynamic method and plate is modeled by finite element analysis. For failure Hashin failure criteria is employed. Two cutout diameters, 100 mm and 200 mm, are considered under various combinations of clamped, simply supported, and free boundary conditions. Helicoidal layup showed better impact resistance against the conventional quasi-isotropic layups. The results show that the helicoidal layups exhibit better impact resistance than the conventional quasi-isotropic layup. These findings provide critical insights and practical design guidelines for developing impact-resistant and damage-tolerant composite structures subjected to foreign object impact, supporting the development of more efficient and reliable bio-inspired composite designs for aerospace and related engineering applications.**

***Keywords*—Bird strike, bio-inspired laminated composite, SPH-FEA**

The first author is supported by th Scientific and Technological Research Council of Türkiye (TÜBİTAK) under the 2211-A Ntional Ph.D Scholarship Program

I. INTRODUCTION

The worldwide load effect on the aerospace, aeronautics, and automotive engineering fields has critical importance. During service life, especially aerospace and aeronautical materials have to be exposed to dynamic and static loading. Dynamic loadings can be hard body impact such as ballistic impacts or soft body impact such as bird strike. Bird strikes not only cause financial damage amounting to billions of dollars annually but also endanger human lives. The extent of the damage depends on various parameters, such as the speeds of the bird and the aircraft at the moment of impact. Most of the strike happened during the landing/take-off process [1]. To overcome the weight disadvantage of conventional metallic alloys, carbon fiber-reinforced polymer (CFRP) composites have become widely used in primary structures. While CFRP composites substantially reduce structural weight, they provide superior specific strength and stiffness. Despite all these higher mechanical advantages, laminated composite materials are vulnerable to dynamic impact loadings, frequently suffering from severe matrix cracking, fiber breakage, and delamination. This critical weakness has driven researchers to explore novel material architectures and bio-inspired design strategies to enhance impact tolerance and energy dissipation. The nature provides a great source. Nature adapts structures to withstand various loading conditions without severe failure [1]. For understand the impact behavior of laminated composites different researchers was made studies. Zhao et al. [2] investigates bio-inspired CFRP laminates based on the sinusoidal structure found in trap-jaw ant mandibles. Zhang et al. [3] study on the combined sinusoidal and helicoidal fiber structure using basalt fibers structure to create laminate composites with higher impact toughness. The researchers found out that bio-inspired layup increased the impact toughness comparing the traditional laminated composite layup. Chew et al. [4] demonstrated that out-of-plane strength benefits markedly from helical layup, noting that a 9° helical layup reached 72% and 52% higher peak loads than cross-ply and quasi-isotropic layup laminates. Berhie et al. [5] conducted a study on bio-inspired laminated composite panels through both numerical and experimental methods. Li et al. [6] studied the quasi-static tensile and bending performance of bio-inspired suture-interlocked composite laminates. Garg et al. [7] investigated the damage behaviour of the sinusoidal corrugated bio-inspired helicoidal laminated composite plates under the 200 m/s bird strike with different boundary conditions. Kop and Garg [8] studied bio-inspired laminated composite plates under various strike angles, comparing the helicoidal layup with the

ISBN: 978-605-80193-6-2

conventional quasi-isotropic layup. Chen et al. [9] investigated the high-velocity impact behavior of Z-pinned composite laminates using both experimental testing and numerical modeling.

Although previous studies have highlighted the superior energy dissipation and damage tolerance of bio-inspired helicoidal laminates, the coupled effects of functional cutouts and various edge boundary conditions under dynamic bird strike events remain insufficiently explored. To address this gap, the present study investigates the strain energy capacity and damage dissipation of bio-inspired helicoidal CFRP plates (linear, recursive, and semi-circular) with cutout diameters of 100 mm and 200 mm under various boundary conditions using a coupled SPH-FEA approach in Abaqus/Explicit.

II. MATERIAL AND METHODS

Numerical study was conducted using the finite element software Abaqus/Explicit. Bird performed different impact behaviour from the hard body impact. Depending on this reason SPH method was used to modelled the bird. Plates was modelled using the Finite element. Bird was modelled revolved solid cylinder. The height and diameter are 200 mm and 100 mm. Plate was modelled shell planar. 1m x 1m 64 ply with two cutouts diameters are100mm and 200 mm. 4 different layup were investigated in this study. These are linear helicoidal, helicoidal recursive, helicoidal semi-circular and conventional quasi-isotropic. Stacking sequences of helicoidal laminated composites are:

Linear helicoidal: $\theta_1/\theta_2/\theta_3 .../\theta_n = (n-1)\xi$

Helicoidal recursive: $\theta_1/\theta_2/\theta_3 .../\theta_n = \theta_{n-1} + \alpha(n-1)$

Helicoidal semi: $\theta_1/\theta_2/\theta_3 .../\theta_n = \sqrt{\theta^2{}_{max} - (\gamma(n-1) - \theta_{max})^2}$

For helicoidal semi-circular $(\theta_{max}, \gamma) = (45, 1.5)$, Linear helicoidal layup $\xi = 4$, and Helicoidal recursive α=1.Q

To ensure the accuracy and reliability of the numerical methodology, a validation study was conducted prior to the parametric analyses. Firstly, strain energy and time graph are compared with the study of Garg et al. [10].

Strike analyses was performed using frictionless general contact interaction within dynamic explicit approach. Birs was strike plate wit 100 m/s velocity. In this study 5 different boundary conditions were examined. Specifically, five edge boundary conditions namely fully clamped (CCCC), simply supported (SSSS), clamped-free-clamped-free (CFCF), clamped-clamped-free-free (CCFF), and clamped-free-free-free (CFFF)—were evaluated across both cutout configurations. For plate material CFRP properties used. Material properties taken from the study of Rahimian Koloor et al. [11]. The material properties of CFRP are given in the Table I.

Table I. Material properties of CFRP

Property	Value
E_1	105.5 GPa
$E_2 = E_3$	7.2 GPa
$G_{12} = G_{13}$	3.4 GPa
G_{23}	2.52 GPa
$\nu_{12} = \nu_{13}$	0.34
ν_{23}	0.378
ρ	1500 kg/m³

For falure investigation Hashin failure criteria was employed. Hashin failure parameters are givien in Table II.

Table II. Hashin failure parameters

Parameter	Value
X_T (Tensile strength, longitudinal)	1340 MPa
X_C (Compressive strength, longitudinal)	1192 MPa
Y_T (Tensile strength, transverse)	19.6 MPa
Y_C (Compressive strength, transverse)	92.3 MPa
S_{12} (Shear strength, in-plane)	51 MPa
S_{23} (Shear strength, out-of-plane)	23 MPa
G_{XT} (Fracture energy, longitudinal tensile)	48.4 N/mm
G_{XC} (Fracture energy, longitudinal compressive)	60.3 N/mm
G_{YT} (Fracture energy, transverse tensile)	4.5 N/mm
G_{YC} (Fracture energy, transverse compressive)	8.5 N/mm

For bird modelling material properties obtained from the study of Zhou et al. [12]

III. RESULTS AND DISCUSSIONS

In this study, the mechanical behaviour of CFRP bio-inspired helicoidal laminated composite plates—specifically, linear helicoidal (LH), helicoidal recursive (HR), and helicoidal semi-circular (HS) types—was examined under dynamic bird strike loads. This investigation involved comparing these bio-inspired designs with conventional quasi-isotropic (QI) laminated composite plates across five boundary conditions: CCCC, CCFF, CFCF, CFFF, and SSSS. Additionally, two circular cutout sizes, measuring 100 mm and 200 mm in diameter, were considered.

Firstly, strain energy-time relationships were investigated. The strain energy–time histories of the QI (Figs. 1 and 2), HR (Figs. 3 and 4), HS (Figs. 5 and 6) and LH (Figs. 7 and 8) plates with 100 mm and 200 mm diameter cutouts under the five boundary conditions shows similar behavior: the strain energy increase sharply during the initial contact phase, reaches a peak between approximately 0.002 s and 0.004s, and then decreases as part of the stored elastic energy is released and part is converted into damage dissipation energy. The boundary conditions effects the peak strain energy magnitudes more than to reach these magnitudes. For all layup and cutout combinations CFCF conditions gives the highest peak magnitude. It is 2070.83J for HR layup with 200 mm diameter cutouts. For all layup and cutouts SSSS boundary conditions gives the lowest peak strain energy.

ISBN: 978-605-80193-6-2

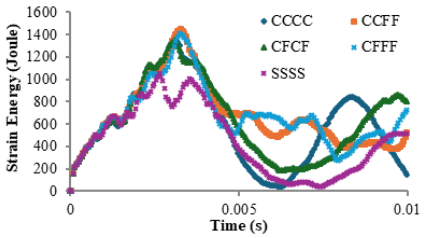


Fig. 1. Strain energy-time history of QI layup 100 mm diameter cutouts.

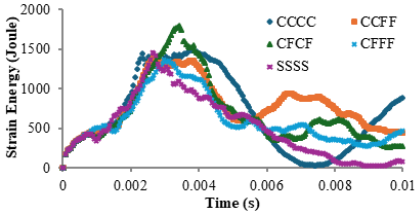


Fig. 2. Strain energy-time history of QI layup 200 mm diameter cutouts.

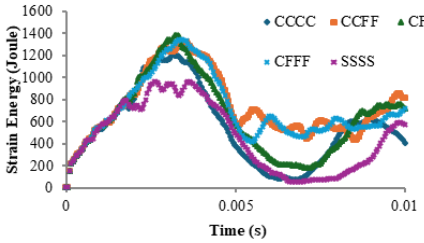


Fig. 3. Strain energy-time history of HR layup 100 mm diameter cutouts.

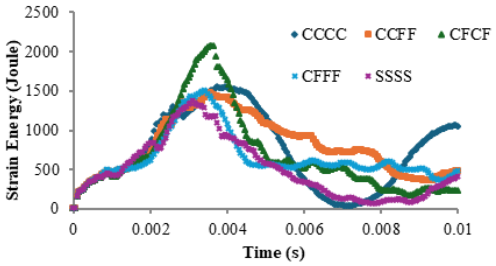


Fig. 4. Strain energy-time history of HR layup 200 mm diameter cutouts.

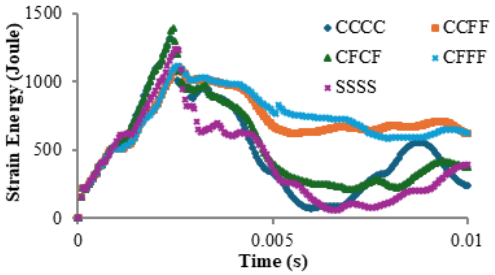


Fig. 5. Strain energy-time history of HS layup 100 mm diameter cutouts.

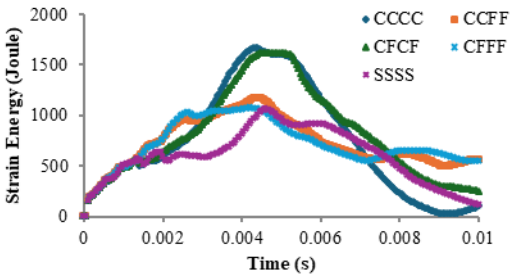


Fig. 6. Strain energy-time history of HS layup 200 mm diameter cutouts.

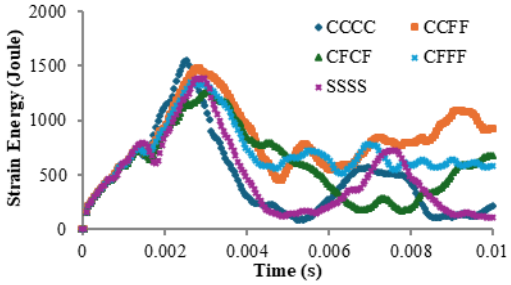


Fig. 7. Strain energy-time history of LH layup 100 mm diameter cutouts.

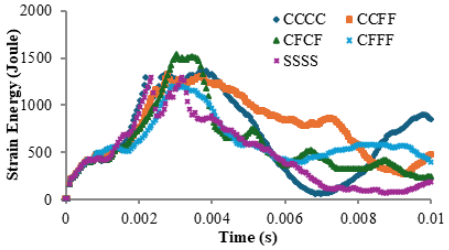


Fig. 8. Strain energy-time graph of LH layup 200 mm diameter cutouts.

After the strain energy-Time relationship, damage dissipation energy-time relationships were investigated. The damage dissipation energy – time graphs were demonstrated that QI (Figs. 9 and 10), HR (Figs. 11 and 12), HS (Figs. 13 and 14) and LH (Figs. 15 and 16) averaged over the five boundary conditions, the HS layup dissipates more energy than the QI layup at both cutout diameters, by approximately 5% for the 200 mm cutouts and 20% for the 100 mm diameter cutouts. The LH layup dissipates less energy than QI, by approximately 13% and 38% for the 200 mm and 100 mm cutouts, respectively. The HR layup behaves differently at the two diameters, dissipating about 10% less energy than QI for the 200 mm cutouts but about 12% more for the 100 mm cutouts. For the 200 mm cutouts, the CFCF condition gives the highest damage dissipation energy for all four layups, whereas no single boundary condition dominates for the 100 mm cutouts, indicating that the effect of the edge restraint on damage dissipation depends on the cutout size.

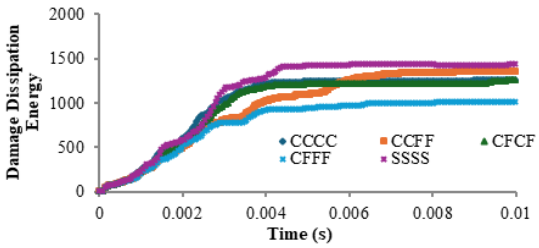


Fig. 9. Damage dissipation energy-time history of QI layup 100 mm diameter cutouts.

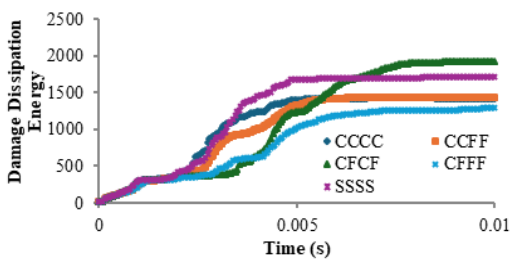


Fig. 10. Damage dissipation energy-time history of QI layup 200 mm diameter cutouts.

ISBN: 978-605-80193-6-2

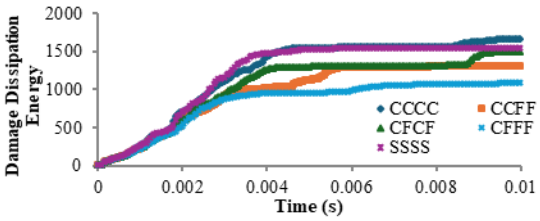


Fig. 11. Damage dissipation energy-time history of HR layup 100 mm diameter cutouts.

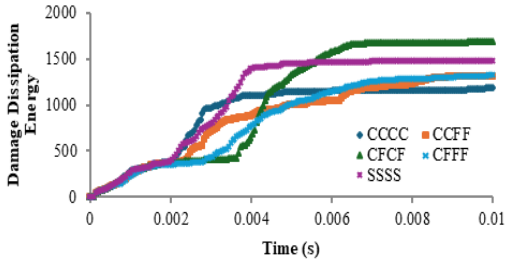


Fig. 12. Damage dissipation energy-time history of HR layup 200 mm diameter cutouts.

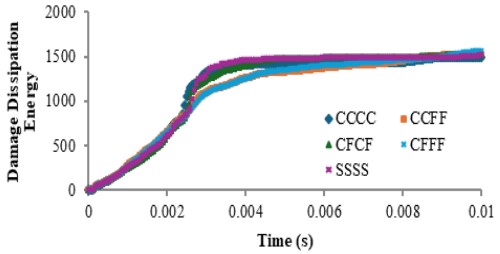


Fig. 13. Damage dissipation energy-time history of HS layup 100 mm diameter cutouts.

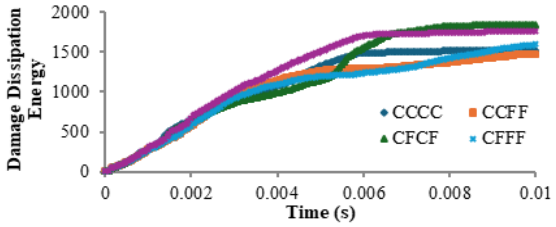


Fig. 14. Damage dissipation energy-time history of HS layup 200 mm diameter cutouts.

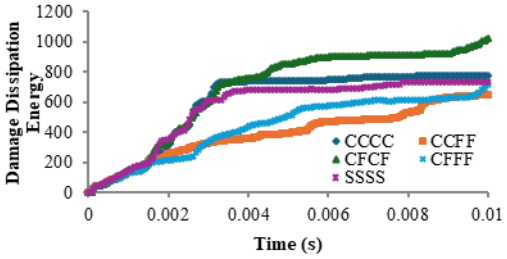


Fig. 15. Damage dissipation energy-time history of LH layup 100 mm diameter cutouts.

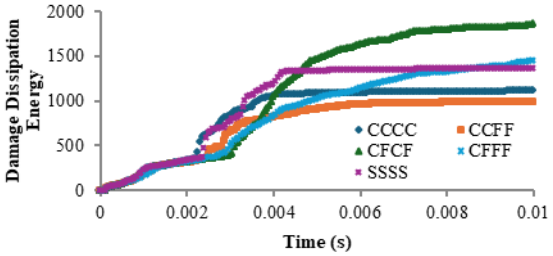


Fig. 16. Damage dissipation energy-time history of LH layup 200 mm diameter cutouts.

The failure modes evaluated through Hashin failure criteria provide physical explanations for these energy dissipation mechanisms. The DAMAGEFT and DAMAGEMT contours (Figs. 17–21 and 22–26) show that the four layup develop distinctly different damage morphologies, substantially determined by the cutouts's diameter. QI plates exhibit a four-armed, cross-shaped fracture pattern at a 100 mm diameter cutout, whereas a single, nearly linear crack between cutouts at a 200 mm diameter; CCCC and SSSS boundary conditions show slight secondary branching at 200 mm characteristics absent in CCFF, CFCF, and CFFF configurations. For HR layup, while CCCC and SSSS conditions produce noticeably longer crack branches extending toward the plate edges at 100 mm, CCFF, CFCF, and CFFF conditions confine the damage closer to the cuts. HS plates do not show such diameter dependence, instead forming numerous long, cross-fiber cracks across the majority of the plate at both diameters; CCCC and CFFF conditions create the most extensive and branched networks, whereas CCFF and SSSS concentrate the damage within a narrower band, However LH layup remain limited to a single, narrow crack across both diameters and all five boundary conditions. Overall, fully clamped (CCCC)and for HR plates, simply supported (SSSS) boundary conditions promote more extensive or branched crack propagation compared to free-edge conditions, though the specific outcome depends on the layup structure. In contrast, matrix tension damage does not show boundary condition dependency.

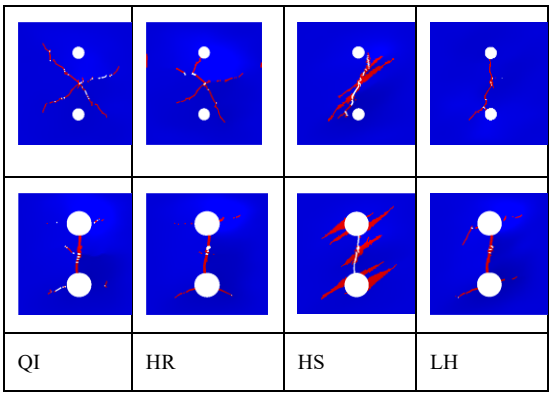


Fig. 17. DAMAGEFT contours for CCCC boundary condition.

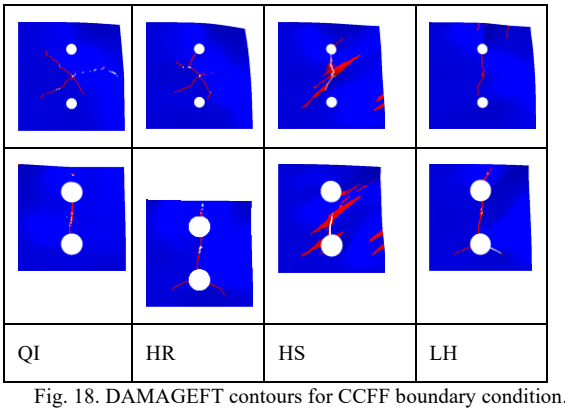


Fig. 18. DAMAGEFT contours for CCFF boundary condition.

ISBN: 978-605-80193-6-2

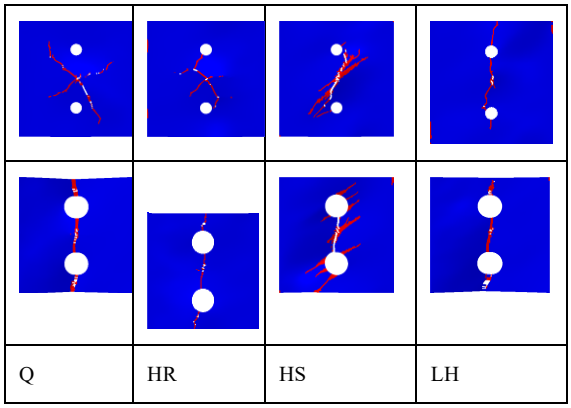


Fig. 19. DAMAGEFT contours for CFCF boundary condition.

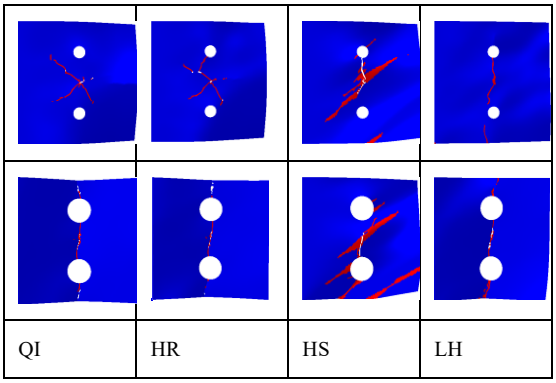


Fig. 20. DAMAGEFT contours for CFFF boundary condition.

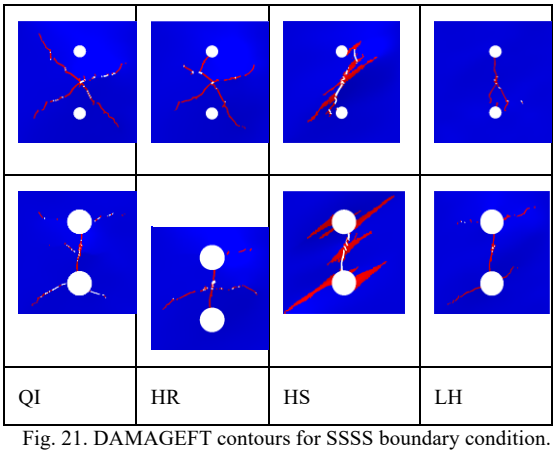


Fig. 21. DAMAGEFT contours for SSSS boundary condition.

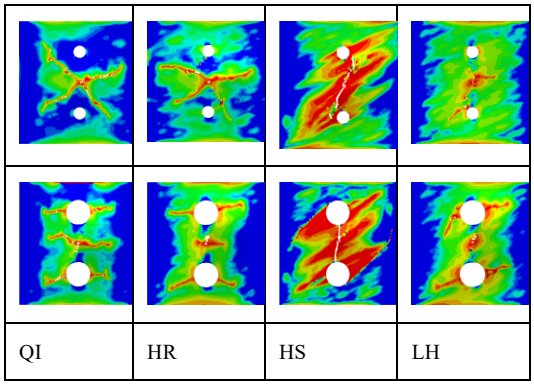


Fig. 22. DAMAGEMT contours for CCCC boundary condition.

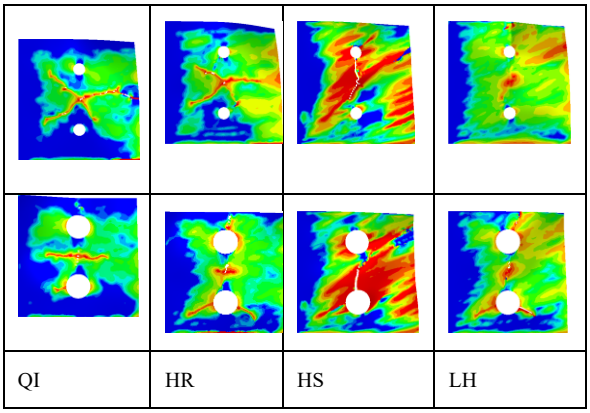


Fig. 23. DAMAGEMT contours for CCFF boundary condition.

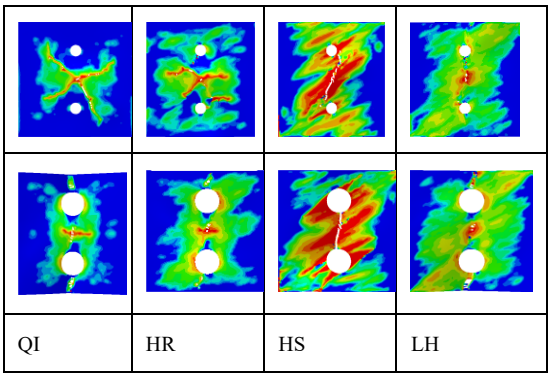


Fig. 24. DAMAGEMT contours for CFCF boundary condition.

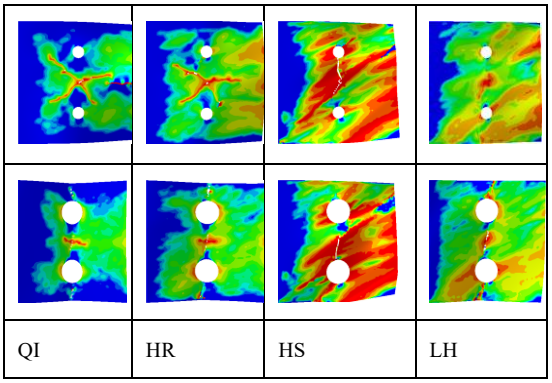


Fig. 25. DAMAGEMT contours for CFFF boundary condition.

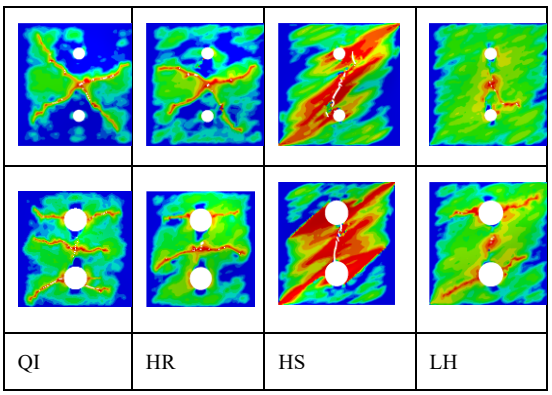


Fig. 26. DAMAGEMT contours for SSSS boundary condition.

IV. CONCLUSION

In this study, the dynamic soft-body bird strike behavior of carbon fiber-reinforced polymer (CFRP) composite plates with circular cutouts was investigated numerically using a hybrid SPH-FEA methodology. The effects of bio-inspired helicoidal layup (linear, recursive, and semi-circular) were compared with conventional quasi-isotropic layup laminates under five

ISBN: 978-605-80193-6-2

boundary conditions (CCCC, CCFF, CFCF, CFFF, and SSSS) and two cutout sizes (100 mm and 200 mm).

Boundary constraints have a dominant effect on dynamic compliance; the CFCF support permits greater global flexure, leading to the highest peak strain energy absorption for most layup and cutout combinations (reaching 2070.83J in HR for 200 mm cutouts), whereas simply supported (SSSS) conditions of plates exhibit stiffer behavior with significantly lower peak strain energy.

Increasing the cutout diameter from 100 mm to 200 mm decreases net-section structural stiffness and promotes larger transverse deflections, resulting in higher peak strain energy capacity across the remaining ligament regions.

Limitations of this study are:

The numerical analyses were conducted at a single impact velocity 100 m/s. For future researches strike velocities and strike angle will changed

The current study examined a single overall plate geometry (1 m x 1 m) and two circular cutout diameters (100 mm and 200 mm.

References

[1] Garg, A., Zheng, W., Belarbi, M.-O., & Li, L. (2026). A review on the hybrid SPH-FEM-based bird impact analysis of laminated composite panels. Journal of Reinforced Plastics and Composites, 45(15-16), 3806-3840. 10.1177/07316844251384370

[2] Zhao, S., Yin, X., & Zhang, D. (2023). A study of a bio-inspired impact resistant carbon fiber laminate with a sinusoidal helicoidal structure in the mandibles of trap-jaw ants. Acta Biomaterialia, 169, 179-191. https://doi.org/10.1016/j.actbio.2023.07.047

[3] Zhang, B., Yang, J., Li, Y., Zhang, J., Niu, S., Han, Z., & Ren, L. (2023). Bioinspired basalt fiber composites with higher impact resistance through coupling sinusoidal and helical structures inspired by mantis shrimp. International Journal of Mechanical Sciences, 244, 108073. https://doi.org/10.1016/j.ijmecsci.2022.108073

[4] Chew, E., Liu, J. L., Tay, T. E., Tran, L. Q. N., & Tan, V. B. C. (2021). Improving the mechanical properties of natural fibre reinforced laminates composites through Biomimicry. Composite Structures, 258, 113208. https://doi.org/10.1016/j.compstruct.2020.113208

[5] Berhie, S. E., Chunt, C., Kassa, M. K., & Boddduru, K. (2026). Mechanical Behavior of Bio-Inspired Laminated MXene Reinforced GFRP Composite Core of Structural Sandwich Panels: Numerical and Experimental Study. Polymer Composites, 47(1), 212-230.

[6] Li, Y., Cheng, X., Yan, J., Shi, Z., Shi, H., Ding, R., Sun, J., Yue, D., & Han, Q. (2026). Bio-inspired suture interface for enhancing strength and damage tolerance of repaired composite laminates under tensile and bending tests. Composites Part B: Engineering, 310, 113163. https://doi.org/10.1016/j.compositesb.2025.113163

[7] Garg, A., Shukla, N. K., Avcar, M., Garg, A., & Li, L. (2025). Bird Strike-Induced Damage in Bio-Inspired Sinusoidal Corrugated Helicoidal Laminates. AIAA Journal, 64(4), 2299-2315. 10.2514/1.J066182

[8] Kop, M. D., & Garg, A. (2026). A comparative study on bird strike of helicoidal and quasi-isotropic CFRP laminate plates: Effect of impact angle and circular cutouts. Journal of Ceramics and Composites, 2(2), 1-7. https://dergipark.org.tr/en/pub/jcc/article/1968732

[9] Chen, L., Xu, Y., Li, G., Cheng, J., Zhang, W., & Wang, J. (2026). Experiment and numerical modeling of high-velocity impact behavior of Z-pinned composite laminates. Mechanics of Advanced Materials and Structures, 33(1), 2491840. 10.1080/15376494.2025.2491840

[10] Garg, A., Shukla, N. K., Belarbi, M.-O., Mukherjee, D., Pushpavalli, M., Raman, R., Avcar, M., Singh Narula, M., & Li, L. (2025). Bird strike-induced damage studies on bio-inspired laminated plates with holes. Aerospace Science and Technology, 162, 110200. https://doi.org/10.1016/j.ast.2025.110200

[11] Rahimian Koloor, S. S., Karimzadeh, A., Yidris, N., Petrů, M., Ayatollahi, M. R., & Tamin, M. N. (2020). An Energy-Based Concept for Yielding of Multidirectional FRP Composite Structures Using a Mesoscale Lamina Damage Model. Polymers, 12(1), 157. https://www.mdpi.com/2073-4360/12/1/157

[12] Zhou, Y., Sun, Y., & Huang, T. (2020). Bird-Strike Resistance of Composite Laminates with Different Materials. Materials, 13(1), 129.

ISBN: 978-605-80193-6-2

Free Vibration Analysis of Functionally Graded Porous Nanobeams

Burak İKİNCİ[1] Mehmet AVCAR[1*]

[1]Süleyman Demirel University, Faculty of Engineering and Natural Sciences, Department of Civil Engineering, Isparta, Türkiye

* Corresponding author e-mail: mehmetavcar@sdu.edu.tr

***Abstract*— The free vibration behavior of the functionally graded porous (FGP) nanobeams is investigated incorporating Eringen's nonlocal elasticity theory and a higher-order shear deformation theory, which satisfies the traction-free conditions at the top and bottom surfaces without requiring a shear correction factor. The material properties of FGP nanobeam are modeled using one uniform and two symmetric non-uniform porosity distributions. The equations of motion of FGP nanobeams are derived using Hamilton's principle, and the natural frequencies of simply supported FGP nanobeams are obtained using Navier's method. The present formulation is validated against results available in the literature. The influence of the porosity distribution, nonlocal parameter, slenderness ratio and mode numbers on the natural frequencies of the FGP nanobeams are examined through parametric studies.**

***Keywords*—Free vibration, FGP, nanobeams, porosity, nonlocal elasticity.**

I. INTRODUCTION

FGP structures are modeled and manufactured by using open or closed cell metal foams, which are graded through a continuous variation of pore density or morphology along a specific direction. These structures are lightweight elements that have applications in aerospace, automotive industries as well as in civil engineering structures. Combining exceptionally high strength-to-low weight ratio properties, these materials possess unique capabilities such as energy absorption, vibration damping, thermal insulation and more [1–3].

There are several studies are conducted on mechanical behaviors of FGP beams. Chen et al. [4] studied the bending and buckling behaviors of FGP beams, employing the Ritz method. Jamshidi and Arghavani [5] studied the design problem to optimize the free vibration and buckling behaviors of FGP beams. Civalek et al. [6] studied the free vibration of FGP microbeams resting on a foundation for different boundary conditions, using nonlocal strain gradient theory. Wu et al. [7] investigated buckling and free vibration characteristics of FGP beams using different beam theories.

Kaptan and Ozkol [8] conducted free vibration analysis of FGP Euler and Timoshenko beams using Differential Transform Method (DTM). Nano-sized FGP structures are heavily used in micro/nano-electromechanical systems thanks to their versatility. Consequently, mechanical behaviors of FGP structures remain a subject of interest. At the nanoscale, where the classical continuum theories fail to consider small-scale effects, nonlocal elasticity theories are utilized by researchers.

Although FGP nanobeams have attracted attention over the past decade, studies remain limited. Pia and Delogu [9] conducted research on the bending response of nanoporous metal foam beams, utilizing hyperbolic shear deformation theory. Chen et al. [10] examined the nonlinear free vibration of sandwich nanobeams with an FGP core within the framework of Timoshenko beam theory (TBT). Barati and Zenkour [11] investigated the post-buckling behavior of porous metal foam nanobeams within the scope of a nonlocal nonlinear shear deformable beam model, considering the effects of foundation and porosity distributions. Sahmani et al. [12] analyzed the nonlinear bending of FGP micro- and nanobeams reinforced with graphene platelets (GPL) employing nonlocal strain gradient theory and third-order shear deformation beam theory. Al-Maliki et al. [13] investigated the vibrational characteristics of metal foam nanobeams resting on an elastic foundation. Ozalp and Esen [14] presented a Navier solution for the thermomechanical vibration of sandwich nanobeams with an FGP core under the influence of a magnetic field. Nassir et al. [15] examined the nonlinear vibration of metal foam nanobeams resting on a nonlinear foundation.

To the best of the authors' knowledge, there is a gap in the works focused on the free vibration of FGP nanobeams that consider shear deformation effects and various porosities. In the present study, the free vibration behavior of FGP nanobeam is investigated within the framework of a higher-order shear deformation theory and Eringen's nonlocal elasticity theory using Navier's method. The validity of the present formulation is verified by comparing the results with those of available in the literature. Parametric studies conducted to examine the impact of porosity, small-scale effect, slenderness ratio and mode number on the natural frequencies of the FGP nanobeam.

II. FORMULATION OF THE PROBLEM

A. The Material Properties of FGP Nanobeam

The origin of the coordinate system is located at the mid-plane of the left end of the FGP nanobeam; where x,y,z-axes coincide along the length (L), width (b) and thickness (h), respectively. The geometry and origin placement is given in Fig. 1 (a). The cross-section of FGP nanobeam that displays porosity distribution, mass density and elasticity modulus is shown in Fig. 1 (b).

ISBN: 978-605-80193-6-2

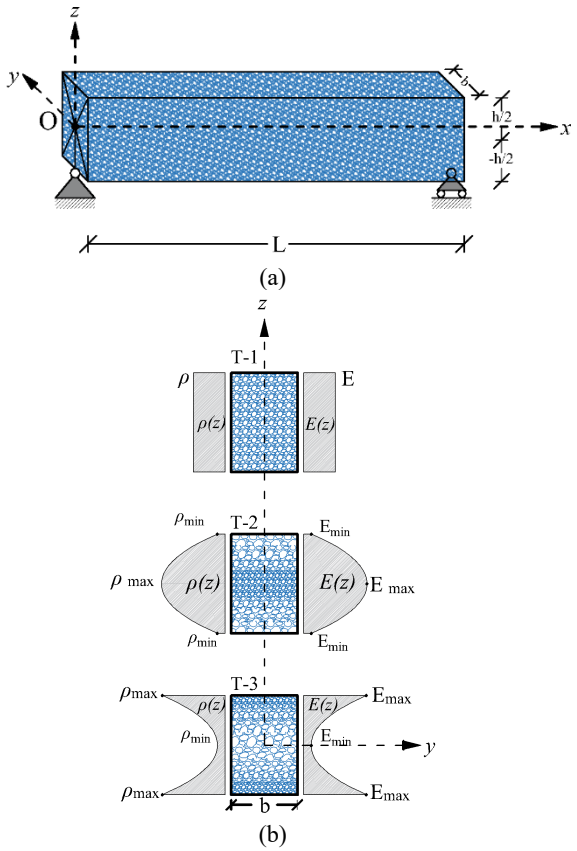


Fig. 1. a) The geometry of the FGP nanobeam; b) The variation of the porosity distribution, mass density and Young's modulus through the thickness direction

The effective material properties of FGP nanobeams vary continuously along the thickness direction or remain uniform depending on the porosity distribution [7]:

• T-1 (Uniform porosity):

$$E(z)=E_s\left[1-\alpha\chi\right] \tag{1a}$$

$$\rho(z)=\rho_s\sqrt{1-\alpha\chi} \tag{1b}$$

• T-2 (Highest porosity at the top and bottom surfaces):

$$E(z)=E_s\left[1-\alpha\cos\left(|\pi z/h|-\pi/2\right)\right] \tag{2a}$$

$$\rho(z)=\rho_s\sqrt{1-\alpha\cos\left(|\pi z/h|-\pi/2\right)} \tag{2b}$$

• T-3 (Highest porosity at the mid-plane):

$$E(z)=E_s\left[1-\alpha\cos\left(\pi z/h\right)\right] \tag{3a}$$

$$\rho(z)=\rho_s\sqrt{1-\alpha\cos\left(\pi z/h\right)} \tag{3b}$$

Here, E_s and ρ_s are elasticity modulus and mass density of solid (non-porous) material. E, ρ, α and χ are Young's modulus, mass density, porosity parameter and uniform porosity coefficient, respectively. Porosity parameter is given as:

$$\alpha=1-\frac{E_p}{E_s} \tag{4}$$

E_p is the minimum value of the Young's modulus (highest porosity). The relation for open-cell foams is expressed [16] as:

$$\frac{E_p}{E_s}=\left(\frac{\rho_p}{\rho_s}\right)^2 \tag{5}$$

Uniform porosity coefficient is calculated by equal-mass condition:

$$\chi=\frac{1}{\alpha}\left[1-\left(\frac{1}{h}\int_{-h/2}^{h/2}\sqrt{1-\alpha\cos\left(\pi z/h\right)}dz\right)^2\right] \tag{6}$$

B. Kinematics and Constitutive Relations

The displacement field is expressed according to the higher-order shear deformable beam theory as:

$$u(x,z,t)=u_0(x,t)-z\frac{\partial w_0(x,t)}{\partial x}+f(z)\varphi(x,t) \tag{7a}$$

$$w(x,z,t)=w_0(x,t) \tag{7b}$$

Here, u and w are axial and transverse displacement of an arbitrary point on the FGP nanobeam. u_0 and w_0 are axial and transverse displacements of a point on the mid-plane. φ is the shear slope and $f(z)$ is the shape function that defines transverse shear strains and stresses [17] :

$$f(z)=(7z/8)-(2z^3/h^2)+(2z^5/h^4) \tag{8}$$

The strain components can be expressed as:

$$\varepsilon_x=\frac{\partial u_0}{\partial x}-z\frac{\partial^2 w_0}{\partial x^2}+f(z)\frac{\partial\varphi}{\partial x} \tag{9a}$$

$$\gamma_{xz}=g(z)\varphi \tag{9b}$$

$$g(z)=\frac{df(z)}{dz} \tag{10}$$

where ε_x and γ_{xz} are axial strain and transverse shear strain, respectively. $g(z)$ is the first derivative of the shape function. The variation of $g(z)$ through the thickness is shown in Fig. 2.

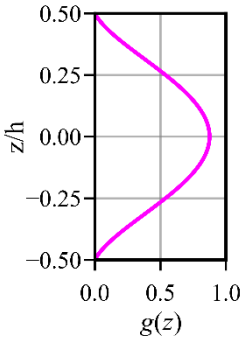


Fig. 2. The variation of derivative of the shape function through thickness

The normal and transverse stresses at any point in nanobeam can be expressed as:

$$\sigma_x=E(z)\varepsilon_x \tag{11a}$$

$$\tau_{xz}=G(z)\gamma_{xz} \tag{11b}$$

$$G(z)=\frac{E(z)}{2(1+\nu)} \tag{11c}$$

$G(z)$ and ν represent the shear modulus and Poisson's ratio of the FGP nanobeam, respectively.

ISBN: 978-605-80193-6-2

C. The Equations of Motion

Hamilton's principle is used to derive the equations of motion [18]:

$$\int_{t_0}^{t_1}\left(\delta\Pi-\delta K\right)dt=0 \tag{12}$$

t, $\delta\Pi$ and δK are the time, the virtual variations of the strain energy and kinetic energy. The virtual variation of the strain energy of the beam is:

$$\begin{aligned}\delta\Pi &= b\int_0^L\int_{-h/2}^{h/2}\left(\sigma_x\delta\varepsilon_x+\tau_{xz}\,\delta\gamma_{xz}\right)dzdx= \\ &=\int_0^L\left(N\frac{\partial\delta u_0}{\partial x}-M\frac{\partial^2\delta w_0}{\partial x^2}+P\frac{\partial\delta\varphi}{\partial x}+Q\delta\varphi\right)dx\end{aligned} \tag{13}$$

N, M, P and Q are the stress and moment resultants and can be defined as follows:

$$\left\{N,M,P\right\}=b\int_{-h/2}^{h/2}\left\{1,z,f\left(z\right)\right\}\sigma_x dz \tag{14a}$$

$$Q=b\int_{-h/2}^{h/2}g\left(z\right)\tau_{xz}dz \tag{14b}$$

$$\begin{Bmatrix}N\\M\\P\\Q\end{Bmatrix}=\begin{bmatrix}A&B&C&0\\B&D&F&0\\C&F&H&0\\0&0&0&J\end{bmatrix}\begin{Bmatrix}\dfrac{\partial u_0}{\partial x}\\-\dfrac{\partial^2 w_0}{\partial x^2}\\\dfrac{\partial\varphi}{\partial x}\\\varphi\end{Bmatrix} \tag{14c}$$

here, A, B, C, D, F, H and J are the stiffness components and can be defined as follows:

$$\begin{Bmatrix}A\\B\\C\\D\\F\\H\end{Bmatrix}=b\int_{-h/2}^{h/2}E\left(z\right)\begin{bmatrix}1\\z\\f\left(z\right)\\z^2\\zf\left(z\right)\\f^2\left(z\right)\end{bmatrix}dz \tag{15a}$$

$$J=b\int_{-h/2}^{h/2}G\left(z\right)g^2\left(z\right)dz \tag{15b}$$

The virtual variation of the kinetic energy is:

$$\delta K=b\int_0^L\int_{-h/2}^{h/2}\rho\left(z\right)\left[\frac{\partial u}{\partial t}\frac{\partial\delta u}{\partial t}+\frac{\partial w}{\partial t}\frac{\partial\delta w}{\partial t}\right]dzdx \tag{16}$$

By substituting (16) and (13) into (12) integrating by parts with respect to time and x, and setting the coefficients of δu, δw and $\delta\varphi$ to zero, the equations of motion are obtained:

$$\delta u:\frac{\partial N}{\partial x}=I_0\frac{\partial^2 u_0}{\partial t^2}-I_1\frac{\partial^3 w_0}{\partial x\partial t^2}+I_2\frac{\partial^2\varphi}{\partial t^2} \tag{17a}$$

$$\begin{aligned}\delta w:&\frac{\partial^2 M}{\partial x^2}=I_0\frac{\partial^2 w_0}{\partial t^2}+I_1\frac{\partial^3 u_0}{\partial x\partial t^2}- \\ &I_3\frac{\partial^4 w_0}{\partial x^2\partial t^2}+I_4\frac{\partial^3\varphi}{\partial x\partial t^2}\end{aligned} \tag{17b}$$

$$\delta\varphi:\frac{\partial P}{\partial x}-Q=I_2\frac{\partial^2 u_0}{\partial t^2}-I_4\frac{\partial^3 w_0}{\partial x\partial t^2}+I_5\frac{\partial^2\varphi}{\partial t^2} \tag{17c}$$

where $I_0, I_1, I_2, I_3, I_4, I_5$ are inertia coefficients and can be expressed as:

$$\begin{Bmatrix}I_0\\I_1\\I_2\\I_3\\I_4\\I_5\end{Bmatrix}=b\int_{-h/2}^{h/2}\rho\left(z\right)\begin{bmatrix}1\\z\\f\left(z\right)\\z^2\\zf\left(z\right)\\f^2\left(z\right)\end{bmatrix}dz \tag{18}$$

D. Eringen's Nonlocal Elasticity Theory

In this study, Eringen's nonlocal elasticity theory is utilized to introduce the small size-effect (softening) of the FGP nanobeam. The differential form of Eringen's nonlocal elasticity for the FGP nanobeam can be defined as [19] :

$$\left(1-\lambda\frac{\partial^2}{\partial x^2}\right)\sigma_x=E(z)\varepsilon_x \tag{19a}$$

$$\left(1-\lambda\frac{\partial^2}{\partial x^2}\right)\tau_{xz}=G(z)\gamma_{xz} \tag{19b}$$

$$\lambda=\left(e_0 a\right)^2\left(nm^2\right) \tag{19c}$$

Here, λ is the nonlocal parameter that accounts for the small-scale effect. Dimensionless material constant e_0 is calibrated by experiments or atomistic simulations and internal characteristic length a is dependent on the material (e.g. lattice size and granular distance) [20]. Note that the local formulation is recovered by setting $\lambda=0$.

By substituting the classical stress resultants (14) into the nonlocal constitutive relations (19), and subsequently inserting the resulting expressions into the classical equations of motion (17), the final form of the nonlocal governing equations can be obtained:

$$\begin{aligned}&A\frac{\partial^2 u_0}{\partial x^2}-B\frac{\partial^3 w_0}{\partial x^3}+C\frac{\partial^2\varphi}{\partial x^2} \\ &=\left(1-\lambda\frac{\partial^2}{\partial x^2}\right)\left(I_0\frac{\partial^2 u_0}{\partial t^2}-I_1\frac{\partial^3 w_0}{\partial x\partial t^2}+I_2\frac{\partial^2\varphi}{\partial t^2}\right)\end{aligned} \tag{20a}$$

$$\begin{aligned}&B\frac{\partial^3 u_0}{\partial x^3}-D\frac{\partial^4 w_0}{\partial x^4}+F\frac{\partial^3\varphi}{\partial x^3}= \\ &\left(1-\lambda\frac{\partial^2}{\partial x^2}\right)\left(I_0\frac{\partial^2 w_0}{\partial t^2}+I_1\frac{\partial^3 u_0}{\partial x\partial t^2}-I_3\frac{\partial^4 w_0}{\partial x^2\partial t^2}+I_4\frac{\partial^3\varphi}{\partial x\partial t^2}\right)\end{aligned} \tag{20b}$$

$$\begin{aligned}&C\frac{\partial^2 u_0}{\partial x^2}-F\frac{\partial^3 w_0}{\partial x^3}+H\frac{\partial^2\varphi}{\partial x^2}-J\varphi \\ &=\left(1-\lambda\frac{\partial^2}{\partial x^2}\right)\left(I_2\frac{\partial^2 u_0}{\partial t^2}-I_4\frac{\partial^3 w_0}{\partial x\partial t^2}+I_5\frac{\partial^2\varphi}{\partial t^2}\right)\end{aligned} \tag{20c}$$

E. Analytical Solution

Navier's method is used to obtain analytical solutions. The boundary conditions for the simply supported FGP nanobeam at $x=0,\ x=L$:

$$w=N=M=P=0 \tag{21}$$

ISBN: 978-605-80193-6-2

To obtain the analytical solution, the displacement variables are expanded in terms of trigonometric series that satisfy the boundary conditions:

$$\begin{Bmatrix} u \\ w \\ \varphi \end{Bmatrix} = \sum_{k=1}^{\infty} \begin{Bmatrix} U_k \cos(\eta x) \\ W_k \sin(\eta x) \\ \Phi_k \cos(\eta x) \end{Bmatrix} e^{i\omega t} \tag{22a}$$

$$\eta = k\pi / L \tag{22b}$$

where $i = \sqrt{-1}$, U_k, W_k, Φ_k are the unknown coefficients to be determined. k is the mode number and ω is the natural frequency.

By substituting (22) into (20), the eigenvalue problem for free vibration of FGP nanobeams can be obtained:

$$\left([S] - \omega^2 \left(1 + \lambda\eta^2\right)[M] \right) \begin{Bmatrix} U_k \\ W_k \\ \Phi_k \end{Bmatrix} = 0 \tag{23a}$$

$$[S] = \begin{bmatrix} A\eta^2 & -B\eta^3 & C\eta^2 \\ & D\eta^4 & -F\eta^3 \\ Sym. & & H\eta^2 + J \end{bmatrix} \tag{23b}$$

$$[M] = \begin{bmatrix} I_0 & -I_1\eta & I_2 \\ & I_0 + I_3\eta^2 & -I_4\eta \\ Sym. & & I_5 \end{bmatrix} \tag{23c}$$

III. RESULTS AND DISCUSSION

In this section firstly, to validate the accuracy, findings of the proposed theory are compared with those in the existing literature. Secondly, dimensionless natural frequencies of the FGP nanobeams versus different nonlocal parameter, slenderness ratio and the porosity parameter and distributions is investigated. The following parameters are taken: $E_s = 200$ GPa, $\rho_s = 7850\,\mathrm{kg/m^3}$, $\nu = 1/3$ $L = 10\,\mathrm{nm}$,. b=1 nm. For dimensionless natural frequency (DNF), the following form is assumed, unless otherwise stated:

$$\bar{\omega} = \omega L^2 \sqrt{\rho_s A / E_s I} \tag{24}$$

In Table I, the dimensionless fundamental frequencies (DFF) of isotropic nanobeams versus λ for different L/h are compared with those of Hadji and Avcar [21].

TABLE I. Comparison of DFFs of isotropic nanobeams respect to λ

$\lambda\left[nm^2\right]$	Ref.	L/h		
		5	10	20
0	Present	9.2716	9.7066	9.8279
1		8.8454	9.2603	9.3761
2		8.4730	8.8705	8.9813
3		8.1440	8.5261	8.6326
4		7.8506	8.2189	8.3216
0	[21]	9.2749	9.7076	9.8282
1		8.8485	9.2613	9.3763
2		8.4759	8.8714	8.9816
3		8.1469	8.5269	8.6329
4		7.8534	8.2197	8.3218

In Table II, the DFF of FGP beams ($\lambda = 0$) versus α for slenderness ratio $L/h = 10$ are compared with those of Wu et al. [7]. Here, DFF is taken as:

$$\bar{\omega} = \omega 10h \sqrt{\rho_s / E_s} \tag{25}$$

TABLE II. Comparison of DFFs of FGP beams respect to α

Ref.	Present Study			[7]		
α	T1	T2	T3	T1	T2	T3
0.00	0.2802	0.2802	0.2802	0.2802	0.2802	0.2802
0.90	0.2201	0.1629	0.2861	0.2201	0.1629	0.2874
0.99	0.2062	0.1362	0.2928	0.2062	0.1362	0.2961

It can be seen from Table I and Table II, present formulation is valid and accurate.

In Fig. 3, the variation of first three modes of DNFs of FGP nanobeams versus α are presented for first three modes, $L/h = 5$ and $\lambda = 2\left(\mathrm{nm}^2\right)$.

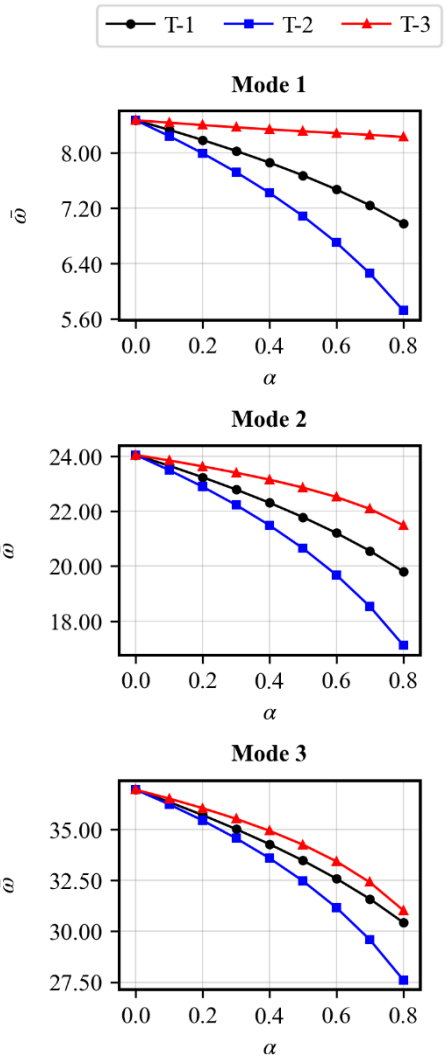


Fig. 3. The variation of first three modes of DNFs of FGP nanobeams respect to α

According to Fig. 3, DNFs decrease with the increment of α, and the impact of α on the DNFs of FGP nanobeams depends on the porosity distribution. Particularly for T-3 FGP nanobeam, the effect of α on DNFs is more pronounced at higher mode numbers.

The variation of DNFs of FGP nanobeams versus λ is visualized for first three modes, $L/h = 10$ and $\alpha = 0.6$ in Fig. 4.

ISBN: 978-605-80193-6-2

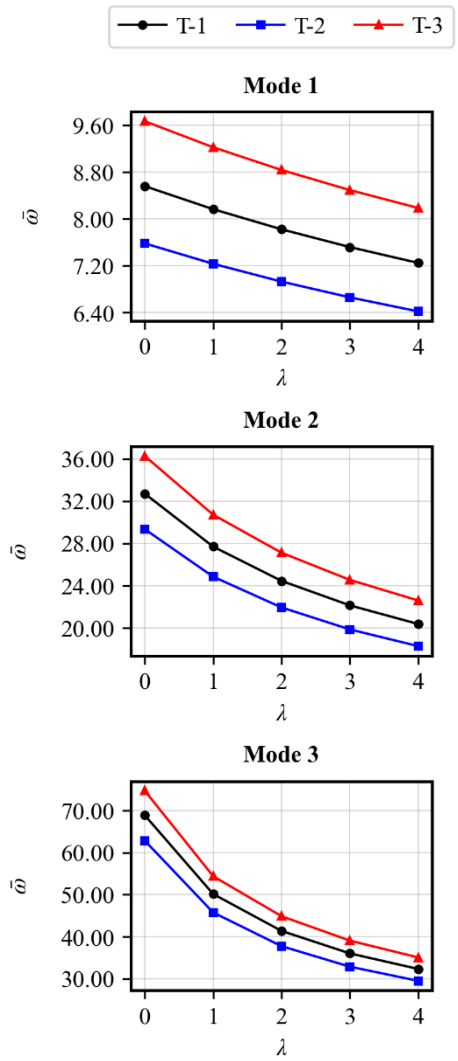


Fig. 4. The variation of first three modes of DNFs of FGP nanobeams respect to λ

As shown in Fig. 4, an increase in λ decreases DNFs due to stiffness-softening effect. The DNF reduction is highly dependent on specific vibration mode.

In Fig. 5, the variation of first three modes of DNFs of FGP nanobeams versus L/h for $\alpha = 0.5$ and $\lambda = 3\left(\mathrm{nm}^2\right)$ is given.

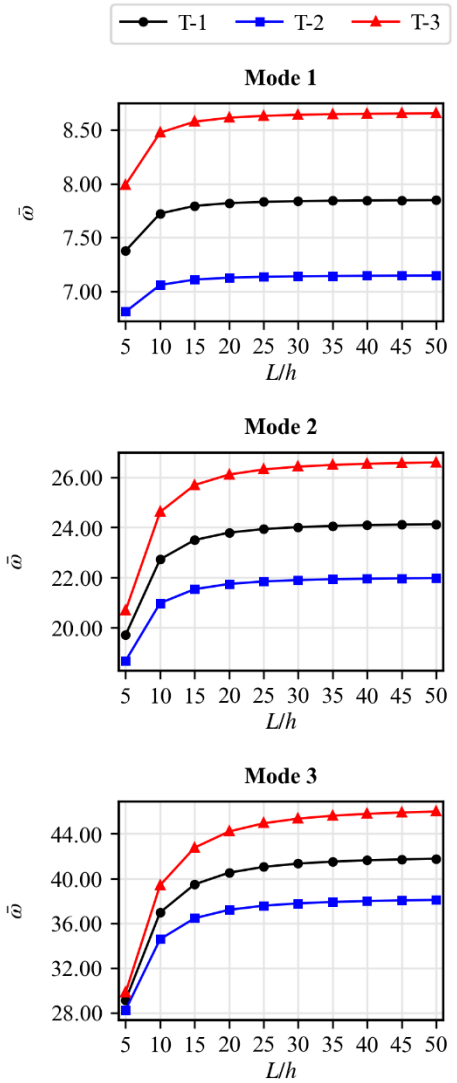


Fig. 5. The variation of first three modes of DNFs of FGP nanobeams respect to L/h

Fig. 5 shows that DNFs increase as L/h values increase, for all porosity distributions and modes. The increment of DNFs is more pronounced for the values between $L/h = 5$ and $L/h = 10$.

Fig. 6 examines the variation of DNFs of FGP nanobeams versus mode numbers for $\alpha = 0.8$ and $\lambda = 4\left(\mathrm{nm}^2\right)$.

ISBN: 978-605-80193-6-2

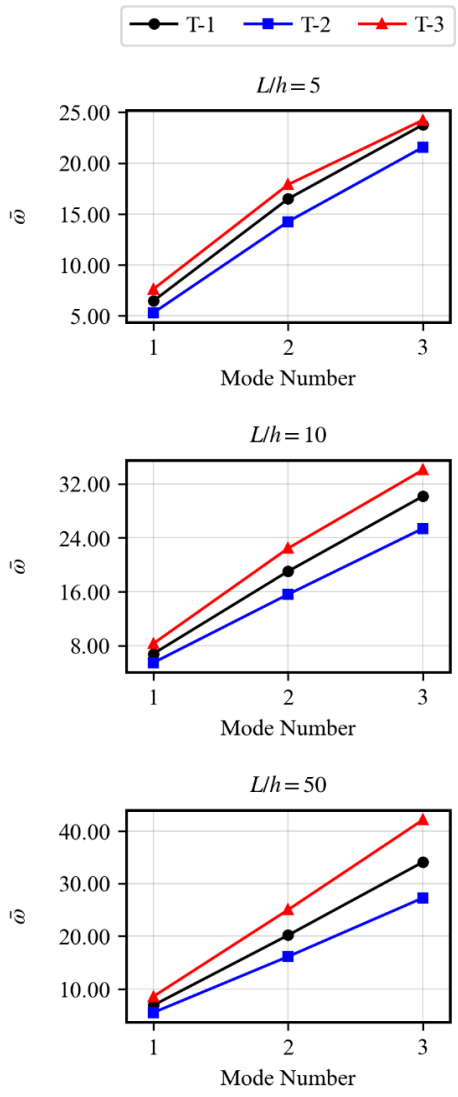


Fig. 6. The variation of DNFs of FGP nanobeams respect to mode numbers

Fig. 6 shows that DNFs increase with the mode number for all porosity distributions and L/h. All the results indicate that the T-3 distribution yields the highest frequencies, whereas T-2 yields the lowest. This is because T-3 preserves structural stiffness by minimizing porosity at the top and bottom FGP beam surfaces, while T-2 inherently weakens it by maximizing porosity at these surfaces.

IV. CONCLUSIONS

In this paper, the free vibration behavior of FGP nanobeams was investigated utilizing a higher-order shear deformation theory. The present model satisfies the zero transverse shear stress conditions at the top and bottom surfaces, without requiring a shear correction factor. Three distinct porosity distributions for FGP nanobeams were assumed and the governing equations of FGP nanobeams were derived via Hamilton's principle. The small-scale effect was taken into account by employing Eringen's nonlocal elasticity theory. Navier-type solutions for the free vibration analysis of simply supported FGP nanobeams were obtained. The results of present formulation were compared and found to be in good agreement with existing literature and comprehensive parametric studies were conducted to examine the impact of nonlocal parameter, porosity and slenderness ratio on the natural frequencies of the FGP nanobeams. Based on the numerical results of current problem, the following conclusions can be made:

- The small-scale effect plays a remarkable role in the dimensionless natural frequencies of FGP nanobeams. An increase in the nonlocal parameter introduces a stiffness-softening effect, which decreases the natural frequencies.
- The porosity decreases the dimensionless natural frequencies of FGP nanobeams, an effect that is highly dependent on the distribution type: the T-3 distribution exhibited the highest frequencies, while the T-2 profile yields the lowest.
- The impact of porosity on the dimensionless natural frequencies remarkably increases for higher vibration modes of the T-3 FGP nanobeams.
- The increase in the slenderness ratio leads to higher dimensionless natural frequencies.

This paper specifically addresses the free vibration behavior of simply supported FGP nanobeams. The scope of the current study can be broadened in future investigations to explore the bending, buckling, and dynamic responses of various FGP nanostructures under distinct boundary conditions and different size-dependent continuum theories.

References

[1] Pollien, A., Conde, Y., Pambaguian, L., & Mortensen, A. (2005). *Graded open-cell aluminium foam core sandwich beams*. *Materials Science and Engineering: A*, *404*(1–2), 9–18. https://doi.org/10.1016/j.msea.2005.05.096

[2] Banhart, J. (2001). *Manufacture, characterisation and application of cellular metals and metal foams*. *Progress in Materials Science*, *46*(6), 559–632. https://doi.org/10.1016/S0079-6425(00)00002-5

[3] García-Moreno, F. (2016). *Commercial Applications of Metal Foams: Their Properties and Production*. *Materials*, *9*(2), 85. https://doi.org/10.3390/ma9020085

[4] Chen, D., Yang, J., & Kitipornchai, S. (2015). *Elastic buckling and static bending of shear deformable functionally graded porous beam*. *Composite Structures*, *133*, 54–61. https://doi.org/10.1016/j.compstruct.2015.07.052

[5] Jamshidi, M., & Arghavani, J. (2018). *Optimal material tailoring of functionally graded porous beams for buckling and free vibration behaviors*. *Mechanics Research Communications*, *88*, 19–24. https://doi.org/10.1016/j.mechrescom.2018.01.006

[6] Civalek, Ö., Ersoy, H., Uzun, B., & Yaylı, M. Ö. (2023). *Dynamics of a FG porous microbeam with metal foam under deformable boundaries*. *Acta Mechanica*, *234*(11), 5385–5404. https://doi.org/10.1007/s00707-023-03663-7

[7] Wu, S., Li, Y., Bao, Y., Zhu, J., & Wu, H. (2024). *Examination of Beam Theories for Buckling and Free Vibration of Functionally Graded Porous Beams*. *Materials*, *17*(13), 3080. https://doi.org/10.3390/ma17133080

ISBN: 978-605-80193-6-2

[8] Kaptan, S., & Ozkol, I. (2026). *Free vibration analysis of functionally graded porous beams using the differential transform method under Euler–Bernoulli and Timoshenko theories*. *Discover Mechanical Engineering*, *5*(1), 16. https://doi.org/10.1007/s44245-026-00187-x

[9] Pia, G., & Delogu, F. (2013). *On the elastic deformation behavior of nanoporous metal foams*. *Scripta Materialia*, *69*(11–12), 781–784. https://doi.org/10.1016/j.scriptamat.2013.08.027

[10] Chen, D., Kitipornchai, S., & Yang, J. (2016). *Nonlinear free vibration of shear deformable sandwich beam with a functionally graded porous core*. *Thin-Walled Structures*, *107*, 39–48. https://doi.org/10.1016/j.tws.2016.05.025

[11] Barati, M. R., & Zenkour, A. M. (2017). *Investigating post-buckling of geometrically imperfect metal foam nanobeams with symmetric and asymmetric porosity distributions*. *Composite Structures*, *182*, 91–98. https://doi.org/10.1016/j.compstruct.2017.09.008

[12] Sahmani, S., Aghdam, M. M., & Rabczuk, T. (2018). *Nonlinear bending of functionally graded porous micro/nano-beams reinforced with graphene platelets based upon nonlocal strain gradient theory*. *Composite Structures*, *186*, 68–78. https://doi.org/10.1016/j.compstruct.2017.11.082

[13] Al-Maliki, A. F. H., Faleh, N. M., & Alasadi, A. A. (2019). *Finite element formulation and vibration of nonlocal refined metal foam beams with symmetric and non-symmetric porosities*. *Structural Monitoring and Maintenance*, *6*(2).

[14] Ozalp, A. F., & Esen, I. (2025). *Magnetic field effects on the thermomechanical vibration behavior of functionally graded biocompatible material sandwich nanobeams*. *Mechanics of Advanced Materials and Structures*, *32*(3), 459–477. https://doi.org/10.1080/15376494.2024.2349966

[15] Layla M. Nassir, Nadhim M. Faleh, Raad M. Fenjan, Mustafa Qahtan Hadi Al-Ghrari, & Mamoon A. A. Al-Jaafari. (2025). *Nonlinear Vibration Analysis of Metal Foam Nanobeams with Varying Porosity*. *International Journal of Mechanical Engineering and Robotics Research*, *14*(4), 407–417. https://doi.org/10.18178/ijmerr.14.4.407-417

[16] Gibson, I. J., & Ashby, M. F. (1982). *The mechanics of three-dimensional cellular materials*. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, *382*(1782), 43–59. https://doi.org/10.1098/rspa.1982.0088

[17] Nguyen-Xuan, H., Thai, C. H., & Nguyen-Thoi, T. (2013). *Isogeometric finite element analysis of composite sandwich plates using a higher order shear deformation theory*. *Composites Part B: Engineering*, *55*, 558–574. https://doi.org/10.1016/j.compositesb.2013.06.044

[18] Reddy, J. N. (2002). *Energy Principles and Variational Methods in Applied Mechanics* (2nd Edition). John Wiley & Sons.

[19] Eringen, A. C. (1983). *On differential equations of nonlocal elasticity and solutions of screw dislocation and surface waves*. *Journal of Applied Physics*, *54*(9), 4703–4710. https://doi.org/10.1063/1.332803

[20] Wang, Q., & Shindo, Y. (2006). *Nonlocal continuum models for carbon nanotubes subjected to static loading*. *Journal of Mechanics of Materials and Structures*, *1*(4), 663–680. https://doi.org/10.2140/jomms.2006.1.663

[21] Hadji, L., & Avcar, M. (2021). *Nonlocal free vibration analysis of porous FG nanobeams using hyperbolic shear deformation beam theory* . *Advances in Nano Research*, *10*(3), 281–293.

 ISBN: 978-605-80193-6-2

TTHM and HAA5 as Related but Non-Equivalent Disinfection By-Product Groups: Implications for Drinking-Water Monitoring and Data-Driven Modelling

Berkin Üstünyıldız[1]

[1]Düzce University, Kaynaşlı Vocational School, Department of Environmental Protection Technologies, Düzce, Türkiye

* Corresponding author e-mail: berkinuy@gmail.com

***Abstract*—Total trihalomethanes (TTHM) and the five haloacetic acids grouped as HAA5 are major disinfection by-product groups monitored in chlorinated drinking water. Although they arise from related chlorination chemistry and respond to many of the same water-quality and operational factors, their shared drivers do not imply interchangeable behaviour. This focused review examines how precursor characteristics, pH, temperature, bromide-related speciation, formation kinetics, water age, disinfectant exposure, and constituent-specific loss processes can produce non-parallel TTHM and HAA5 patterns. Evidence from controlled studies and full-scale distribution systems shows that the magnitude, timing, and location of elevated concentrations can differ between the two groups, while coincident behaviour can also occur under some conditions. These distinctions have practical implications for drinking-water monitoring because locations or sampling periods representative of higher TTHM concentrations need not provide equivalent information for HAA5. Quantitative surrogacy should therefore be demonstrated rather than assumed. The same principle applies to data-driven modelling. Common predictors and modelling frameworks may be applicable to both outcomes, but predictor relevance, transferability, and predictive performance should be evaluated separately for each target and dataset context. Overall, TTHM and HAA5 are best treated as related but non-equivalent disinfection by-product groups whose observed relationship depends on source-water chemistry, treatment, and distribution-system conditions.**

***Keywords*—Disinfection by-products, TTHM, HAA5, drinking-water monitoring, data-driven modelling**

I. INTRODUCTION

Disinfection remains essential for controlling microbial risks in drinking-water treatment and distribution. Chlorine is widely used for this purpose, but its reaction with natural organic matter and other constituents in water also produces disinfection by-products (DBPs). Among these compounds, trihalomethanes (THMs) and haloacetic acids (HAAs) are two of the most frequently monitored groups in chlorinated drinking water. Their formation is influenced by organic matter characteristics, water chemistry, and disinfection conditions [1]. Subsequent distribution-system behaviour is considered below.

Regulatory practice also treats the two groups separately. Under the U.S. Stage 2 Disinfectants and Disinfection Byproducts Rule (Stage 2 DBPR), total trihalomethanes (TTHM) and five haloacetic acids (HAA5) have maximum contaminant levels of 0.080 and 0.060 mg/L, respectively. Compliance is evaluated using a locational running annual average at each monitoring location rather than a single system-wide average [2]. This spatially resolved approach is relevant because the locations associated with higher THM and HAA concentrations do not necessarily coincide. Speight and Singer, for example, reported that maximum THM and HAA concentrations may occur at different locations and that their relationships with residence time can differ [3].

THM and HAA concentrations can nevertheless be strongly correlated under some conditions. Villanueva et al., who analysed nine HAAs in drinking waters from heterogeneous sources, reported a strong association between total THM and total HAA concentrations [4]. Such associations, however, do not establish that the regulated groups behave as interchangeable indicators. Full-scale studies have documented non-parallel spatial and temporal patterns. In the distribution system examined by Rodríguez et al., THM concentrations generally increased and stabilised toward system extremities, whereas the measured HAA concentrations initially increased and then declined, a pattern driven largely by decreasing dichloroacetic acid [5]. In a more recent rural system, the relationship between modelled water age and TTHM or HAA5 varied seasonally, while HAA5 and TTHM concentrations reached their seasonal peaks at different times [6].

This focused review examines why TTHM and HAA5 can exhibit non-identical responses to physicochemical and distribution-system conditions and considers the practical implications of this distinction. The discussion focuses on three questions: why the two regulated DBP groups may behave differently, what this means for drinking-water monitoring, and how such non-equivalence should be considered in data-driven modelling. The central argument is not that TTHM and HAA5 are unrelated, but that their relationship should be demonstrated rather than assumed to imply interchangeability.

II. WHY TTHM AND HAA5 DO NOT BEHAVE IDENTICALLY

TTHM and HAA5 share many of the same broad precursor sources, but this does not mean that they respond identically to changes in source-water composition. Natural organic matter (NOM) is chemically heterogeneous, and different fractions can favour different DBP pathways. In fractionation experiments, trihaloacetic acid precursors were generally more hydrophobic than THM precursors, while hydrophilic and low-molecular-weight fractions

ISBN: 978-605-80193-6-2

produced particularly high dihaloacetic acid yields [7]. Results from Istanbul reservoir waters similarly showed that pH, contact time, chlorine dose and specific ultraviolet absorbance (SUVA) were significant factors for both THM and HAA formation, while the relative formation potentials differed among waters with different organic-matter characteristics [1]. Thus, bulk measures of organic matter can be informative without implying an equivalent response of the two regulated DBP groups.

pH provides a clearer example of non-identical formation behaviour. During free chlorination, Hua and Reckhow observed a strong increase in THM formation as pH increased from 5 to 10. In the same experiments, however, trihaloacetic acids were suppressed at high pH, whereas dihaloacetic acids increased under alkaline conditions [8]. These findings should not be translated into a simple rule that higher pH increases TTHM while decreasing HAA5. Rather, they show that even constituent HAA classes can respond differently to the same pH change. This internal heterogeneity is one reason why an aggregate HAA5 response cannot be inferred directly from the behaviour of a single HAA subclass.

Temperature can also affect DBP groups to different extents. Under the chlorination conditions tested by Hua and Reckhow, increasing temperature from 5 to 30 °C produced a greater than 100% increase in THMs, compared with smaller increases in dihaloacetic and trihaloacetic acids [8]. The important implication is not that TTHM is universally more temperature-sensitive than HAA5, but that a common temperature increase need not produce proportional responses across THMs and individual HAA classes. Field-scale seasonal behaviour is therefore considered separately in the following section rather than inferred directly from laboratory temperature effects.

Bromide adds another source of divergence because it affects DBP composition as well as concentration. In the Istanbul reservoir study, the water containing the highest bromide concentration showed a greater contribution of brominated species, including dibromochloromethane and dibromoacetic acid (DBAA) [1]. In this higher-bromide water, DBAA, which is included in the regulated HAA5 group, accounted for more than 13% of the measured total HAA [1]. This illustrates why changes in bromide should not be reduced to a simple increase or decrease in total TTHM or HAA5. Similar group totals can conceal substantial differences in constituent composition.

Disinfectant exposure and contact time provide the link between formation chemistry and distribution-system behaviour. Different DBP classes can form at different rates during chlorination; Hua and Reckhow found that dihaloacetic acids formed more rapidly than THMs and trihaloacetic acids during the early stages of reaction [8]. Sfynia et al. likewise combined full-scale observations with simulated distribution-system (SDS) tests and found that, in chlorinated SDS experiments, THMs increased markedly between 6 and 106 h, whereas the measured aggregate of nine haloacetic acids (HAA9) showed a much smaller net increase and intermediate increases and decreases [9]. These results indicate that continued disinfectant exposure does not necessarily produce parallel concentration trajectories.

Overall, TTHM and HAA5 arise within a shared disinfection chemistry, but differences in precursor characteristics, pH response, temperature sensitivity, brominated speciation and formation kinetics prevent their behaviour from being assumed equivalent. These distinctions become especially important once treated water enters the distribution system.

III. NON-PARALLEL BEHAVIOUR WITHIN DRINKING-WATER DISTRIBUTION SYSTEMS

Once treated water enters a distribution system, continuing reactions, water age and hydraulic conditions can modify the DBP profile observed at the treatment-plant outlet. However, increasing residence time does not impose a common trajectory on TTHM and HAA5. In the full-scale system studied by Rodríguez et al., THM concentrations generally increased and then stabilised toward system extremities, whereas the measured HAA concentrations initially increased and subsequently declined. The HAA decrease was driven largely by declining dichloroacetic acid, and residence time was an important explanatory variable for both DBP groups [5]. This finding is useful as an example of non-parallel behaviour, but it should not be interpreted as a general rule that increasing water age always raises TTHM or lowers HAA5.

Observed HAA concentrations can reflect both continuing formation and constituent-specific loss. Field monitoring by Tung and Xie across five small drinking-water systems found that HAA reduction at some locations was associated with low chlorine residuals and elevated heterotrophic bacterial populations, supporting a role for biological loss under suitable distribution-system conditions [10]. Loss within the HAA group is also not uniform. In the Rodríguez study, the decline in measured HAAs was mainly attributable to dichloroacetic acid rather than to an equivalent decrease across all HAA species [5]. Consequently, species-specific loss should not be converted into a universal statement about HAA5 behaviour.

Distribution-system storage and disinfectant residual further illustrate how the two groups can separate spatially. Speight and Singer compiled occurrence data from several distribution-system studies and showed that maximum THM and maximum HAA concentrations did not necessarily occur at the same location or time [3]. In a Morehead case included in their analysis, older water associated with storage and low disinfectant residual coincided with lower HAA concentrations, while THM concentrations remained comparatively high [3]. Such observations demonstrate that hydraulically remote or storage-influenced locations can produce different balances of continued formation and loss, without implying that storage universally decreases HAA5.

Temporal patterns are similarly system-dependent. Rodríguez et al. reported average THM concentrations in summer and autumn that were about five times winter levels, whereas measured HAA concentrations were highest in spring and about four times winter levels [5].

ISBN: 978-605-80193-6-2

In a much larger Massachusetts dataset, Parvez et al. found the highest mean concentrations of the four regulated trihalomethanes (THM4) in the third quarter, whereas HAA5 was highest in the second and third quarters; their seasonal minima also occurred in different quarters [11]. Unrine et al. likewise observed HAA5 peaking in mid-summer and TTHM in early autumn, with water-age relationships changing by season [6]. Yet non-equivalence does not require opposite seasonal behaviour: in three Istanbul distribution systems, both THM4 and HAA9 showed their highest median concentrations during spring and summer [12].

The distinction, therefore, is not that TTHM consistently increases while HAA5 decreases. Their observed distributions can instead reflect different balances of formation, persistence and constituent-specific loss, producing spatial and temporal patterns that are not reliably interchangeable.

IV. IMPLICATIONS FOR DRINKING-WATER MONITORING

The spatial variability described above has direct implications for interpreting monitoring data. A location selected to characterise high TTHM concentrations may not characterise high HAA concentrations to the same extent. Speight and Singer showed that maximum THM and maximum HAA concentrations need not occur simultaneously at the same site, reflecting differences in their behaviour within distribution systems [3]. This locational issue also has regulatory relevance. Under the U.S. Stage 2 DBPR, compliance with TTHM and HAA5 maximum contaminant levels is assessed using a locational running annual average for each monitoring location rather than a system-wide running annual average [2]. The regulation does not itself establish chemical non-equivalence, but it illustrates the practical importance of interpreting DBP concentrations in a location-specific context.

Sampling time presents a related issue. In the Massachusetts dataset analysed by Parvez et al., mean THM4 concentrations were highest in the third quarter and lowest in the first, whereas HAA5 concentrations were highest in the second and third quarters and lowest in the fourth. Some systems also showed substantial differences between successive monthly means [11]. Unrine et al. similarly observed HAA5 concentrations peaking in mid-summer and TTHM concentrations in early autumn in a rural distribution system [6]. These findings do not imply that quarterly monitoring is inadequate or that a higher sampling frequency is universally required. They instead show that the temporal representativeness of a sampling period can differ between the two DBP groups.

The same caution applies to quantitative surrogacy. Across 9,003 matched samples, Parvez et al. reported a THM4–HAA5 correlation of 0.52, with correlations varying according to source-water type [11]. This demonstrates that the strength of the relationship can vary among systems rather than establishing that one group is an unreliable surrogate for the other. Accordingly, use of TTHM or HAA5 as a quantitative surrogate for the other should be demonstrated rather than assumed. Monitoring design can likewise depend on the specific objective and observed spatiotemporal pattern: Delpla et al. showed in an intensively monitored residential sector that alternative monitoring objectives led to different solutions for sampling locations and periods [13]. Such optimisation remains system- and dataset-specific.

Overall, non-equivalent spatial and temporal behaviour does not call for a universal monitoring prescription. It instead supports parameter-specific interpretation of monitoring locations, sampling periods and any proposed quantitative relationship between TTHM and HAA5.

V. IMPLICATIONS FOR DATA-DRIVEN MODELLING

The non-equivalent behaviour of TTHM and HAA5 does not imply that completely different modelling architectures are required. Under uniform chlorination conditions, Ersan et al. applied several nonlinear and linear modelling approaches to both THM4 and HAA5 using a common set of water-quality and operational inputs [14]. The same least-squares support vector machine (LS-SVM) approach performed best for both regulated outcomes in the combined water-source analysis, but test performance differed, with R^2 values of 0.80 for THM4 and 0.72 for HAA5 [14]. This illustrates an important distinction: a common modelling framework can be suitable for both targets without implying equivalent predictive performance. Good performance for TTHM therefore does not, by itself, establish comparable performance for HAA5.

Predictor relevance can also depend on the DBP outcome and the type of dataset. In a decade-long full-scale study of Seoul's drinking-water distribution system, Park et al. modelled several DBP groups using common water-quality, temporal and distribution-system variables. Water temperature dominated feature importance for THM prediction, whereas pipe distance emerged as an important predictor for HAAs [15]. This does not show that TTHM and HAA5 require different predictor sets, particularly because the Seoul HAA outcome was not reported specifically as HAA5. It does, however, provide field-scale evidence that shared candidate predictors need not make equal contributions across DBP targets.

Dataset composition is another source of variation. In the Ersan study, modelling performance changed when heterogeneous water sources were analysed together versus when more homogeneous NOM waters were considered separately, and the authors concluded that performance depended on both dataset characteristics and the DBP being modelled [14]. These results were obtained from datasets generated under uniform formation conditions and should not be interpreted as validation in full-scale distribution systems.

Accordingly, predictor relevance and model performance should be evaluated for each DBP outcome rather than assumed to transfer from TTHM to HAA5. Data-driven modelling should retain the physicochemical and distribution-system context of the target group while allowing shared predictors and

 ISBN: 978-605-80193-6-2

modelling frameworks where the data support them. The principal distinctions and their monitoring and modelling implications are summarised in Table I.

TABLE I. Key Distinctions Between TTHM and HAA5 Relevant to Drinking-Water Monitoring and Data-Driven Modelling

Factor	TTHM	HAA5	Monitoring implication	Modelling implication
Precursor characteristics	Formation depends on NOM composition and precursor fractions.	Constituent HAA classes can reflect precursor fractions that differ from those favouring THM formation [7].	Bulk NOM measures do not establish equivalent TTHM and HAA5 behaviour.	Organic-matter predictors should be evaluated separately for each outcome.
pH	Formation can be strongly pH-sensitive during chlorination.	No single directional response should be assumed; constituent HAA classes can respond differently to pH [8].	Changes in pH may alter the relationship between the two regulated groups.	The contribution of pH should be evaluated separately for each outcome.
Temperature	Formation kinetics can increase with temperature under chlorination conditions.	Temperature sensitivity can differ among HAA constituents and across conditions [8].	Temperature-related changes should not be assumed to affect the two groups proportionally.	Temperature may contribute differently to target-specific models.
Bromide and speciation	Bromide can shift composition toward brominated THM species.	Bromide can alter HAA composition, including brominated constituents of HAA5 [1].	Similar group totals can conceal changes in constituent composition.	Models of aggregate TTHM or HAA5 do not by themselves represent changes in constituent speciation.
Water age, contact time and loss processes	Continued formation can occur as contact or residence time increases, but the magnitude is system-dependent [5].	The aggregate response can reflect continued formation together with constituent-specific loss; no universal HAA5 trajectory should be assumed [5], [10].	Locations associated with high TTHM need not coincide with locations associated with high HAA concentrations [3].	Water-age and distribution-system variables should be evaluated for each target rather than assumed to transfer equally.
Spatial and temporal variability	Spatial and seasonal maxima are system-specific.	Spatial and seasonal maxima are also system-specific and may coincide with or diverge from those of TTHM [5], [6], [11], [12].	Monitoring locations, sampling periods and any quantitative surrogacy should be interpreted separately rather than assumed interchangeable.	Predictor relevance and predictive performance should be evaluated separately for each outcome.

Note: Mechanistic statements based on individual HAA species or HAA subclasses describe constituent-level behaviour and should not be interpreted as a universal directional response of aggregate HAA5.

VI. CONCLUSIONS

TTHM and HAA5 are produced within the same broad disinfection chemistry and are influenced by many of the same water-quality and operational factors, but this shared context does not make them interchangeable indicators. Differences in precursor characteristics, pH response, temperature sensitivity, speciation and formation or loss processes can lead to non-parallel behaviour within drinking-water distribution systems. Their spatial maxima and seasonal patterns may therefore coincide in some systems but diverge in others.

These differences have practical implications for both monitoring and data-driven modelling. Monitoring locations or sampling periods that characterise high TTHM concentrations may not provide equivalent information for HAA5, and any quantitative surrogacy between the two groups should be demonstrated rather than assumed. Similarly, common predictors and modelling frameworks can be used for both outcomes, but predictor relevance and predictive performance should be evaluated separately. TTHM and HAA5 should therefore be treated as related regulatory DBP groups whose degree of equivalence remains system- and context-dependent.

References

[1] Uyak, V., Ozdemir, K., & Toroz, I. (2007). Multiple linear regression modeling of disinfection by-products formation in Istanbul drinking water reservoirs. *Science of the Total Environment*, *378*(3), 269–280. https://doi.org/10.1016/j.scitotenv.2007.02.041

[2] U.S. Environmental Protection Agency. (2010). Comprehensive disinfectants and disinfection byproducts rules (Stage 1 and Stage 2): Quick reference guide (EPA 816-F-10-080).

[3] Speight, V. L., & Singer, P. C. (2005). Association between residual chlorine loss and HAA reduction in distribution systems. *Journal AWWA*, *97*(2), 82–91. https://doi.org/10.1002/j.1551-8833.2005.tb10827.x

[4] Villanueva, C. M., Kogevinas, M., & Grimalt, J. O. (2003). Haloacetic acids and trihalomethanes in finished drinking waters from heterogeneous sources. *Water Research*, *37*(4), 953–958. https://doi.org/10.1016/S0043-1354(02)00411-6

[5] Rodriguez, M. J., Sérodes, J.-B., & Levallois, P. (2004). Behavior of trihalomethanes and haloacetic acids in a drinking water distribution system. *Water Research*, *38*(20), 4367–4382. https://doi.org/10.1016/j.watres.2004.08.018

[6] Unrine, J. M., et al. (2024). Spatial and seasonal variation in disinfection byproducts concentrations in a rural public drinking water system: A case study of Martin County, Kentucky, USA. *PLOS Water*, *3*(3), e0000227. https://doi.org/10.1371/journal.pwat.0000227

ISBN: 978-605-80193-6-2

[7] Hua, G., & Reckhow, D. A. (2007). Characterization of disinfection byproduct precursors based on hydrophobicity and molecular size. *Environmental Science & Technology*, *41*(9), 3309–3315. https://doi.org/10.1021/es062178c

[8] Hua, G., & Reckhow, D. A. (2008). DBP formation during chlorination and chloramination: Effect of reaction time, pH, dosage, and temperature. *Journal AWWA*, *100*(8), 82–95. https://doi.org/10.1002/j.1551-8833.2008.tb09702.x

[9] Sfynia, C., Bond, T., Kanda, R., & Templeton, M. R. (2022). Simultaneous prediction of trihalomethanes, haloacetic acids, haloacetonitriles and haloacetamides using simulated distribution system tests. *Environmental Science: Water Research & Technology*, *8*, 742–756. https://doi.org/10.1039/D1EW00824B

[10] Tung, H.-H., & Xie, Y. F. (2009). Association between haloacetic acid degradation and heterotrophic bacteria in water distribution systems. *Water Research*, *43*(4), 971–978. https://doi.org/10.1016/j.watres.2008.11.041

[11] Parvez, S., Rivera-Núñez, Z., Meyer, A., & Wright, J. M. (2011). Temporal variability in trihalomethane and haloacetic acid concentrations in Massachusetts public drinking water systems. *Environmental Research*, *111*(4), 499–509. https://doi.org/10.1016/j.envres.2010.12.008

[12] Uyak, V., et al. (2014). Spatial and seasonal variations of disinfection byproducts (DBPs) in drinking water distribution systems of Istanbul City, Turkey. *Environmental Forensics*, *15*(2), 190–205. https://doi.org/10.1080/15275922.2014.890145

[13] Delpla, I., Florea, M., Pelletier, G., & Rodriguez, M. J. (2018). Optimizing disinfection by-product monitoring points in a distribution system using cluster analysis. *Chemosphere*, *208*, 512–521. https://doi.org/10.1016/j.chemosphere.2018.06.009

[14] Ersan, G., Goz, E., & Karanfil, T. (2025). Performance analysis of machine learning algorithms for the prediction of disinfection byproducts formation during chlorination: Effect of background water characteristics. *Journal of Environmental Management*, *389*, 126144. https://doi.org/10.1016/j.jenvman.2025.126144

[15] Park, J. W., Lee, Y., Noh, J. H., & Maeng, S. K. (2025). Historical, seasonal, and spatial variations in chlorination disinfection byproducts within Seoul's drinking water distribution system: A decade-long assessment. *Journal of Environmental Chemical Engineering*, *13*(3), 116390. https://doi.org/10.1016/j.jece.2025.116390

ISBN: 978-605-80193-6-2

Machine Learning in Struvite Recovery: A Focused Review of Process Inputs and Prediction Targets

Berkin Üstünyıldız[1]*

[1]Düzce University, Kaynaşlı Vocational School, Department of Environmental Protection Technologies, Düzce, Türkiye

* Corresponding author e-mail: berkinuy@gmail.com

***Abstract*—Machine learning has been applied to struvite recovery to relate physicochemical and operational conditions to recovery-related process outcomes. This focused review examines the variables used as model inputs and the process quantities selected as prediction targets in published machine-learning studies of struvite recovery. Across the studies considered, recurring inputs include nutrient concentrations, molar ratios, pH, reaction conditions and, in full-scale applications, broader reactor and operational information. However, input sets are not standardized and vary with data source, process configuration and modelling objective. Prediction targets are also heterogeneous and include nutrient recovery, phosphorus concentrations, removal-based metrics, product purity, Conversion and Yield. These quantities are not interchangeable and represent different aspects of process performance or process state. Reported model accuracy should therefore be interpreted together with the dataset context and the exact definition of the predicted target. The reviewed applications include process prediction, model interpretation, optimization and operational decision support. Recent full-scale studies extend the evidence base to operational prediction and decision support, while broader operational evidence and independent validation remain important for assessing model transferability. Overall, meaningful interpretation of predictive results requires clear identification of both the information supplied to the model and the specific process quantity that the model is designed to predict.**

***Keywords*—Struvite recovery, machine learning, phosphorus recovery, process variables, prediction targets, wastewater treatment**

I. INTRODUCTION

Phosphorus recovery from wastewater can reduce phosphorus losses while converting part of the nutrient load into reusable products. Among the available recovery routes, controlled struvite crystallization has received sustained attention because phosphorus can be precipitated as magnesium ammonium phosphate hexahydrate ($MgNH_4PO_4 \cdot 6H_2O$). Struvite formation has been investigated from laboratory to full scale, and controlled precipitation can combine phosphorus removal from wastewater with production of a phosphorus-containing solid with fertilizer value [1].

Struvite precipitation, however, is governed by several interacting chemical and operational conditions rather than by a single process variable. The availability of magnesium, ammonium and phosphate, together with pH, supersaturation and mixing conditions, influences crystallization. Wastewater composition introduces further variability. Foreign ions, organic matter and other matrix constituents can affect reaction kinetics, nutrient recovery, crystal characteristics and product purity, while operational variables such as chemical dosing and reactor conditions also influence process performance [2]. These interacting factors make struvite recovery a multivariable problem in which the relevant inputs may differ according to wastewater matrix, process configuration and the outcome being evaluated.

Data-driven modelling provides one approach for examining these multivariable relationships. Machine-learning models can relate measured physicochemical and operational inputs to selected process outputs without requiring the complete reaction mechanism to be explicitly represented in the predictive model. Leng et al. [3] described this capability in the context of struvite precipitation, where multiple process variables were used as inputs to predict recovery-related outcomes and subsequently support process optimization. More broadly, recent struvite studies have applied machine learning to datasets derived from published experiments as well as operational treatment systems, creating models for different aspects of nutrient recovery and reactor performance. This expanding use of data-driven methods also makes the definition of both model inputs and predicted outputs increasingly important when results are interpreted across studies.

This focused review examines the process variables used as inputs in machine-learning studies of struvite recovery and compares the process outcomes selected as prediction targets. Particular attention is given to how chemical, operational and reactor-related information is represented in predictive models and to the distinction among recovery, removal, concentration-based, product-quality and reactor-performance outcomes. By examining these two elements together, the review aims to clarify what information current models use and what, precisely, their predictions represent.

II. DATA SOURCES AND PROCESS INPUTS IN MACHINE-LEARNING STUDIES

The data used for predictive modelling of struvite recovery span substantially different experimental and operational contexts. Early work by Forrest et al. [4] used historical operating data from technical-scale crystallization reactors treating digester supernatant and centrate at two wastewater treatment plants. In contrast, several later studies assembled datasets from previously published experiments. Nageshwari et al. [5] compiled 100 input-output records from approximately 85 published studies covering struvite precipitation from different wastewater sources, whereas Leng et al. [3] collected 504 phosphorus-recovery records from 20 studies using synthetic wastewater, with smaller datasets used for nitrogen-recovery and multi-target modelling. Aldaach et al. [6] extended this literature-derived approach to organic-waste digestates, drawing on

ISBN: 978-605-80193-6-2

laboratory- and pilot-scale studies involving feedstocks such as sewage sludge, swine manure and chicken slurry. Zhang et al. [7] used a broader compiled dataset covering several wastewater and sludge-related substrates, while more recent studies by Aguiar et al. [8] and Lybik et al. [9] used full-scale operational data. Aguiar et al. analyzed more than five years of records from a municipal fluidized-bed struvite crystallizer, whereas Lybik et al. developed predictive models for a full-scale MagPrex reactor at a water resource recovery facility in Denver.

Despite differences in data origin, several physicochemical inputs recur across the representative studies. Variables describing the concentrations or relative availability of the principal struvite-forming ions are particularly prominent. Nageshwari et al. [5] included phosphate, ammonium and externally added magnesium concentrations among the model inputs, together with pH and operating variables. Leng et al. [3] represented nutrient and magnesium availability partly through molar ratios, using Mg:P, N:P and initial P concentration for phosphorus-recovery modelling, with Mg:N replacing Mg:P in the nitrogen-recovery model. Zhang et al. [7] used a wider chemical description that included initial P, N, Mg, Ca and K concentrations as well as Mg:P and N:P ratios. These examples show that models may describe related chemical conditions through absolute concentrations, molar ratios, or both. Consequently, apparently similar models should not be assumed to represent nutrient availability in an identical way.

Operating conditions constitute a second recurring group of predictors. pH appears across several of the representative studies, but its role should be interpreted at the level of each individual model rather than treated as universally dominant. Reaction temperature, reaction or retention time, and mixing or stirring intensity are also repeatedly represented. Nageshwari et al. [5], for example, included reaction temperature, stirring rate, reaction time and retention time among the process-related inputs. Leng et al. [3] used reaction time, temperature and stirring speed alongside composition-related variables. In the digestate-focused analysis of Aldaach et al. [6], the predictor combinations ranged from relatively compact sets based on pH and molar ratios to broader combinations including reaction temperature, mixing speed and hydraulic retention time. This variation reflects differences in both the available source data and the modelling objective rather than a standardized input specification across the literature.

Studies using more compositionally complex matrices also incorporated variables describing potentially competing ions and broader matrix conditions. Aldaach et al. [6] included sodium, potassium, calcium, magnesium, ammonium and phosphate concentrations in the more extensive predictor combinations. Zhang et al. [7] similarly incorporated Ca and K alongside the main struvite-forming components. Such variables can help represent matrix conditions associated with competing precipitation and recovered-product composition, particularly in studies that also consider product purity. Calcium-phosphate co-precipitation, for example, has been associated with mixed calcium-phosphate (CaP)-struvite solids and reduced product purity [10].

The input space becomes broader in the full-scale examples considered here. Aguiar et al. [8] assembled water-quality measurements, crystallizer operating variables, chemical-dosing information, recycle conditions and environmental variables, and subsequently generated additional features such as mineral saturation indices, molar ratios, hydraulic retention times, seasonal terms and lagged reactor-performance variables. Feature selection was then used to retain only the highest-ranked variables for individual models. The importance of distinguishing the initial feature pool from the final model inputs is therefore especially clear in this type of study. Lybik et al. [9] also applied feature selection to multivariable plant-monitoring data in a full-scale MagPrex system; effluent soluble magnesium was identified as the most important feature of the effluent orthophosphate (OP) eXtreme Gradient Boosting (XGBoost) model, and the resulting predictions were subsequently used to support $MgCl_2$ dosing recommendations.

Taken together, the studies considered here do not use a single standardized input set. The laboratory- and literature-derived studies examined in this review are largely structured around chemical composition, molar ratios and controlled operating conditions, whereas the full-scale examples considered here incorporate a broader range of reactor, recycle, dosing and process-state information. This distinction is important when comparing model results, because differences in predictive performance may reflect not only the modelling algorithm but also substantial differences in what information was available to the model. Representative studies and their main inputs, prediction targets and applications are summarized in Table I.

TABLE I. Representative Machine-Learning Studies in Struvite Recovery and Their Main Inputs and Prediction Targets

Study	Data source	Main model inputs	Main predictive model(s)	Prediction target(s)	Main application
Forrest et al. [4]	Historical operational data from technical-scale struvite crystallization reactors treating digester supernatant and centrate at two wastewater treatment plants	Conductivity; reactor upflow velocity; recycle ratio; Mg^{2+}; NH_4-N; PO_4-P*	Back-propagation artificial neural network (ANN; NeuStruvite)	Final effluent PO_4-P concentration	Effluent-phosphate prediction and sensitivity analysis in technical-scale struvite-recovery systems

ISBN: 978-605-80193-6-2

Study	Data source	Main model inputs	Main predictive model(s)	Prediction target(s)	Main application
Nageshwari et al. [5]	Literature-derived published struvite data from multiple wastewater sources	pH; reaction temperature; phosphate, ammonium and externally added Mg concentrations; stirring rate; reaction time; retention time; drying temperature; drying time	Extra Trees; Random Forest; eXtreme Gradient Boosting (XGBoost)	Phosphate recovery; ammonium recovery	Prediction of phosphate and ammonium recovery with experimental validation
Leng et al. [3]	Literature-derived data from synthetic-wastewater studies	Mg:P; N:P; initial P; pH; reaction time; temperature; stirring speed; Mg:N replaced Mg:P in the N model	Random Forest; Gradient Boosting Regressor	P recovery; N recovery; simultaneous P/N recovery	Prediction and optimization with experimental verification
Aldaach et al. [6]	Literature-derived laboratory- and pilot-scale organic-waste digestate data	Combination-dependent inputs including pH, Mg:P, N:P and reaction temperature; broader combinations also included mixing speed, hydraulic retention time (HRT) and Na, K, Ca, Mg, NH_4 and PO_4 concentrations	k-nearest neighbors (KNN); Random Forest; XGBoost	PO_4^{3-} recovery efficiency; NH_4^+ recovery efficiency	Nutrient-recovery prediction across digestates and parameter combinations
Zhang et al. [7]	Literature-derived multi-substrate experimental records	Substrate; pH; temperature; initial P, N, Mg, Ca and K; stirring speed; reaction time; Mg:P; N:P	Random Forest; Histogram-based Gradient Boosting Regressor (HGBR); XGBoost	Struvite recovery efficiency (phosphate-removal-based); product purity	Prediction and interpretation followed by multi-objective optimization
Aguiar et al. [8]	More than five years of full-scale municipal fluidized-bed reactor operational data	Target- and boundary-specific retained features after feature selection, drawn from feed/recycle chemistry, chemical dosing, flow/recycle variables, and engineered thermodynamic/operational variables	XGBoost; Ridge regression	Conversion; Yield	Full-scale prediction of soluble-P conversion/removal and P recovered as struvite, with distinct drivers
Lybik et al. [9]	Full-scale MagPrex struvite precipitation reactor operational data	Selected multivariable plant-monitoring inputs after feature selection; effluent soluble Mg was the most important feature in the effluent-OP XGBoost model	Ridge regression; XGBoost	Influent OP; effluent OP; OP removal	Performance prediction and $MgCl_2$ dosing decision support, including full-scale deployment

* Inputs listed for Forrest et al. are those explicitly evaluated in the ANN sensitivity analysis; additional predictors were not inferred from the mechanistic-model description.

III. PREDICTION TARGETS

The prediction targets used in struvite-related machine-learning studies represent different process quantities and are not interchangeable indicators of struvite-recovery performance. Several studies have focused directly on nutrient recovery. Nageshwari et al. [5] modelled phosphate and ammonium recovered as struvite, while Leng et al. [3] developed both single-target and multi-target models for phosphorus and nitrogen recovery. Aldaach et al. [6] similarly predicted phosphate and ammonium recovery percentages from organic-waste digestates. These targets describe nutrient recovery performance, but they should not automatically be treated as equivalent to the amount or quality of recovered struvite product.

Other studies have predicted phosphorus concentrations or removal rather than a recovered-product quantity. Forrest et al. [4], for example, used an artificial neural network to predict final effluent phosphate concentration in struvite crystallization systems. Lybik et al. [9] extended this type of operational prediction at full scale by modelling influent OP concentration, effluent OP concentration and OP removal in a MagPrex reactor. The effluent-OP predictions were subsequently used to support $MgCl_2$ dosing recommendations. These outcomes are operationally useful, but an influent or effluent concentration, or a removal-based metric, does not by itself specify how much phosphorus is retained as a recoverable struvite product.

The distinction becomes especially clear when recovery-related outcomes are modelled together with product quality. Zhang et al. [7] treated recovery efficiency and product purity as separate targets. In that study, the reported recovery efficiency was calculated from phosphate removal, whereas purity represented the percentage of struvite in the recovered solid. The two targets showed different model responses and a trade-off during subsequent multi-objective optimization, indicating that high phosphate removal does not necessarily imply equally high struvite purity.

Aguiar et al. [8] introduced another important distinction in a full-scale fluidized-bed crystallizer by modelling Conversion and Yield separately. Conversion represented phosphorus removal from the solution phase, whereas Yield represented the fraction of phosphorus recovered as struvite. This distinction addresses a limitation of treating soluble phosphorus removal as equivalent to product recovery. A reactor may achieve substantial phosphorus conversion while losing fine struvite particles or other particulate phosphorus with the effluent, resulting in a different recovered-product yield.

Taken together, the representative studies use prediction targets that span nutrient-recovery percentages, influent or effluent phosphorus concentrations, removal-based metrics, product purity, Conversion, and Yield. These outcomes answer different process questions and therefore should not be compared as if they represented the same endpoint. Reported model accuracy must be interpreted together with the

ISBN: 978-605-80193-6-2

exact target definition, particularly when comparing results across studies that use different datasets, reactor contexts and performance measures.

IV. CURRENT APPLICATIONS OF MACHINE LEARNING

Across the studies considered here, the most direct application of machine learning is the prediction of struvite-related process quantities from physicochemical and operational inputs. These predictions include nutrient recovery, effluent phosphorus concentration, product purity, and full-scale reactor performance indicators. In several studies, predictive models were also used beyond simple estimation. Nageshwari et al. [5], for example, experimentally compared model predictions with struvite precipitation results, while Leng et al. [3] used trained models to identify operating conditions for high phosphorus and nitrogen recovery and subsequently verified selected conditions experimentally.

A second application is the identification of influential process variables and the use of predictive models to support process optimization. Feature-importance and interpretation methods have been applied to examine how individual inputs relate to predicted targets, although these relationships should be interpreted as model-specific rather than causal. Leng et al. [3] used feature-importance, SHapley Additive exPlanations (SHAP), and partial-dependence analyses to examine the effects of process conditions on phosphorus and nitrogen recovery. Zhang et al. [7] separately examined the drivers of recovery efficiency and product purity and then coupled the selected predictive models with the Non-dominated Sorting Genetic Algorithm II (NSGA-II) as a distinct multi-objective optimization step.

Full-scale studies also demonstrate the use of predictive modelling for operational decision support. Aguiar et al. [8] used historical full-scale data to predict Conversion and Yield and to identify different process variables associated with these outcomes. Lybik et al. [9] linked influent and effluent OP predictions to $MgCl_2$ dosing recommendations in a full-scale MagPrex reactor and evaluated the approach during full-scale deployment. Together, these studies illustrate an application spectrum from prediction and model interpretation to optimization and operational decision support.

V. CURRENT LIMITATIONS AND RESEARCH NEEDS

Among the studies considered here, a recurring limitation is the representativeness of the data used for model development. Several models have been developed from literature-derived or synthetic-wastewater datasets rather than long-term operational records. Leng et al. [3], for example, explicitly limited their analysis to synthetic wastewater and identified testing with more complex synthetic matrices and real wastewater as a priority for future work. Zhang et al. [7] similarly noted that their compiled dataset was derived mainly from laboratory-scale studies and lacked important information associated with full-scale operation. The purity subset was also substantially smaller than the recovery-efficiency dataset, which may limit model generalizability.

A second issue is heterogeneity in both model inputs and prediction targets. Literature-derived datasets combine experiments conducted under different chemical conditions, reactor configurations and reporting practices, while the studies reviewed here predict quantities ranging from nutrient recovery to concentrations, removal-based metrics, purity, Conversion and Yield. Consequently, model-performance values reported across different studies should not be treated as directly comparable. Differences in accuracy may reflect differences in data composition, target definition and validation design as much as differences between algorithms.

Future work would therefore benefit from larger and more diverse datasets from real wastewater and operating facilities, together with validation on data that are independent of the model-development dataset. Full-scale studies such as those of Aguiar et al. [8] and Lybik et al. [9] demonstrate the feasibility of operational prediction, but evidence from additional facilities and process conditions is needed to evaluate transferability. Clear reporting of target definitions is equally important, particularly where terms such as removal, recovery, conversion and yield describe different process quantities. Expanding full-scale evidence while preserving explicit target definitions would make comparisons among predictive models more meaningful and improve their practical interpretation.

VI. CONCLUSIONS

Machine-learning studies of struvite recovery use recurring physicochemical and operational variables, including nutrient concentrations, molar ratios, pH, reaction conditions and, in full-scale applications, broader process and reactor information. However, the input sets are not standardized and reflect differences in data source, process configuration and modelling objective. Prediction targets are similarly heterogeneous. Nutrient recovery, phosphorus concentrations, removal-based metrics, product purity, Conversion and Yield represent different process quantities and should not be treated as interchangeable measures of struvite-recovery performance. Consequently, reported model accuracy cannot be interpreted independently of the dataset context and the exact definition of the predicted target. Although recent full-scale studies show that data-driven models can be applied to operational struvite-recovery systems, broader evidence from diverse operating conditions and independent validation would strengthen assessment of model transferability and practical relevance. Overall, meaningful interpretation of predictive results requires clear identification of both the information supplied to the model and the specific process quantity that the model is designed to predict.

References

[1] Le Corre, K. S., Valsami-Jones, E., Hobbs, P., & Parsons, S. A. (2009). Phosphorus recovery from wastewater by struvite crystallization: A review. *Critical Reviews in Environmental Science and*

 ISBN: 978-605-80193-6-2

Technology, 39(6), 433–477. https://doi.org/10.1080/10643380701640573

[2] Li, B., et al. (2019). Phosphorus recovery through struvite crystallisation: Recent developments in the understanding of operational factors. *Journal of Environmental Management*, 248, 109254. https://doi.org/10.1016/j.jenvman.2019.07.025

[3] Leng, L., et al. (2024). Machine-learning-aided prediction and optimization of struvite recovery from synthetic wastewater. *Journal of Water Process Engineering*, 58, 104896. https://doi.org/10.1016/j.jwpe.2024.104896

[4] Forrest, A. L., Fattah, K. P., Mavinic, D. S., & Koch, F. A. (2007). Application of artificial neural networks to effluent phosphate prediction in struvite recovery. *Journal of Environmental Engineering and Science*, 6(6), 713–725. https://doi.org/10.1139/S07-023

[5] Nageshwari, K., Senthamizhan, V., & Balasubramanian, P. (2022). Sustaining struvite production from wastewater through machine learning based modelling and process validation. *Sustainable Energy Technologies and Assessments*, 53, 102608. https://doi.org/10.1016/j.seta.2022.102608

[6] Aldaach, H., Zaki, M. T., & Orner, K. D. (2025). Improving prediction of nutrient recovery via struvite precipitation from organic waste digestate. *Environmental Engineering Science*, 42(2), 53–67. https://doi.org/10.1089/ees.2024.0116

[7] Zhang, X., Sun, S., Han, X., & Shu, S. (2025). Optimizing struvite recovery from wastewater treatment: Insights from machine learning models. *Journal of Water Process Engineering*, 77, 108520. https://doi.org/10.1016/j.jwpe.2025.108520

[8] Aguiar, S. E., et al. (2026). Elucidating drivers of full-scale struvite recovery system performance with thermodynamic and data-driven modeling. *Water Research*, 290, 125122. https://doi.org/10.1016/j.watres.2025.125122

[9] Lybik, J. E., Love, N. G., Maltos, R., Goldman, J., & Newhart, K. B. (2026). Development of data-driven models for performance prediction and chemical dosing of a full-scale controlled phosphorus precipitation reactor. *ACS ES&T Engineering*, 6(8), 2298–2310. https://doi.org/10.1021/acsestengg.6c00351

[10] Le Corre, K. S., Valsami-Jones, E., Hobbs, P., & Parsons, S. A. (2005). Impact of calcium on struvite crystal size, shape and purity. *Journal of Crystal Growth*, 283(3–4), 514–522. https://doi.org/10.1016/j.jcrysgro.2005.06.012

ISBN: 978-605-80193-6-2

Ağır Betonların TPU veya PLA Hasır Donatı Kullanılarak Optimizasyonu

Ömer Şimşek[1], Prof. Dr. Sadık Alper Yıldızel[2], Doç. Dr. Selin Erzin[3]
[1]Karamanoğlu Mehmetbey Üniversitesi, Mühendislik Fakültesi, İnşaat Mühendisliği Bölümü, Karaman, Türkiye
[2]Karamanoğlu Mehmetbey Üniversitesi, Fen Bilimleri Enstitüsü, Karaman, Türkiye
[3]Dokuz Eylül Üniversitesi, Fen Fakültesi, Fizik Bölümü, İzmir, Türkiye
Makale Notu: Bu çalışma, bir yüksek lisans tezinden türetilmiştir.
* Sorumlu yazar e-mail: omersimsek@windowslive.com

ÖZET—Ağır betonlar, yüksek birim hacim kütleleri nedeniyle nükleer tesisler, tıbbi görüntüleme ve radyoterapi alanları, radyoaktif atık depolama birimleri ve bazı özel mühendislik uygulamalarında kullanılmaktadır. Bu betonların üretiminde barit, hematit ve manyetit gibi yüksek yoğunluklu agregalar tercih edilmektedir. Bununla birlikte ağır betonlarda yoğunluğun artırılması ile birlikte işlenebilirlik, mekanik davranış ve kırılma karakterinin birlikte değerlendirilmesi gerekmektedir. Son yıllarda üç boyutlu baskı teknolojisi, beton ve çimento esaslı kompozitlerde polimerik kafes ve hasır yapıların üretiminde kullanılmaya başlanmıştır. Bu çalışmada, yüksek yoğunluklu agregalar kullanılarak hazırlanan ağır beton içerisine üç boyutlu baskı yöntemiyle üretilen TPU veya PLA esaslı polimerik hasır donatı yerleştirilmiştir. Deney esnasında döküm malzemeden imal edilmiş Küp kalıp ve Üç boyutlu yazıcıdan imal edilmiş dikdörtgen kalıp kullanılmıştır. Bu numunelerin birim hacim kütlesi, mekanik özellikleri ve çatlak-kırılma davranışı deneysel olarak değerlendirilmiştir. Deneysel çalışmada çimento, silis dumanı, süperakışkanlaştırıcı, ince barit, iri barit, mikro barit ve hematit kullanılmıştır. Küp ve dikdörtgen numuneler hazırlanarak belirlenen kür koşullarında bekletilmiş ve mekanik deneylere tabi tutulmuştur. Deneysel süreçte numunelerin birim hacim kütleleri belirlenmiş ve mekanik davranışları incelenmiştir. Ayrıca deney sonrası çatlak gelişimi ve kırılma biçimleri incelenmiştir. Elde edilen deney sonuçları, yükleme sürecindeki davranış ve kırılma sonrası fiziksel görünümler birlikte dikkate alınarak değerlendirilmiştir. Bu değerlendirmede polimerik hasırların çatlak ilerlemesiyle ilişkisi de ele alınmıştır. Çalışma, üç boyutlu baskılı polimerik hasırların yüksek yoğunluklu betonlarda kullanılabilirliğinin değerlendirilmesine yönelik deneysel bir temel oluşturmuştur.

Anahtar Kelimeler—Ağır beton, barit, 3B baskı, polimerik hasır, TPU veya PLA

I. GİRİŞ

Ağır betonlar, normal betonlara göre daha yüksek birim hacim kütlesine sahip özel beton türleridir. Bu betonlar yüksek yoğunluğun gerekli olduğu nükleer tesisler, tıbbi görüntüleme ve radyoterapi odaları, radyoaktif atık depolama alanları ve bazı özel mühendislik yapılarında kullanılmaktadır. Barit, hematit ve manyetit gibi yüksek yoğunluklu agregalar, ağır beton üretiminde yaygın olarak tercih edilmektedir [1], [2], [5].

Ağır betonların önemli kullanım alanlarından biri radyasyon zırhlamadır. Betonun radyasyon azaltma kapasitesi; yoğunluk, agrega türü, beton kalınlığı ve radyasyon enerjisi gibi değişkenlerden etkilenmektedir. Barit ve hematit içeren betonlar üzerine yapılan araştırmalar, yüksek yoğunluklu agregaların radyasyon zayıflatma kapasitesinin artırılmasına katkı sağlayabildiğini göstermektedir [6]–[10], [17], [18].

Bununla birlikte ağır beton tasarımında yalnızca yoğunluğun artırılması yeterli değildir. İşlenebilirlik, segregasyon, dayanım ve kırılma davranışının birlikte değerlendirilmesi gerekmektedir [1], [2]. Betonun çekme kapasitesinin sınırlı olması ve gevrek kırılma eğilimi, çatlak kontrolünü önemli hâle getirmektedir.

Üç boyutlu baskı teknolojisinin gelişmesiyle birlikte polimer esaslı kafes ve hasır yapıların çimento esaslı kompozitlerde donatı olarak kullanılması yeni bir araştırma alanı oluşturmuştur. Bu yapıların geometrisi ve üretim yönü kontrollü biçimde değiştirilebilmekte; bazı araştırmalarda süneklik, çatlak ilerlemesi ve eğilme davranışında olumlu değişimler raporlanmaktadır [3], [12]–[16].

Bu çalışmada, yüksek yoğunluklu agregalarla hazırlanan ağır beton içerisine üç boyutlu baskılı polimerik hasır donatı yerleştirilerek numunelerin mekanik ve kırılma davranışı deneysel olarak incelenmiştir.

II. LİTERATÜR VE ARAŞTIRMA FİKRİ

Ağır beton üzerine yapılan araştırmalarda yüksek yoğunluklu agregaların betonun birim hacim kütlesi, mekanik özellikleri ve radyasyon zırhlama kapasitesi üzerindeki etkileri incelenmiştir. Khalaf vd. ağır betonların bileşenlerini, özelliklerini ve kullanım alanlarını kapsamlı biçimde değerlendirmiştir [1]. Topçu ise barit agregalı ağır betonların fiziksel ve mekanik özelliklerini incelemiştir [5].

Barit ve hematit içeren betonların radyasyon zırhlama özellikleri de çeşitli çalışmalarda ele alınmıştır. Akkurt vd. farklı agregaların gama ışını zırhlama üzerindeki etkisini incelemiş; Demir vd. barit içeren betonların zayıflatma özelliklerini değerlendirmiştir [6], [7]. Gencel vd. ise farklı hematit oranlarının gama ve nötron zırhlama özellikleri üzerindeki etkisini araştırmıştır [8].

Betonun çatlak ve süneklik davranışının geliştirilmesine yönelik çalışmalarda 3B baskılı polimerik kafeslerin kullanımı dikkat çekmektedir.

ISBN: 978-605-80193-6-2

Salazar vd. PLA ve ABS esaslı polimerik kafeslerin betonun sünekliğini artırabildiğini ve çatlak yollarını değiştirebildiğini bildirmiştir [3]. Xu vd. fonksiyonel olarak derecelendirilmiş polimerik kafeslerin çimento esaslı kompozitlerde kullanımını incelemiş; Tang vd. ise derecelendirilmiş kafeslerin eğilme davranışı üzerindeki etkisini değerlendirmiştir [12], [14].

Bu çalışmalar doğrultusunda, ağır betonun yüksek yoğunluk özelliği ile üç boyutlu baskılı polimerik hasır donatının çatlak gelişimi ve kırılma davranışı üzerindeki etkisinin birlikte değerlendirilmesi ele alınmıştır.

III. AĞIR BETON KAVRAMI VE KULLANIM ALANLARI

A. Ağır Betonlarda Agrega Seçimi

Ağır betonların temel bileşenlerinden biri yüksek yoğunluklu agregalardır. Barit, hematit ve manyetit gibi agregalar yüksek özgül ağırlıkları nedeniyle betonun birim hacim kütlesini artırmak amacıyla kullanılmaktadır [1], [2], [5]. Bu çalışmada ince barit, iri barit, mikro barit ve hematit kullanılmıştır. Farklı tane boyutlarının birlikte kullanılmasıyla agrega dağılımının dengeli biçimde oluşturulması amaçlanmıştır. Farklı tane boyutlarına sahip agregaların genel görünümü Şek. 1'de verilmiştir.



Şek. 1. Deneysel çalışmada kullanılan mikro barit, ince barit ve iri barit agregalarının görünümü.

B. Radyasyon Zırhlama Açısından Ağır Beton

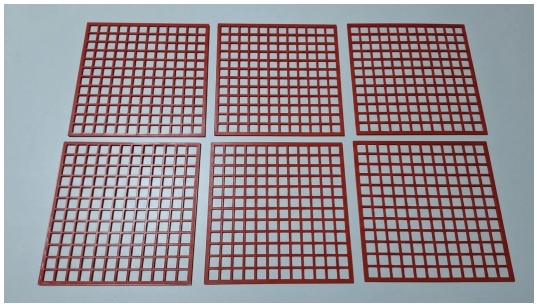
Ağır betonların önemli uygulama alanlarından biri radyasyon zırhlamadır. Özellikle yüksek yoğunluklu agregaların kullanılması, betonun gama ışınlarına karşı zayıflatma kapasitesinin artırılmasına katkı sağlayabilmektedir [6]–[10]. Bu çalışma kapsamında doğrudan radyasyon zırhlama deneyi yapılmamış; radyasyon zırhlama, kullanılan ağır betonun potansiyel uygulama alanlarından biri olarak ele alınmıştır.

C. 3B Baskılı Polimerik Kafeslerin Betonda Kullanımı

Üç boyutlu baskı teknolojisi, karmaşık geometrilerin kontrollü şekilde üretilmesine olanak sağlamaktadır. Bu özellik, beton içerisinde kullanılabilecek kafes veya hasır geometrilerin üretiminde avantaj sağlamaktadır. Literatürde 3B baskılı polimerik kafeslerin çimento esaslı kompozitlerde süneklik, eğilme davranışı ve çatlak gelişimini etkileyebildiği bildirilmektedir [3], [12]–[15].

D. PLA'nın Donatı Malzemesi Olarak Değerlendirilmesi

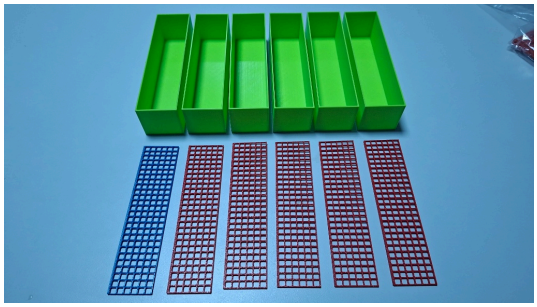
PLA, üç boyutlu baskı uygulamalarında yaygın kullanılan termoplastik bir polimerdir. Üretim kolaylığı ve farklı geometrilerde baskıya uygunluğu nedeniyle polimerik kafes donatı çalışmalarında kullanılabilmektedir. Salazar vd. PLA esaslı 3B baskılı kafesleri beton içerisinde kullanarak özellikle süneklik davranışındaki değişimi incelemiştir [3]. Bu çalışmada TPU/PLA esaslı hasır donatı küp numunelerde kullanılmıştır. Küp numuneler için hazırlanan polimerik hasır donatı Şek. 2'de gösterilmiştir.



Şek. 2. Küp numuneler için 3B yazıcı ile hazırlanan polimerik hasır donatının görünümü.

E. TPU'nun Donatı Malzemesi Olarak Değerlendirilmesi

TPU, esnekliği ve deformasyon kapasitesi yüksek bir termoplastik polimerdir. Üç boyutlu baskı ile farklı kafes geometrilerinde üretilebilmektedir. TPU kafeslerin mekanik davranışı, yapı mimarisi ve baskı yönü gibi değişkenlerden etkilenmektedir [4]. Bu özellik, TPU esaslı hasırların beton içerisindeki çatlak ve deformasyon davranışının incelenmesi açısından önem taşımaktadır. Bu çalışmada TPU/PLA esaslı hasır donatı dikdörtgen numunelerde kullanılmıştır. Dikdörtgen numuneler için hazırlanan polimerik hasır donatılar Şek. 3'te gösterilmiştir.



Şek. 3. Dikdörtgen numuneler için 3B yazıcı ile hazırlanan TPU esaslı polimerik hasır donatıların görünümü.

IV. ÇALIŞMANIN DENEYSEL YAKLAŞIMI

A. Beton Karışımının Hazırlanması

Deneysel çalışmada ağır beton karışımı çimento, silis dumanı, su, süperakışkanlaştırıcı, ince barit, iri barit, mikro barit ve hematit kullanılarak hazırlanmıştır. Karışım bileşenlerinin homojen dağılması ve ağır agregaların beton içerisinde dengeli biçimde yayılması gözetilmiştir. Beton karışımının hazırlanması Şek. 4'te gösterilmiştir.

ISBN: 978-605-80193-6-2



Şek. 4. Ağır beton karışımının deneysel çalışma kapsamında hazırlanması.

B. Polimerik Hasırın Hazırlanması

Polimerik hasır donatı üç boyutlu yazıcı ile üretilmiştir. Küp numunelerde ve dikdörtgen numunelerde TPU/PLA esaslı polimerik hasır donatı kullanılmıştır. TPU fiber (8 mm) / PLA fiber (6 mm) deney programında ilgili numune gruplarında kullanılmış; hasır geometrileri numune boyutlarına uygun olacak şekilde tasarlanmış ve baskı sonrasında geometrik bütünlüğün korunması gözetilmiştir.

C. Numunelerin Üretilmesi ve Kür Edilmesi

Hazırlanan beton karışımı uygun kalıplara yerleştirilmiştir. Küp numunelerde ve dikdörtgen numunelerde TPU/PLA esaslı polimerik hasır donatı kullanılmıştır. Hasırların kalıp içerisindeki konumu deney programına uygun biçimde ayarlanmıştır. Numuneler kalıptan çıkarıldıktan sonra belirlenen kür koşullarında bekletilmiştir. Küp numune kalıbına beton karışımının yerleştirilmesi Şek. 5'te gösterilmiştir.



Şek. 5. Beton karışımının küp numune kalıbına yerleştirilmesi.

D. Mekanik Deneyler

Deneysel çalışmada numunelerin birim hacim kütlesi ve basınç dayanımı belirlenmiştir. Elde edilen deney sonuçları numunelerin mekanik davranışları açısından değerlendirilmiştir. Basınç deneyi sırasında TPU/PLA hasır donatılı küp numunenin kırılma davranışı Şek. 6'da, elde edilen yük-deplasman grafiği ise Şek. 7'de verilmiştir.



Şek. 6. TPU/PLA hasır donatılı küp numunenin basınç deneyi sırasında kırılma davranışı.

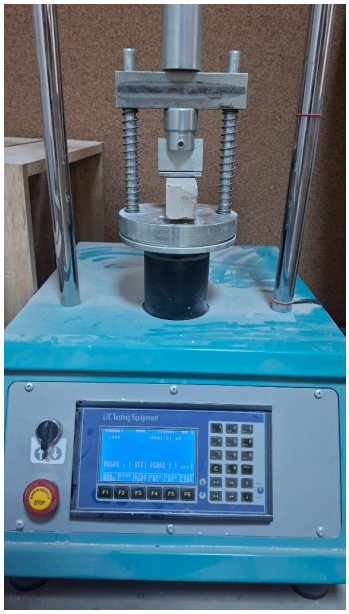


Şek. 7. TPU/PLA hasır donatılı küp numunenin basınç deneyi sırasında elde edilen yük-deplasman grafiği.

E. Kırılma Davranışının İncelenmesi

Mekanik deneyler sırasında oluşan çatlakların başlangıcı, ilerleme biçimi ve nihai kırılma davranışı gözlemlenmiştir. Deney sonrasında kırılma yüzeyleri incelenerek polimerik hasırların çatlak ilerlemesi üzerindeki davranışı değerlendirilmiştir. TPU/PLA hasır donatılı dikdörtgen numunenin mekanik deney sırasındaki kırılma davranışı Şek. 8'de, deney sonrasındaki kırılmış görünümü ise Şek. 9'da gösterilmiştir.

ISBN: 978-605-80193-6-2



Şek. 8. TPU/PLA hasır donatılı dikdörtgen numunenin mekanik deney sırasında kırılma davranışının incelenmesi.



Şek. 9. TPU/PLA hasır donatılı dikdörtgen numunenin deney sonrasında kırılmış görünümü.

TABLO I. Çalışmada Kullanılan Beton Bileşenleri

Bileşen	Miktar ve çalışmadaki kullanım
Çimento	420 kg/m³; Bağlayıcı faz
Silis dumanı	30 kg/m³; Bağlayıcı yapının iyileştirilmesi
Su	150–160 kg/m³; Karışım ve hidratasyon
Süperakışkanlaştırıcı	4,5–6,0 kg/m³; İşlenebilirliğin düzenlenmesi
İnce barit	850 kg/m³; Yüksek yoğunluklu agrega
İri barit	550 kg/m³; Yüksek yoğunluklu agrega
Hematit/Magnetit	450 kg/m³; Yüksek yoğunluklu agrega
Mikro barit	150 kg/m³; Yüksek yoğunluklu agrega
TPU fiber (8 mm)	4,8 kg/m³; Küp numunelerde polimerik hasır donatı
PLA fiber (6 mm)	2,5 kg/m³; Dikdörtgen numunelerde polimerik hasır donatı
Toplam	2600–2700 kg/m³

V. ARAŞTIRMANIN AMACI VE ÇALIŞMANIN KAPSAMI

Ağır betonların yüksek birim hacim kütlesi sağlayan yapısının, farklı donatı yaklaşımlarıyla birlikte değerlendirilmesi bu çalışmanın temel araştırma çerçevesini oluşturmaktadır. Bu doğrultuda barit ve hematit gibi yüksek yoğunluklu agregalar kullanılarak hazırlanan beton matrisi içerisine üç boyutlu baskı yöntemiyle üretilen polimerik hasır yapı dahil edilmiştir [1], [3], [4], [12]–[16]. Polimerik hasır kullanılan numunelerin birim hacim kütlesi, mekanik davranışı ve kırılma karakteri deneysel olarak incelenmiştir.

Araştırmanın deneysel kapsamı; beton bileşenlerinin belirlenmesi, ağır beton karışımının hazırlanması, polimerik hasırların üç boyutlu yazıcı ile üretilmesi, hasırların numune kalıpları içerisindeki uygun konumlara yerleştirilmesi ve beton numunelerinin hazırlanarak kür edilmesi aşamalarından oluşmaktadır. Hazırlanan numuneler üzerinde birim hacim kütlesi ve mekanik deneyler gerçekleştirilmiş; elde edilen veriler yükleme davranışı ve kırılma sonrasındaki fiziksel durumlarla birlikte değerlendirilmiştir.

Polimerik hasırların değerlendirilmesinde yalnızca deney sonunda ulaşılan dayanım değerleri dikkate alınmamış; yükleme sürecindeki davranış, çatlakların oluşumu ve gelişimi ile kırılma biçimi de ele alınmıştır. Böylece üç boyutlu baskı yöntemiyle oluşturulan hasır geometrisinin beton matrisi içerisindeki davranışı, özellikle çatlak oluşumu ve kırılma süreciyle olan ilişkisi açısından değerlendirilmiştir [3], [12]–[16].

Araştırmanın kapsamını; kullanılan beton bileşenleri, polimer türü, hasır geometrisi, numune düzeni ve deney koşulları doğrultusunda hazırlanan numunelerin deneysel olarak incelenmesi oluşturmaktadır. Elde edilen sonuçlar, numunelerin mekanik davranışları, çatlak gelişimleri ve kırılma karakterleri üzerinden değerlendirilmiştir.

VI. SONUÇ

Deneysel çalışma kapsamında yüksek yoğunluklu agregalar kullanılarak hazırlanan ağır beton içerisine üç boyutlu baskı yöntemiyle üretilen polimerik hasır yapı dahil edilmiştir. Beton karışımının hazırlanmasından polimerik hasırların üretimine, numunelerin oluşturulmasından kür sürecine ve mekanik deneylere kadar ilerleyen uygulama aşamaları bir bütün halinde değerlendirilmiştir.

Deneylerden elde edilen sonuçların değerlendirilmesinde numunelerin nihai dayanım değerlerinin yanı sıra yükleme sırasında meydana gelen davranışlar ve deney sonrasında ortaya çıkan kırılma biçimleri de incelenmiştir. Çatlakların oluşumu ve ilerlemesi ile numunelerin kırılma sonrasındaki fiziksel görünümleri, polimerik hasırların beton içerisindeki davranışının değerlendirilmesinde dikkate alınmıştır. Deneylerden elde edilen grafiksel veriler de bu gözlemlerle birlikte değerlendirilmiştir.

Ağır beton ile polimerik hasırın aynı kompozit sistem içerisinde değerlendirilmesi, yüksek yoğunluklu betonların mekanik ve kırılma davranışının alternatif bir donatı yaklaşımıyla birlikte incelenmesine olanak sağlamıştır. Elde edilen bulgular kullanılan malzemeler, numune geometrisi, polimerik hasır düzeni ve deney koşulları kapsamında değerlendirilmelidir.

İlerleyen çalışmalarda farklı hasır geometrilerinin ve yerleşim düzenlerinin kullanılması, polimer türlerinin çeşitlendirilmesi ve deney koşullarının genişletilmesiyle beton-polimer etkileşiminin daha ayrıntılı biçimde

ISBN: 978-605-80193-6-2

incelenmesi mümkündür. Bunun yanında basınç deneylerine ek olarak eğilme, yarma çekme, enerji yutma kapasitesi ve kırılma enerjisi gibi parametrelerin değerlendirilmesi, söz konusu kompozit sistemin mekanik davranışının daha kapsamlı biçimde açıklanmasına katkı sağlayabilir.

Sonuç olarak, üç boyutlu baskılı polimerik hasırların ağır beton içerisinde kullanılmasına yönelik gerçekleştirilen deneysel çalışma, yüksek yoğunluklu betonların mekanik ve kırılma davranışının alternatif bir donatı yaklaşımıyla birlikte incelenebileceğini ortaya koymuştur. Üç boyutlu baskı teknolojisinin sağladığı geometrik tasarım ve üretim esnekliği, beton-polimer kompozitlerinin farklı düzenlemelerle araştırılmasına olanak sağlamaktadır.

Kaynaklar

[1] Khalaf, M. A., Cheah, C. B., Ramli, M. (2019). The constituents, properties and application of heavyweight concrete: A review. Construction and Building Materials, 215, 73–89. https://doi.org/10.1016/j.conbuildmat.2019.04.146

[2] Palou, M. T., Podhorská, J., Ju, M., Park, K., Čepčianska, J., et al. (2024). Formulation of mixture proportions and experimental study of heavyweight self-compacting concrete based on magnetite and barite. Journal of Thermal Analysis and Calorimetry, 149, 10545–10558. https://doi.org/10.1007/s10973-024-13418-8

[3] Salazar, B., Aghdasi, P., Williams, I. D., Ostertag, C. P., Taylor, H. K. (2020). Polymer lattice-reinforcement for enhancing ductility of concrete. Materials & Design, 196, 109184. https://doi.org/10.1016/j.matdes.2020.109184

[4] Beloshenko, V., Beygelzimer, Y., Chishko, V., Savchenko, B., Sova, N., et al. (2021). Mechanical Properties of Flexible TPU-Based 3D Printed Lattice Structures: Role of Lattice Cut Direction and Architecture. Polymers, 13(17), 2986. https://doi.org/10.3390/polym13172986

[5] Topçu, İ. B. (2003). Properties of heavyweight concrete produced with barite. Cement and Concrete Research, 33(6), 815–822. https://doi.org/10.1016/S0008-8846(02)01063-3

[6] Akkurt, İ., Basyigit, C., Kilincarslan, S., Mavi, B., Akkurt, A. (2006). Radiation shielding of concretes containing different aggregates. Cement and Concrete Composites, 28(2), 153–157. https://doi.org/10.1016/j.cemconcomp.2005.09.006

[7] Demir, F., Budak, G., Sahin, R., Karabulut, A., Oltulu, M., et al. (2011). Determination of radiation attenuation coefficients of heavyweight- and normal-weight concretes containing colemanite and barite for 0.663 MeV γ-rays. Annals of Nuclear Energy, 38(6), 1274–1278. https://doi.org/10.1016/j.anucene.2011.02.009

[8] Gencel, O., Bozkurt, A., Kam, E., Korkut, T. (2011). Determination and calculation of gamma and neutron shielding characteristics of concretes containing different hematite proportions. Annals of Nuclear Energy, 38(12), 2719–2723. https://doi.org/10.1016/j.anucene.2011.08.010

[9] Akkurt, İ., Basyigit, C., Kilincarslan, S., Mavi, B. (2005). The shielding of γ-rays by concretes produced with barite. Progress in Nuclear Energy, 46(1), 1–11. https://doi.org/10.1016/j.pnucene.2004.09.015

[10] Mavi, B., Akkurt, İ., Akyildirim, H., Kılınçarslan, Ş., Basyigit, C. (2010). Gamma-ray shielding properties of concrete including barite at different energies. Progress in Nuclear Energy, 52(7), 620–623. https://doi.org/10.1016/j.pnucene.2010.04.006

[11] Awadeen, M., Amin, M., Bakr, R. H., Tahwia, A. M. (2024). Mechanical properties, attenuation coefficient, and microstructure of ultra high-performance heavyweight concrete for radiation shielding applications. Journal of Building Engineering, 82, 108395. https://doi.org/10.1016/j.jobe.2023.108395

[12] Xu, Y., Zhang, H., Gan, Y., Šavija, B. (2021). Cementitious composites reinforced with 3D printed functionally graded polymeric lattice structures: Experiments and modelling. Additive Manufacturing, 39, 101887. https://doi.org/10.1016/j.addma.2021.101887

[13] Liu, J., Kanwal, H., Tang, C., Hao, W. (2022). Study on flexural properties of 3D printed lattice-reinforced concrete structures using acoustic emission and digital image correlation. Construction and Building Materials, 333, 127418. https://doi.org/10.1016/j.conbuildmat.2022.127418

[14] Tang, C., Liu, J., Hao, W., Wei, Y. (2023). Flexural properties of 3D printed graded lattice reinforced cementitious composites using digital image correlation. Materials & Design, 227, 111734. https://doi.org/10.1016/j.matdes.2023.111734

[15] Chen, M., Yao, X., Zhu, L., Yin, M., Xiong, Y., Hu, N. (2023). Geometric design and performance of single and dual-printed lattice-reinforced cementitious composite. Cement and Concrete Composites, 143, 105266. https://doi.org/10.1016/j.cemconcomp.2023.105266

[16] Bol, R. J. M., Xu, Y., Luković, M., Šavija, B. (2025). Does printing direction influence the bond between 3D printed polymeric reinforcement and cementitious matrix? Engineering Failure Analysis, 174, 109471. https://doi.org/10.1016/j.engfailanal.2025.109471

[17] Onaizi, A. M., Amran, M., Tang, W., et al. (2024). Radiation-shielding concrete: A review of materials, performance, and the impact of radiation on concrete properties. Journal of Building Engineering, 97, 110800. https://doi.org/10.1016/j.jobe.2024.110800

[18] Cheah, C. B., Khalaf, M. A., Ramli, M., et al. (2021). Modern heavyweight concrete shielding: Principles, industrial applications and future challenges; review. Journal of Building Engineering, 39, 102290. https://doi.org/10.1016/j.jobe.2021.102290

ISBN: 978-605-80193-6-2

A Comparative Study of the Number of Stories, Structural Rigidity, and LRB–FPS Hybrid Seismic Isolation Performance in Multi-Story Reinforced Concrete Frame Buildings

Rifat SEZER[1], Muhammet YANGÖZ[2]

[1]Karamanoğlu Mehmetbey University, Faculty of Engineering, Department of Civil Engineering, Karaman, Turkey

[2]Karaman Provincial Special Administration, Directorate of Urban Planning and Urban Renewal, Karaman, Turkey

* Corresponding author e-mail: rsezer@kmu.edu.tr

***Abstract*—In this study, the effects of the number of stories and the seismic isolation system on the dynamic behavior of 6-, 12-, 18-, and 24-story reinforced concrete hospital buildings were examined comparatively. The structures were modeled in the SAP2000 environment for fixed-support (FS), lead-rubber-core (LRB), friction pendulum system (FPS), and hybrid LRB+FPS configurations. All models feature five 5-meter spans in both the X and Y directions, a story height of 3 meters, C30 concrete, and B420Ç reinforcing steel. The first-mode period, base shear force, top–first-story displacement difference, and performance variation as a function of the number of stories were used as the primary comparison criteria. The isolation systems extended the natural period at all floor levels and generally reduced the base shear force. In the 24-story model, the hybrid LRB+FPS system yielded the lowest base shear force among the isolated systems studied, at Vt = 16,994 kN. The results indicate that a single isolator type does not constitute a universal optimum for all building heights; rather, the optimal selection must be based on a combined evaluation of period, base shear, relative displacement, isolation displacement, and the effects of the top mode.**

***Keywords*— Seismic isolation, LRB, FPS, hybrid isolation, multi-story reinforced concrete building, SAP2000, base shear force.**

I. INTRODUCTION

Seismic isolation is a passive control approach that aims to extend the fundamental vibration period by increasing the structure's horizontal flexibility and to dissipate a significant portion of seismic energy at the isolation level. While LRB systems combine the re-centering and horizontal flexibility properties of elastomeric layers with the hysteretic energy dissipation of a lead core, FPS systems operate through friction and geometric re-centering mechanisms on a curved sliding surface. Since the mechanical behaviors of these two systems differ, their relative advantages may vary as building height, stiffness distribution, soil conditions, and seismic hazard levels change [1–4].

Isolation design in tall and high-rise buildings is more complex than in low-rise structures. Increasing contributions from higher modes, overturning moments, P–Δ effects, isolation displacement, and wind service conditions can narrow the design range. Recent reviews indicate that in tall buildings, not only base isolation but also inter-story isolation and additional damping strategies must be evaluated on a performance-based basis [4,5]. Studies from 2026 also demonstrate that superstructure characteristics can significantly alter the seismic response of reinforced concrete structures isolated with LRBs [6].

In experimental comparisons of LRB and FPS, energy dissipation, equivalent damping, and effective stiffness are among the key performance indicators. Son et al. [2] reported that FPS specimens demonstrated a distinct advantage over LRB specimens in energy dissipation per cycle, while equivalent damping ratios could be comparable. Özer and İnel [3], on the other hand, demonstrated that combination systems using LRB, FPS, and fluid viscous dampers—either individually or in combination—may be more effective in terms of certain performance criteria. These findings support further research into hybrid systems.

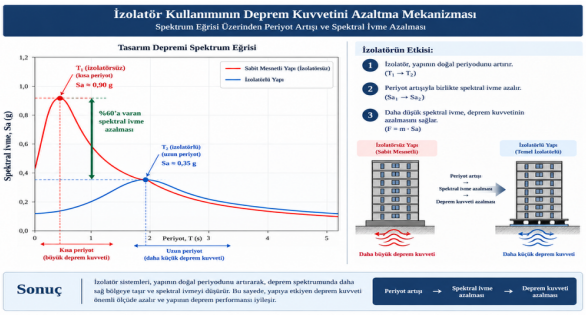


Fig. 1. Period extension and earthquake force reduction mechanism of seismic isolation

The originality of this study lies in the comparison of 6, 12, 18, and 24 story reinforced concrete structures under fixed-base, LRB, FPS, and hybrid LRB+FPS configurations, while maintaining the same plan geometry and material classes. Thus, the effect of building height on performance variations depending on the isolator type can be directly observed.

A. Research Objective and Scope

The objectives are to determine: (i) the change in the fundamental period as the number of stories increases; (ii) the effect of LRB, FPS, and hybrid LRB+FPS systems on the base shear force; (iii) the change in superstructure deformation; and (iv) the relative advantages of single and hybrid isolation systems in high-rise buildings. The study results were discussed in light of the current 2024–2026 literature, and design recommendations were developed for more advanced hybrid configurations, such as LRB+FPS+VD, in the future.

ISBN: 978-605-80193-6-2

II. NUMERICAL MODEL AND METHOD

The models are assumed to be located in the Istanbul/Bayrampaşa region. The plan dimensions consist of five 5-meter spans in both the X and Y directions, with a story height of 3 meters and a slab thickness of 15 cm. The concrete class is C30, and the reinforcing steel is B420Ç. Column and wall dimensions were increased in parallel with the increase in the number of stories ( ). Fixed-base and seismically isolated models were created in the SAP2000 environment, and the modal characteristics were compared with the base shear and displacement results. When performing the analysis in SAP2000, the **Mode Combination Method** was used as the horizontal load method. The number of modes was set to 12, and the mass participation ratios reached 99%.

Rubber Isolator Properties: Link/Support Type: Rubber Isolator. **For 6-story buildings:** U1 Linear Effective Stiffness: 11258027 kN/m, Effective Damping: 0.1, U2 and U3 Linear Effective Stiffness: 7860. 16 kN/m, Effective Damping: 0.1, U2 and U3 Nonlinear Stiffness: 6,625.51 kN/m, U2 and U3 Yield Strength: 345.7 kN, U2 and U3 Post-Yield Stiffness Ratio: 0.1.

For 12-story buildings: U1 Linear Effective Stiffness: 7,456,038 kN/m, Effective Damping: 0.1, U2 and U3 Linear Effective Stiffness: 4,737 kN/m, Effective Damping: 0.1, U2 and U3 Nonlinear Stiffness: 3,993 kN/m, U2 and U3 Yield Strength: 432 kN, U2 and U3 Post-Yield Stiffness Ratio: 0.1.

For 18–24-story buildings: U1 Linear Effective Stiffness: 6,807,262 kN/m, Effective Damping: 0.1, U2 and U3 Linear Effective Stiffness: 4,167. 36 kN/m, Effective Damping: 0.1, U2 and U3 Nonlinear Stiffness: 3,504.4 kN/m, U2 and U3 Yield Strength: 510.59 kN, U2 and U3 Post-Yield Stiffness Ratio: 0.1.

Friction Isolator Properties: Link/Support Type: Friction Isolator. **For 6-story buildings:** U1 Linear Effective Stiffness: 3.142×10^{8} kN/m, Effective Damping: 0, U1 Nonlinear Effective Stiffness: 4394.6 kN/m, Effective Damping: 0, U2 and U3 Linear Effective Stiffness: 5178. 64 kN/m, Effective Damping: 0, U2 and U3 Nonlinear Stiffness: 4,394.6 kN/m, Friction Coefficient, Slow: 0.03, Friction Coefficient, Fast: 0.05, Rate Parameter: 50, Net Pendulum Radius: 1.

For 12-story buildings: U1 Linear Effective Stiffness: 1.414×10^{8} kN/m, Effective Damping: 0, U1 Nonlinear Effective Stiffness: 4750.07 kN/m, Effective Damping: 0, U2 and U3 Linear Effective Stiffness: 5663. 23 kN/m, Effective Damping: 0, U2 and U3 Nonlinear Stiffness: 4750.07 kN/m, Friction Coefficient, Slow: 0.03, Friction Coefficient, Fast: 0.05, Rate Parameter: 50, Net Pendulum Radius: 2.23.

For 18–24-story buildings: U1 Linear Effective Stiffness: 2.513×10^{8} kN/m, Effective Damping: 0, U1 Nonlinear Effective Stiffness: 4194.07 kN/m, Effective Damping: 0, U2 and U3 Linear Effective Stiffness: 5000. 63 kN/m, Effective Damping: 0, U2 and U3 Nonlinear Stiffness: 4194.07 kN/m, Friction Coefficient, Slow: 0.01, Friction Coefficient, Fast: 0.03, Rate Parameter: 50, Net Pendulum Radius: 3.

Table 1. Cross-sections of the structural system for the building models

Story	Column (cm)	Beam (cm)
6	40×40	30×45
12	50×50	30×50
18	60×60	30×55
24	70×70	30×60

A. Isolation Configurations

Four configurations were evaluated for each building height: fixed-supported (FS), LRB, FPS, and hybrid LRB+FPS. In the LRB configuration, horizontal stiffness and hysteretic energy dissipation are critical; in the FPS configuration, the radius of curvature, friction, and recentring are critical; and in the hybrid configuration, the combined behavior of the two mechanisms is critical. Recent research indicates that these parameters must be considered in conjunction with seismic hazard and soil class [1,7].

B. Performance Criteria

For comparison, the fundamental period T1, base shear force Vt, displacement difference Δmax between the top story and the first story, and the variation of these quantities with story number were used. The reduction in base shear was calculated using the equation $\eta V = (V_{FS} - V_{iso}) / V_{FS} \times 100$. The displacement difference was taken as the difference between the top-floor and first-floor displacements from the SAP2000 results.

III. RESULTS FOR A 6-STORY REINFORCED CONCRETE BUILDING

In the 6-story FS structure, the first-mode period is T1 = 0.73224 s and the base shear force is Vt = 14,531 kN. With isolation, the period increased to the range of 1.245–1.330 s; the base shear forces in the LRB, FPS, and hybrid systems were 9,056.45, 9,056.45, and 9,054.66 kN, respectively. Thus, all three isolation systems achieved a reduction in base shear of approximately 37.7%. The difference in displacement between the top floor and the first floor is 1.621 cm for FS, 0.510 cm for LRB, 0.408 cm for FPS, and 0.489 cm for the hybrid system. Consequently, this indicates that in 6-story models with isolators, relative displacements are reduced by 69%, 75%, and 70%, respectively, and that the superstructure undergoes rigid-body motion.

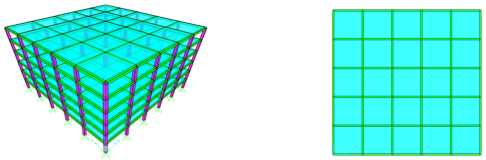


Fig. 2. 6-story SAP2000 building model and typical floor plan

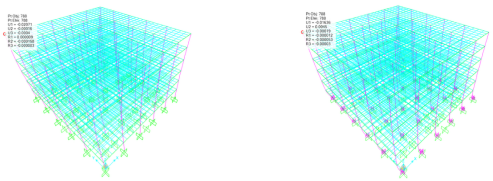


Fig. 3. First mode shape of the 6-story FS and LRB model

ISBN: 978-605-80193-6-2

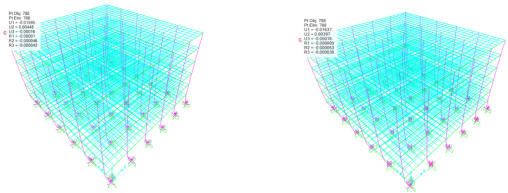


Fig. 4. First mode shape of the 6-story FPS and Hybrid model

Table 2. Comparison of results for the 6-story models

System	T1 (s)	Vt (kN)	Δmax (m)	Vt reduction (%)
FS	0.732	14,531	0.01621	—
LRB	1.245	9,056	0.00510	37.68
FPS	1.330	9,056	0.00408	37.68
Hybrid	1.251	9055	0.00489	37.69

In the low-rise model, the FPS system yielded the longest period and the smallest superstructure displacement difference; however, in terms of base shear, the FPS and hybrid systems are nearly equivalent. This result indicates that, for low-rise buildings, the difference between the systems may be limited to a single criterion, and that the optimal selection should be made in conjunction with design criteria such as device displacement, cost, and maintenance.

IV. RESULTS FOR THE 12-STORY REINFORCED CONCRETE BUILDING

In the 12-story embedded model, T1 = 1.19147 s and Vt = 20,507.62 kN were obtained. For the LRB, FPS, and hybrid systems, T1 was approximately 2.223, 2.171, and 2.213 s, respectively; Vt was 11,434.96, 11,428.98, and 11,433.89 kN, respectively. The three systems reduced the base shear force by approximately 44.2%. The superstructure displacement difference was 1.308 cm in the built-in model, while it decreased to a range of 0.550–0.592 cm in the isolated models. Consequently, this indicates that relative displacements decreased by 55–58% in the 12-story models with isolators and that the superstructure moved as a rigid mass.

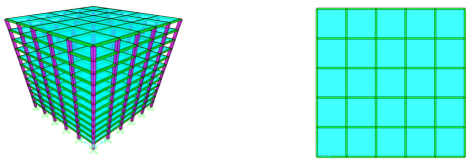


Fig. 5. 12-story SAP2000 building model and typical floor plan

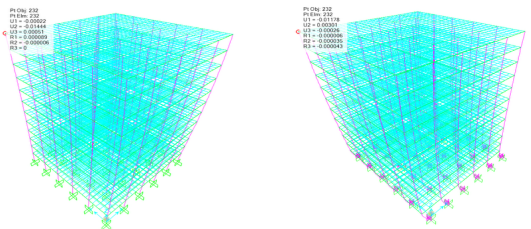


Fig. 6. First mode shape of the 12-story FS and LRB model

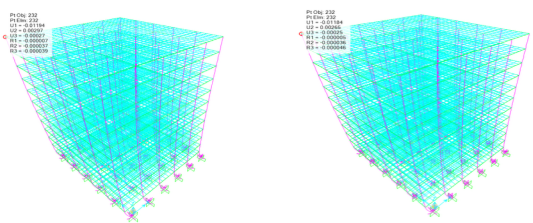


Fig. 7. First mode shape of the 12-story FPS and Hybrid model

Table 3. Comparison of results for the 12-story models

System	T1 (s)	Vt (kN)	Δmax (m)	Vt reduction (%)
FS	1.191	20,508	0.01308	—
LRB	2.223	11,435	0.00550	44.24
FPS	2.171	11,429	0.00592	44.27
Hybrid	2.213	11,434	0.00555	44.24

The 12-story model represents the building height at which the three isolation systems come closest to each other in terms of base shear. The LRB system yields the highest T1 value, while the LRB and hybrid systems result in smaller differences in superstructure displacement. This convergence clearly demonstrates that selecting a system based solely on global base shear is insufficient.

V. RESULTS FOR THE 18-STORY REINFORCED CONCRETE BUILDING

In the 18-story embedded model, T1 = 1.53521 s and Vt = 25,308.76 kN. For the LRB, FPS, and hybrid systems, T1 was found to be 2.311, 2.334, and 2.291 s, respectively; the base shear forces were 16,899.45, 17,909.57, and 17,045.21 kN, respectively. The reduction in base shear is 33.23% for the LRB, 29.23% for the FPS, and 32.65% for the hybrid system. The superstructure displacement difference is 1.124 cm in the cast-in-place model, and approximately 0.241, 0.265, and 0.357 cm in the LRB, FPS, and hybrid systems, respectively. Consequently, this indicates that in 18-story models with isolators, relative displacements are reduced by 79%, 76%, and 68%, respectively, and that the superstructure undergoes rigid-body motion.

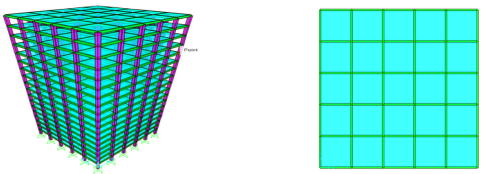


Fig. 8. 18-story SAP2000 building model and typical floor plan

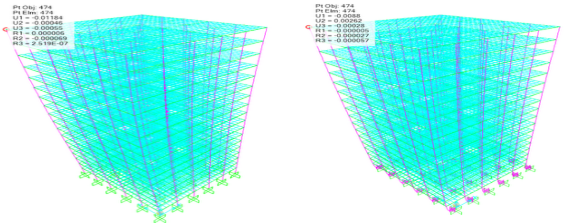


Fig. 9. First mode shapes of the 18-story FS and LRB model

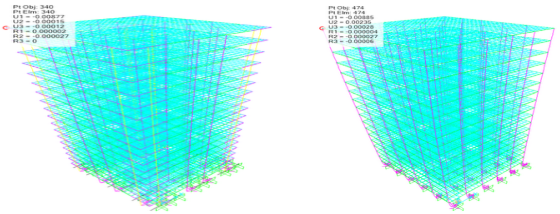


Fig. 10. First-mode shapes of the 18-story FPS and Hybrid model

ISBN: 978-605-80193-6-2

Table 4. Comparison of results for the 18-story models

System	T1 (s)	Vt (kN)	Δmax (m)	Vt reduction (%)
FS	1.535	25,309	0.01124	—
LRB	2.311	16,899	0.00241	33.23
FPS	2.334	17,910	0.00265	29.23
Hybrid	2.291	17,045	0.00357	32.65

At this height, the LRB provides the lowest values in terms of base shear and superstructure displacement difference, while the FPS provides the longest period. Therefore, it is observed that as building height increases, an increase in period alone does not result in lower base shear. This behavior can be explained by the contributions of the upper mode, effective damping, isolator stiffness, and superstructure–isolation interaction.

VI. RESULTS FOR A 24-STORY REINFORCED CONCRETE BUILDING

In the 24-story embedded model, T1 = 1.88891 s and Vt = 29,025.20 kN. For the LRB, FPS, and hybrid systems, T1 is 2.74191, 2.61875, and 2.71781 s, respectively; Vt is 20,085.32, 21,025.21, and 16,994 kN, respectively. Compared to the fixed-base model, the hybrid system reduced the base shear force by approximately 41.45% and yielded the lowest Vt value for this building height. The superstructure displacement difference is 0.997 cm in the cast-in-place model; in the LRB, FPS, and hybrid systems, it is 0.361, 0.405, and 0.366 cm, respectively. Consequently, this indicates that in 24-story models with isolators, relative displacements are reduced by 64%, 59%, and 63%, respectively, and that the superstructure undergoes rigid-body motion.

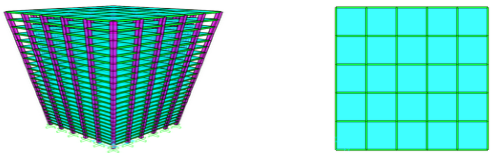


Fig. 11. 24-story SAP2000 building model and typical floor plan

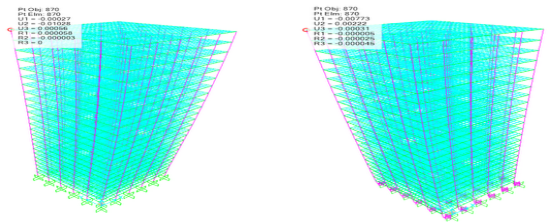


Fig. 12. First mode shapes of the 24-story FS and LRB model

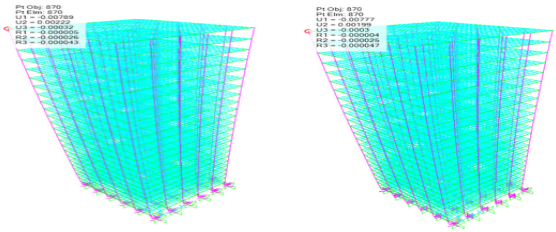


Fig. 13. First-mode shapes of the 24-story FPS and Hybrid model

Table 5. Comparison of results for the 24-story models

System	T1 (s)	Vt (kN)	Δmax (m)	Vt reduction (%)
FS	1.889	29,025	0.00997	—
LRB	2.742	20,085	0.00361	30.80
FPS	2.619	21,025	0.00405	27.56
Hybrid	2.718	16,994	0.00366	41.45

In the 24-story model, the hybrid system's significant base shear advantage is noteworthy. In contrast, the LRB yields a slightly lower value for the superstructure displacement difference. Therefore, the definition of the "best" system for tall buildings should not be limited to base shear; story displacement, relative story drift, story accelerations, overturning moment, and service performance must be evaluated together [4–6].

VII. COMPARATIVE FINDINGS

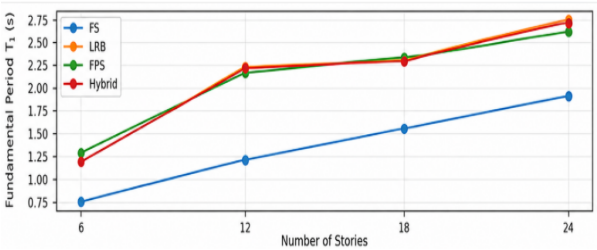


Fig. 14. Variation of the first-mode period as a function of the multiplier and system type.

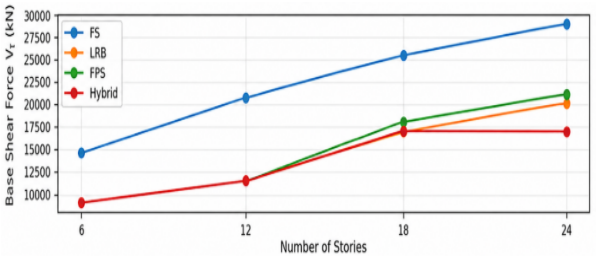


Fig. 15. Variation of the base shear force with the story number and system type.

Figure 20 shows that isolation extends the fundamental period at all story levels. However, it is evident—particularly in the 18- and 24-story models—that the increase in period is not directly proportional to the reduction in base shear. In Figure 21, while the isolation systems in the 12-story building are nearly identical, the hybrid system in the 24-story building stands out distinctly.

Table 6. Summary comparison of base shear forces (kN)

Story	FS	LRB	FPS	Hybrid
6	14,531	9,056	9,056	9,055
12	20,508	11,435	11,429	11,434
18	25,309	16,899	17,910	17,045
24	29,025	20,085	21,025	16,994

VIII. DISCUSSION

A. Effect of Number of Stories on Insulation Performance

The results show that the number of stories affects isolation performance in a nonlinear manner. While the three isolation systems are very close to each other in terms of base shear for the 6-story building, the LRB system is more advantageous for the 18-story building, and the hybrid LRB+FPS system is clearly more advantageous for the 24-story building. This change can

ISBN: 978-605-80193-6-2

be attributed to the superstructure period approaching the isolation period, the activation of higher modes, and changes in the distribution of lateral stiffness. The 2025 review on high-rise building isolation emphasizes that isolation design for high-rise buildings must be approached differently from that of low-rise buildings due to higher modes and overturning effects [4].

B. Comparison of LRB and FPS Behavior

LRB combines energy dissipation based on elastomeric flexibility with the plasticization of the lead core. FPS, on the other hand, largely determines its period through the curvature radius due to its pendulum geometry and dissipates energy through friction. Son et al. [2] demonstrated in an experimental study that the energy dissipation per cycle of the FPS can be higher than that of the LRB. However, in current building analyses, the FPS does not yield the lowest base shear at every floor level. This indicates that energy dissipation at the device level must be evaluated in conjunction with the global response at the building level.

C. The Role of Hybrid Systems

Hybrid systems aim to combine the advantages of different isolation mechanisms within the same structure. The fact that the hybrid system in the 24-story model reached a value of Vt = 16,994 kN demonstrates the potential of this approach in tall buildings. Özer and İnel [3] demonstrated that the combined use of LRB/FPS and " " fluid viscous dampers can be more effective than individual systems in meeting certain seismic performance criteria. A 2026 study involving the combined use of a triple-friction pendulum and a nonlinear viscous damper in tall adjacent structures also supports the effectiveness of the combined control approach under multi-level seismic effects [8].

D. Comparison with Current Literature

Kahvecioğlu et al. [1] demonstrated that the performance of LRBs and FPS varies depending on seismic hazard and soil class; therefore, no single system can be considered superior under all conditions. Kayı et al. [6] revealed that superstructure characteristics significantly influence the isolation response in a reinforced concrete structure retrofitted with LRBs. A 2026 study on segmented friction pendulums, meanwhile, highlights the importance of multi-objective optimization that addresses both seismic and wind demands in tall buildings [9]. These studies support that the current findings are consistent with a performance-based and multi-criteria design approach.

IX. ENGINEERING EVALUATION AND OPTIMAL SYSTEM SELECTION

The selection of a seismic isolation system should not be based solely on the smallest base shear force. In practice, at least five performance criteria must be considered together: (i) base shear force, (ii) relative displacement of the superstructure, (iii) floor accelerations and non-structural element requirements, (iv) isolation story displacement and residual displacement, (v) safety of devices under vertical loads, overturning, and service conditions. In structures with a high importance factor, such as hospitals, maintaining functionality after an earthquake is also a critical design objective.

Table 7. Performance-based system evaluation by number of stories.

Story	Standout in terms of Vt	Notable in terms of Δmax	Engineering commentary
6	Hybrid ≈ FPS ≈ LRB	FPS	The systems are comparable; cost and device displacement are the deciding factors
12	FPS ≈ Hybrid ≈ LRB	LRB	Global responses are very close; a multi-criteria selection is required
18	LRB	LRB	LRB delivers balanced performance in this model
24	Hybrid	LRB	In the Hybrid Vt, the LRB offers an advantage in superstructure deformation

A. Proposal for Multi-Criteria Optimization

In future studies, the optimal design can be defined using a weighted objective function: $J=w_1(V_t/V_{ref})+w_2(\Delta_{drift}/\Delta_{ref})+w_3(aiso/aref)+w_4(D_{iso}/D_{ref})+w_5(C/C_{ref})$. Here, the weights can be determined based on the building's intended use. In hospital buildings, it is appropriate to assign higher weights to floor acceleration and relative drift; in tall structures, it is appropriate to additionally define isolation displacement and overturning safety as boundary conditions.

B. Integration of Viscous Dampers

Although the title and scope definition of the source study mention a hybrid system with viscous dampers, the current document does not contain a verifiable series of LRB+FPS+VD results for all floor levels. Therefore, no quantitative claims regarding damper performance are made in this paper. Instead, based on the current literature, it is recommended that LRB+FVD or FPS+FVD combinations be examined in a future analysis phase [3,8]. Such an extension is particularly important for controlling top-mode and floor acceleration requirements in 18- and 24-story models.

C. Limitations

The results of this study are limited to the existing SAP2000 models and the reported modal/base shear/displacement outputs. Isolator force–displacement cycles, maximum isolation story displacement, story accelerations, device vertical loads, and statistical variation under different earthquake records are not fully reported in the current dataset. Therefore, the results should be evaluated as comparative design trends; all verification steps specified in TBDY 2018 and relevant international standards must be performed separately for the application project [10,11].

X. RESULTS

1) As the number of stories increased, the fundamental period of the fixed-base structures rose from 0.732 s to 1.889 s. LRB, FPS, and hybrid systems further extended this period at all story levels.

ISBN: 978-605-80193-6-2

2) In the 6-story model, the base shear forces of the LRB, FPS, and hybrid systems are very close to each other at approximately 9.06 MN and result in a reduction of approximately 37.7% compared to the rigidly fixed model.

3) In the 12-story model, the three isolation configurations produced a reduction in base shear of approximately 44.2%; the difference between the systems remained negligible.

4) In the 18-story model, the LRB system produced a lower base shear force (Vt = 16.90 MN) than the FPS and hybrid systems. In this model, the LRB system also provided the lowest superstructure displacement difference.

5) In the 24-story model, the hybrid LRB+FPS system produced the lowest base shear force (Vt = 16.994 MN), resulting in a reduction of approximately 41.45% compared to the fixed-base model.

6) When comparing the relative story drift ratios in 6-, 12-, 18-, and 24-story models, it was observed that systems with isolators significantly reduced story drift compared to systems with fixed supports. The average reduction values, expressed as percentages, are 71.33, 56.00, 74.00, and 62.00, respectively. This indicates that in the isolator-equipped models, the superstructure moves as a rigid mass and is not affected by the earthquake.

7) Current literature from 2025–2026 supports that the performance of LRBs and FPS is sensitive to hazard level, soil class, superstructure characteristics, and device parameters; combined isolator–dampener systems can provide additional advantages under certain conditions [1–9].

8) In future studies, the LRB+FPS+VD configuration for the same model family of 6–24-story buildings should be multi-objective optimized using criteria such as multiple earthquake records, floor acceleration, isolation displacement, energy consumption, and cost.

A. Key Design Implications

Rather than defining a single “universal best isolator” for this study, it is more accurate to define an optimal range based on the number of stories and the target performance. While the systems may yield similar results in low- and mid-rise models, the hybrid behavior provided a more pronounced advantage in terms of base shear for the high-rise model. However, the displacements of the devices and the superstructure must be controlled separately.

References

[1] Kahvecioglu, B., Nigdeli, S. M., Bekdaş, G., Kim, S. and Geem, Z. W. (2026). Comparative Assessment of Lead Rubber and Friction Pendulum Seismic Isolation Systems Under Varying Seismic Hazard and Site Conditions. GeoHazards, vol. 7, no. 2, art. 77, 2026. doi:10.3390/geohazards7020077.

[2] Son, J. -S., Chin, W.J., C. B. Cho, C. B. and J. -Y. Lee, J. -Y. (2025). *Comparison of Seismic Resistance Between LRB and FPB in Terms of Energy Dissipation Per Cycle and Equivalent Damping.* Journal of Disaster Science and Management, vol. 1, art. 9, 2025. doi:10.1007/s44367-025-00008-0.

[3] Ozer, E. and Inel, M. (2025). The effect of single and combined use of base isolators and fluid viscous dampers on seismic performance in a conventional reinforced concrete building with torsional irregularity. Journal of Building Engineering, vol. 101, 111898. doi:10.1016/j.jobe.2025.111898.

[4] Avinash, G. C. and Lingeshwaran, N. (2024). *Review on base and inter-story seismic isolation systems for high-rise buildings,* Research on Engineering Structures & Materials, vol. 11, no. 1, pp. 231–271. doi:10.17515/resm2024.156st0117rv.

[5] Kumar, P., Gogineni, A., Sahu, S. K: Animesh Kumar Singh, A. K. & Bamidele Charles Olaiya, B. C. 2025. *A comprehensive review of response control systems for irregular buildings: exploring base isolators and dampers.* Discover Applied Sciences. doi:10.1007/s42452-025-07469-1.

[6] Kayı, D. B., Bayhan, B. and Ozdemir, G. (2026). *Effect of superstructure characteristics on seismic response of an RC building retrofitted with lead rubber bearings.* Bulletin of Earthquake Engineering, vol. 24, pp. 6413–6441. doi:10.1007/s10518-026-02466-0.

[7] Domadzra, Y., Bhandari, M., Hasan, M. 2024. *Seismic response of base-isolated buildings: exploring isolator properties.* Asian Journal of Civil Engineering, vol. 25. doi:10.1007/s42107-024-01041-9.

[8] J. Zhang, M. Wu, Y. Miao, and Z. Zhou, “Seismic Isolation Control of High-Rise Adjacent Structures Based on Triple Friction Pendulum Bearing,” Chinese Quarterly of Mechanics, vol. 47, no. 3, pp. 766–779, 2026. doi:10.15959/j.cnki.0254-0053.2026.03.022.

[9] Li, Z., Huang, G, Jiang, Y., Fan, Y. H., Fu, J. Y., He, Y. C. 2026. *Tall building isolated by segmented friction pendulum bearing: dynamic response analysis and multi-objective optimization.* *Soil Dynamics and Earthquake Engineering*. 110498. doi:10.1016/j.soildyn.2026.110498.

[10] AFAD, *Turkish Building Seismic Code* (TBDY 2018), Ankara, 2018.

[11] ASCE/SEI 7-22. 2022. Minimum Design Loads and Associated Criteria for Buildings and Other Structures, American Society of Civil Engineers, Reston, VA.

[12] *Computers and Structures*, Inc. (CSI), SAP2000 Analysis Reference Manual, Berkeley, CA, 2020.

[13] Y.-P. Zhu, Z. Q. Lang, K. Fujita, and I. Takewaki, “Analysis and design of nonlinear seismic isolation systems for building structures—An overview,” Frontiers in Built Environment, vol. 8, 1084081, 2023. doi:10.3389/fbuil.2022.1084081.

[14] D. Forcellini and K. N. Kalfas, “Inter-story seismic isolation for high-rise buildings,” Engineering Structures, vol. 275, 115175, 2023. doi:10.1016/j.engstruct.2022.115175.

 ISBN: 978-605-80193-6-2

[15] J. Li et al., "Effect of base isolation devices and buckling-restrained brace dampers, used individually and in combination, on the seismic performance of conventional reinforced concrete buildings," Structures, vol. 81, 110323, 2025. doi:10.1016/j.istruc.2025.110323.

[16] Xu, Y., Guo, T., Zhu, R., Liu, Z., Yang, J., Jiang, H. 2025. *"Shaking table tests and numerical studies of a high-rise TFPB-isolated steel frame,"* Journal of Constructional Steel Research, vol. 237B, 110147. doi:10.1016/j.jcsr.2025.110147.

 ISBN: 978-605-80193-6-2

IRQMEPAVE: A Knowledge-Based Intelligent Mechanistic–Empirical Tool for Flexible Pavement Design in Iraq and Its Validation Against AASHTO 1993 Method and MnPAVE

Saadoon Obaid Eyada[1] and Adil N. Abed[1]

[1]Department Civil Engineering Department, Al Mansour University College, 10068, Baghdad, Iraq.

* Corresponding author e-mail: saadoon.o.eyada@muc.edu.iq

***Abstract*—Flexible pavement design in Iraq still relies largely on the empirical AASHTO 1993 procedure, which cannot represent the extreme pavement temperatures and heavy axle loads that govern distress in the country. This paper presents IRQMEPAVE, an offline, single-file mechanistic–empirical (M–E) design tool developed for Iraqi conditions, and compares it with MnPAVE-Flexible 6.405 and AASHTO 1993 for three typical sections (Baghdad, Anbar and Basra; 5, 20 and 40 million ESALs). The two tools share the same computational chain — seasonal climate, temperature-dependent asphalt modulus, multilayer linear-elastic response, Asphalt Institute transfer functions and Miner damage accumulation — so the comparison isolates the effect of the calibration constants and the effective-temperature definition. They ranked the sections identically and their damage ratios were correlated (R^2 = 0.76 between log-damage values), but IRQMEPAVE, which runs uncalibrated global coefficients, predicted 1.7–5.6 times less damage than MnPAVE, which runs Minnesota-calibrated coefficients. Substituting monthly maximum for monthly mean air temperature raises the IRQMEPAVE asphalt temperature from 48 to 59 °C, reproducing the MnPAVE pavement temperature, and accounts for most of the remaining difference. Provisional harmonisation factors of 0.21 for fatigue and 0.35 for rutting, and a staged local-calibration plan, are proposed. IRQMEPAVE additionally models asphalt-layer rutting, which governed every Iraqi case (up to 22.6 mm), reports seven reliability levels simultaneously and includes a knowledge-based Find Ideal Section engine, while AASHTO 1993 rated all three sections adequate.**

***Keywords*— Mechanistic–empirical design, flexible pavement, knowledge-based system, dynamic modulus, MnPAVE, asphalt rutting, Iraq**

I. INTRODUCTION

Road networks carry almost all freight between Iraq's ports, border crossings and production centres, yet most Iraqi pavements are still designed with the AASHTO 1993 guide [1] and the SCRB/R9 specification [11], i.e. with regression equations derived from the AASHO Road Test climate and traffic. Summer air temperatures near 50 °C and heavy trucks produce premature rutting and fatigue cracking that such empirical methods cannot anticipate.

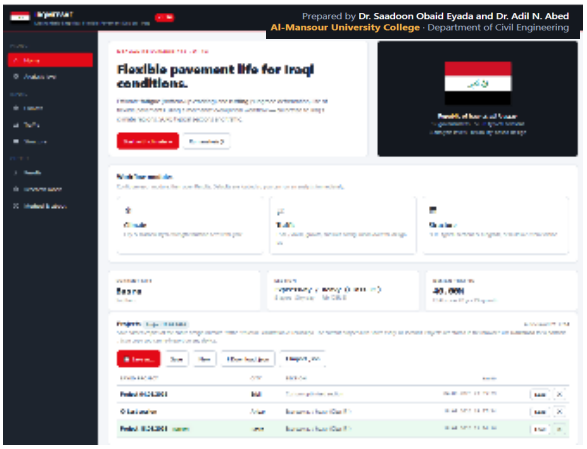
Mechanistic–empirical (M–E) methods compute pavement responses from layered elastic theory and convert them into distress through transfer functions [2, 3]. AASHTOWare Pavement ME implements NCHRP 1-37A but requires a commercial licence and extensive databases [3]; MnPAVE-Flexible, developed by the Minnesota DOT, uses a seasonal layered-elastic approach with reliability applied to layer inputs [6, 7, 14], but is built around Minnesota's freeze–thaw climate and has no asphalt-layer rutting model. Neither is readily applicable to Iraq.

This study therefore (i) presents IRQMEPAVE, an M–E design tool localised for Iraq that combines established mechanistic models with a knowledge-based design engine, and (ii) compares it with MnPAVE-Flexible and AASHTO 1993 for three representative Iraqi sections, so that the sources of agreement and divergence between the tools are identified before local calibration.

II. IRQMEPAVE FRAMEWORK

A. Architecture

IRQMEPAVE is a single self-contained HTML/JavaScript file that runs in any browser without installation or internet connection (Figure 1). It comprises input screens (climate, traffic, structure), an analysis core, a research mode exposing monthly responses, and a results module reporting critical strains, allowable repetitions, damage ratios, rut depth and predicted life at seven reliability levels (Figure 2).



ISBN: 978-605-80193-6-2



Fig. 2. IRQMEPAVE screens: (left) results for CS3 (Basra) with damage check and predicted rut depth; (right) life at seven reliability levels for CS1 (Baghdad)

B. Climate and Asphalt Stiffness

Monthly mean air temperatures for the Iraqi governorates, compiled from official statistical records [16], drive a 12-month analysis; each month is analysed separately rather than being grouped into a small number of seasons. Air temperature is converted to an effective asphalt temperature with the Witczak relation used in MS-1 [4]:

$$MMPT = MMAT\,(1 + 1/(z + 4)) - 34/(z + 4) + 6 \quad (1)$$

where MMPT and MMAT are mean pavement and air temperatures (°F) and z is depth (in.). IRQMEPAVE evaluates Eq. (1) at z = 2 in. (50 mm) using the monthly mean air temperature; back-calculation from its output reproduces the tool's asphalt temperatures to within 0.05 °C. The asphalt dynamic modulus |E*| is obtained from the Witczak 1-37A model [5]:

$$\log|E^*| = -1.249937 + 0.029232p_{200} - 0.001767p_{200}^2 - 0.002841p_4 - 0.058097V_a - 0.802208\cdot V_{beff}/(V_{beff} + V_a) + [3.871977 - 0.0021p_4 + 0.003958p_{38} - 0.000017p_{38}^2 + 0.005470p_{34}] / [1 + \exp(-0.603313 - 0.313351\log f - 0.393532\log \eta)] \quad (2)$$

with |E*| in 10^5 psi, gradation terms p_i (%), air voids V_a and effective binder content V_{beff} (%), loading frequency f (Hz) and binder viscosity η (10^6 P) from log log η = A + VTS·log T_R, with A and VTS selected from the Iraqi penetration grade.

C. Structural Response and Distress Models

Responses under a standard 80 kN single axle with dual tyres are computed with a multilayer linear-elastic (Burmister) engine [8, 9]; an Odemark transformation [10] is available for fast screening. Fatigue and subgrade rutting use the Asphalt Institute transfer functions [4], and asphalt-layer rutting follows the NCHRP 1-37A permanent-deformation model integrated over sub-layers with the depth-confinement factor k_z [2]:

$$N_f = 0.0796 \cdot C \cdot \varepsilon_t^{-3.291} \cdot |E^*|^{-0.854},\ C = 10^M,\ M = 4.84\,[V_b/(V_a + V_b) - 0.69] \quad (3)$$

$$N_d = 1.365 \times 10^{-9} \cdot \varepsilon_v^{-4.477} \quad (4)$$

$$\varepsilon_p/\varepsilon_r = k_z\,\beta_{r1}\,10^{-3.4488}\,T^{1.5606\beta\sim r2\sim}\,N^{0.479244\beta\sim r3\sim} \quad (5)$$

Monthly damage is accumulated with Miner's law, $D = \Sigma\, n_i/N_i$ [13]. Rutting life is the smaller of the subgrade and asphalt-layer values, with a terminal rut depth of 12.5 mm. Reliability is applied by shifting the median life by $10^{-Z\sim R\sim\cdot S\sim 0\sim}$ with S_0 = 0.461, so the structure and moduli are unchanged. Global coefficients are used by default; a calibrated Iraqi set can be substituted [12].

D. Knowledge-Based Intelligent Design Engine

The Find Ideal Section module is a rule-based expert system combined with an optimisation search. Its knowledge base encodes SCRB/R9 section types, thickness limits and increments, binder suitability and unit costs. The inference engine generates admissible sections, screens them with the Odemark approximation, verifies survivors with the elastic engine against all three distress criteria at the chosen reliability, and returns the least-cost section meeting the design life together with the governing distress.

III. VALIDATION METHODOLOGY

Three typical Iraqi sections were analysed over a 20-year period (Table 1). Both tools used the same layers, the same 80 kN standard axle at 552 kPa tyre pressure, the same design traffic and the same governorate air temperatures; MnPAVE was run at Intermediate design level with a user-defined location, PG 70-28 binder, CBR-20 base, CBR-12 subbase and an R-value 52 subgrade, which it converts internally to seasonal resilient moduli, using its five-season scheme, while IRQMEPAVE used its 12-month scheme. AASHTO 1993 [1] served as the empirical baseline with layer coefficients 0.44 (wearing), 0.42 (binder), 0.35 (asphaltic base), 0.14 (granular base) and 0.11 (subbase), S_0 = 0.45 and ΔPSI = 1.7. Agreement was assessed through the damage ratio D at design traffic, the percentage difference PD = 100 ($D_{IRQ} - D_{Mn}$)/D_{Mn}, the ratio D_{Mn}/D_{IRQ} and the coefficient of determination between log-transformed damage values.

Table 1. Case-study inputs

Parameter	CS1 – Baghdad	CS2 – Anbar (Ramadi)	CS3 – Basra (Umm Qasr)
Design ESAL (20 yr)	5×10^6	20×10^6	40×10^6
Reliability (%)	90	95	95
Wearing / binder course (mm)	50 / 60	50 / 70	50 / 80
Asphaltic base (mm)	110	120	160
Granular base / subbase (mm)	200 / 300	200 / 300	200 / 300
Subgrade CBR (%) / M_R (MPa)	5 / 52	5 / 52	5 / 52
Hottest-month air temp. (°C)	45.0 (August)	42.4 (August)	48.1 (August)
Binder grade / mix V_a, V_{beff} (%)	**PG 70-28** / 5.0, 11.0	**PG 70-28** / 5.0, 11.0	**PG 70-28** / 5.0, 11.0

IV. RESULTS AND DISCUSSION

A. AASHTO 1993 Baseline

AASHTO 1993 rates all three sections as adequate, with demand-to-capacity ratios of 0.16–0.65 (Table 2). The empirical method therefore gives no warning of distress for any case, which is the central limitation addressed by M–E analysis.

ISBN: 978-605-80193-6-2

Table 2. AASHTO 1993 check

Case	SN required	SN provided	Allowable W_{18} (10^6)	Demand / capacity
CS1	4.47	5.78	30.4	0.16
CS2	5.74	6.08	30.7	0.65
CS3	6.29	6.80	72.6	0.55

B. Effective Asphalt Temperature and Stiffness

Both tools convert air temperature to pavement temperature with the same MS-1 relation, but they are fed different statistics of the same climate. IRQMEPAVE uses the monthly mean air temperature (38.5 °C in the hottest month at Basra) and obtains 48.1 °C at 50 mm depth, giving |E*| = 1458 MPa; MnPAVE was supplied with the seasonal maximum air temperature (47 °C) and obtained 60 °C, giving a design modulus of about 262 MPa, reduced to 185–199 MPa by its reliability adjustment of inputs (which also trims 220 mm of asphalt to 206 mm in CS1). Substituting the monthly maximum air temperature (47.6 °C) into Eq. (1) at the same depth gives 58.6 °C, i.e. the MnPAVE value, and extrapolating the IRQMEPAVE temperature–modulus curve (R^2 = 0.999) to that temperature gives about 680 MPa. The definition of the effective temperature therefore explains most of the stiffness difference, and the residual factor of roughly 2.6 reflects the Witczak 1-37A model against the MnPAVE regression.

Table 3. Hottest-month stiffness and IRQMEPAVE critical strains (worst month)

Case	IRQ mean air (°C)	IRQ T_{AC} (°C)	IRQ E_{AC} (MPa)	Mn T_{pave} (°C)	Mn E_{AC} design / adjusted (MPa)	IRQ ε_t (με)	IRQ $\varepsilon_{v,sg}$ (με)
CS1	35.9	45.0	1774	60	262 / 199	144	164
CS2	33.6	42.4	2101	60	262 / 185	111	140
CS3	38.5	48.1	1458	60	262 / 185	103	128

C. Damage and Predicted Life

Table 4 and Figure 3 summarise the outcome. Both tools ranked the sections in the same order and produced correlated damage (R^2 = 0.76 between log D values), but MnPAVE's calibrated models gave 1.7–5.6 times more damage than IRQMEPAVE (PD from −42% to −82%, mean absolute difference 69%). Using the Asphalt Institute functions inside MnPAVE, i.e. the same transfer functions as IRQMEPAVE but with MnPAVE's soft summer moduli, increased the damage by a further order of magnitude (7–154 times IRQMEPAVE), which confirms that the stiffness assumption, not the distress equations, drives the divergence.

Table 4. Damage ratios D at design traffic and selected reliability

Case	Distress	IRQMEPAVE D	MnPAVE D (MnPAVE models)	MnPAVE D (AI models)	PD (%)	D_{Mn}/D_{IRQ}	IRQ life (yr)
CS1	Fatigue	0.90	3.53	97.2	−74.5	3.9	21.7
	Rutting	0.56 (asphalt)	1.39	6.37	−59.7	2.5	30.3
CS2	Fatigue	2.43	13.6	375	−82.1	5.6	9.7
	Rutting	1.57 (asphalt)	6.92	40.4	−77.3	4.4	14.0
CS3	Fatigue	3.14	14.2	391	−77.9	4.5	7.7
	Rutting	5.68 (asphalt)	9.74	40.8	−41.7	1.7	4.5

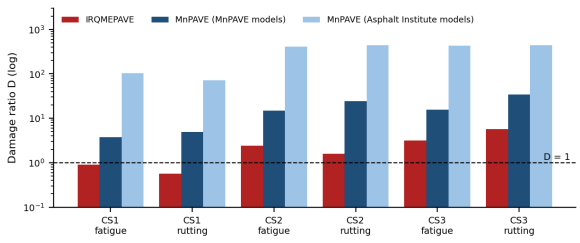


Fig. 3. Damage ratios from IRQMEPAVE and MnPAVE at design traffic (logarithmic scale)

The engineering conclusions are nevertheless consistent. CS1 passes in IRQMEPAVE (fatigue life 21.7 years at 90% reliability, 6.2 mm rut) and is the only section MnPAVE gives a multi-year life (6 years fatigue, 15 years rutting). CS2 and CS3 fail in both tools: IRQMEPAVE predicts 9.7 and 7.7 years to fatigue and 14.0 and 4.5 years to rutting, and MnPAVE predicts 1 year to fatigue for both. In IRQMEPAVE, rutting was controlled by asphalt-layer permanent deformation in every case, with predicted rut depths of 6.2, 10.4 and 22.6 mm against the 12.5 mm limit, so CS3 fails on rut depth alone. MnPAVE evaluates rutting only through the subgrade, and therefore cannot reproduce this mechanism, which is the distress most associated with high-temperature service.

Damage is strongly concentrated in the hot half of the year. In IRQMEPAVE, June–September contributed 68.5–72.8% of fatigue damage and 87.0–90.2% of asphalt-rutting damage, rising to 86–98% for May–October; MnPAVE assigned 77.5–81.9% of fatigue damage to its summer season and a further 0.1–1.8% to a spring-thaw season with granular layers weakened to 23 MPa, a mechanism that does not occur in Iraq.

D. Similarities, Differences and the Calibration Needed

The two tools implement the same computational chain (Table 5), so the divergence in Table 4 is not a difference of theory. It has three identifiable sources: the

ISBN: 978-605-80193-6-2

effective-temperature statistic (Section IV-B), the reliability mechanism (inputs in MnPAVE, life in IRQMEPAVE), and above all the calibration constants. MnPAVE's default functions are calibrated to Minnesota test sections (fatigue $C \cdot k_1 = 0.686$ with a shift relative to the Asphalt Institute form; subgrade rutting $k_1 = 0.0261$, $k_2 = -2.35$), whereas IRQMEPAVE currently runs the uncalibrated global Asphalt Institute and NCHRP 1-37A coefficients. A calibrated tool predicting more damage than an uncalibrated one is the expected result, not an error in either.

Taking the MnPAVE outputs as a provisional reference, the multipliers required on the IRQMEPAVE allowable repetitions are $\beta_f = 1/4.68 \approx 0.21$ for fatigue (range 0.18–0.26 over the three cases) and $\beta_r = 1/2.87 \approx 0.35$ for rutting (range 0.23–0.58). Expressed as strains, these correspond to multiplying ε_t by 1.60 and ε_v by 1.27. These are harmonisation factors to a Minnesota calibration, not an Iraqi calibration; they are reported so that the offset is explicit and reproducible.

The recommended calibration sequence for Iraqi conditions is: (i) fix the effective-temperature definition by installing pavement thermocouples and regressing measured mid-depth temperatures on air temperature, which is the single largest lever; (ii) verify |E*| by laboratory dynamic-modulus testing of Iraqi mixes with PEN 40–50 and 60–70 binders, replacing the predictive model where data exist; (iii) calibrate the fatigue constant $C \cdot k_1$ against cracking surveys on instrumented Iraqi sections; and (iv) calibrate β_{r1} of the asphalt-rutting model against measured rut depths, which is the most urgent step because this distress governed every case analysed here. Steps (iii) and (iv) follow the AASHTO local-calibration procedure [12], with the standard error S_0 re-estimated from the same data.

Table 5. Methodological comparison of the two tools

Feature	IRQMEPAVE	MnPAVE-Flexible 6.405
Response model	Multilayer linear elastic (+ Odemark screening)	Multilayer linear elastic
Climate resolution	12 months, Iraqi governorate database	5 seasons, Minnesota-based scheme
Asphalt modulus	Witczak 1-37A |E*| with binder A–VTS	Temperature–modulus regression by binder grade
Fatigue model	Asphalt Institute (global)	5 models incl. Asphalt Institute; MnPAVE calibrated by default
Subgrade rutting	Asphalt Institute (global)	6 models; MnPAVE calibrated by default
Asphalt-layer rutting	NCHRP 1-37A with k_z; rut depth reported	Not modelled
Damage accumulation	Miner, monthly	Miner, seasonal
Reliability	Life shifted by $10^{-Z \sim R \sim S \sim 0 \sim}$, 7 levels shown at once	Thickness and moduli reduced to the confidence level
Frost / thaw	Not applied (not relevant to Iraq)	Spring-thaw weakening of granular layers
Automated design	Find Ideal Section: knowledge-based least-cost search with governing-distress explanation	Design mode over a fixed set of trials
Deployment	Single offline HTML file, free	Windows installer, free

The practical value of the Iraqi tool lies in the last rows: Find Ideal Section turns the analysis into a design decision without manual iteration, the simultaneous display of seven reliability levels shows the cost of reliability directly, and the asphalt-rutting output gives a mechanism MnPAVE cannot report. The remaining procedural difference is the reliability mechanism, so part of the offset at 95% reliability is definitional rather than physical.

V. CONCLUSIONS

IRQMEPAVE integrates Iraqi monthly climate data, the Witczak 1-37A modulus model, a multilayer elastic engine, three distress models and a knowledge-based Find Ideal Section engine in a free, offline tool that follows the same computational chain as MnPAVE-Flexible. For three typical Iraqi sections: (1) the two tools ranked the sections identically and their damage ratios correlated ($R^2 = 0.76$ on a log scale), and both failed the two heavy-traffic sections, whereas AASHTO 1993 rated all three adequate; (2) IRQMEPAVE, running global coefficients, predicted 1.7–5.6 times less damage than MnPAVE with its Minnesota-calibrated coefficients, and closing this gap requires multipliers of about 0.21 on fatigue and 0.35 on rutting life; (3) the effective-temperature statistic is the largest single lever — using monthly maximum instead of mean air temperature raises the asphalt temperature from 48 to 59 °C and reproduces the MnPAVE pavement temperature; (4) asphalt-layer rutting governed rutting in every case, reaching 22.6 mm for the 40-million-ESAL section, a distress MnPAVE cannot evaluate; and (5) more than two-thirds of fatigue damage and about 90% of asphalt-rutting damage occurred between June and September. A four-step local calibration — pavement temperature, measured |E*|, fatigue constant and β_{r1} — is proposed, after which IRQMEPAVE offers Iraqi agencies an accessible route to mechanistic–empirical design.

Acknowledgements

The authors gratefully acknowledge Al-Mansour University College, Department of Civil Engineering-Baghdad, for supporting the development of IRQMEPAVE.

References

[1] AASHTO (1993). Guide for Design of Pavement Structures. American Association of State Highway and Transportation Officials, Washington, DC.

[2] ARA Inc., ERES Consultants Division (2004). Guide for Mechanistic-Empirical Design of New and Rehabilitated Pavement Structures. Final Report, NCHRP Project 1-37A, Transportation Research Board, Washington, DC.

[3] AASHTO (2020). Mechanistic-Empirical Pavement Design Guide: A Manual of Practice, 3rd ed. AASHTO, Washington, DC.

[4] Asphalt Institute (1982). Research and Development of the Asphalt Institute's Thickness Design Manual (MS-1), 9th ed. Research Report 82-2, Lexington, KY.

[5] Witczak, M.W., Fonseca, O.A. (1996). Revised predictive model for dynamic (complex) modulus of

ISBN: 978-605-80193-6-2

asphalt mixtures. Transportation Research Record 1540, 15–23.

[6] Timm, D., Birgisson, B., Newcomb, D. (1998). Development of mechanistic-empirical pavement design in Minnesota. Transportation Research Record 1629, 181–188.

[7] Timm, D.H., Newcomb, D.E., Galambos, T.V. (2000). Incorporation of reliability into mechanistic-empirical pavement design. Transportation Research Record 1730, 73–80.

[8] Huang, Y.H. (2004). Pavement Analysis and Design, 2nd ed. Pearson Prentice Hall, Upper Saddle River, NJ.

[9] Burmister, D.M. (1945). The general theory of stresses and displacements in layered systems. Journal of Applied Physics 16(2), 89–94.

[10] Odemark, N. (1949). Investigations as to the elastic properties of soils and design of pavements according to the theory of elasticity. Meddelande 77, Statens Väginstitut, Stockholm.

[11] SCRB (2003). General Specification for Roads and Bridges, Section R9: Hot-Mix Asphalt Concrete Pavement (revised ed.). State Corporation for Roads and Bridges, Ministry of Housing and Construction, Baghdad, Iraq.

[12] AASHTO (2010). Guide for the Local Calibration of the Mechanistic-Empirical Pavement Design Guide. AASHTO, Washington, DC.

[13] Miner, M.A. (1945). Cumulative damage in fatigue. Journal of Applied Mechanics 12, A159–A164.

[14] Minnesota Department of Transportation. MnPAVE-Flexible, version 6.405 [software]. MnDOT Office of Materials and Road Research, Maplewood, MN.

[15] Heukelom, W., Klomp, A.J.G. (1962). Dynamic testing as a means of controlling pavements during and after construction. Proc. 1st Int. Conf. on the Structural Design of Asphalt Pavements, Ann Arbor, MI, 667–679.

[16] Central Organization for Statistics and Information Technology (COSIT) (2023). Ministry of Planning, Baghdad, Iraq.

ISBN: 978-605-80193-6-2

Design, Fabrication, and Electromechanical Integration of an Original Experimental Device for Measuring Tissue Friction Force in Medical Needle–Tissue Interactions

Erman ÇELİK[1] Alparslan Ulaş ÇAYDAŞ[1] Süreyya Elif KÖM[1]

[1]Fırat University, Technology Faculty, Department of Mechanical Engineering, Elazığ, Türkiye

* Corresponding author e-mail: ucaydas@firat.edu.tr

***Abstract*— The accurate detection of friction forces generated during the interaction of medical needles with biological tissues in minimally invasive medical procedures, such as biopsy and injection, is of critical importance for procedural success and patient comfort. In this study, the design, fabrication, electromechanical integration, and mathematical modelling processes of an original testing device capable of measuring the friction forces generated during the movement of biopsy needles within tissue in a precise, stable, and repeatable manner are presented. The main mechanical structure of the device was constructed using high-rigidity aluminium sigma profiles, while the linear motion system was driven by a belt-and-pulley mechanism and a stepper motor. The total forces acting in the z-direction are instantaneously measured using a load cell, converted into the normal force exerted on the tissue side, and subsequently used to obtain the tissue friction force as a function of time. Within the scope of the study, the necessary Ethics Committee approvals were obtained, and a user-friendly experimental infrastructure supported by an electronic control panel equipped with an HMI touchscreen was successfully developed.**

***Keywords*— Medical needle, tissue friction force, mechatronic design, biomechanical testing device, load cell, HMI control panel.**

I. INTRODUCTION

Minimizing tissue damage, achieving precise positioning at the target site, and accurately analysing the tissue resistance perceived by the surgeon during minimally invasive surgical procedures and biopsy applications are important areas of research in biomedical engineering. During the advancement of a needle through biological tissue, puncture, cutting, and friction forces between the tissue surface and the needle wall occur simultaneously. Experimentally measuring these forces with high repeatability is essential for optimizing needle geometries and validating simulation models.

The aim of this study is to develop a novel, modular, and highly rigid tissue friction force measurement device, including its mechanical, electronic, and software components, capable of accurately measuring the tissue friction forces generated during the interaction of medical needles with biological tissues.

II. PHYSICAL MODEL

A novel experimental setup was developed to enable the precise and repeatable measurement of the friction forces generated during the interaction of medical needles with biological tissues. The experimental study plan consisted of the stages of device design, system fabrication, calibration procedures, experimental testing (Figure 1), and evaluation of the obtained data.

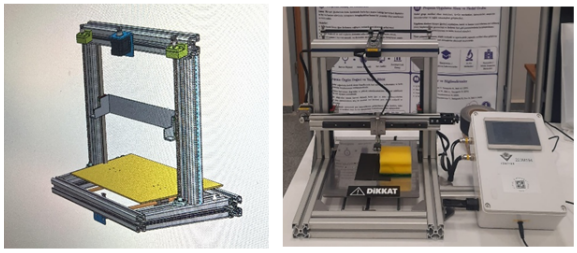


Fig. 1. Design and fabrication of the tissue friction force measurement device.

The mechanical manufacturing process of the tissue friction force measurement device began with the construction of the supporting frame of the designed system. In the first stage, the main frame of the device was constructed using aluminium sigma profiles, which provide high rigidity and modular assembly capabilities. The profiles were assembled using connection plates, corner brackets, T-slot nuts, and bolts. Axial alignment and geometrical accuracy checks were performed to ensure the proper operation of the linear motion system. In this way, mechanical vibrations and deformations that could occur during the experiments were intended to be minimized.

The following fundamental physical model was adopted in the design of the tissue friction force measurement device.

The biopsy needle moves downward:

M → Mass of the biopsy needle

g → Gravitational acceleration

v → Constant velocity of the needle

$F_{friction}$ → Friction force between the needle and the tissue

μ(t) → Time-varying coefficient of friction

N(t) → Normal force, i.e., the perpendicular force exerted by the tissue

Newton's Second Law:

$$\boldsymbol{F_{net} = M.a} \quad (1)$$

Since the needle moves toward the tissue at a constant velocity, the acceleration is a=0, and therefore, the net force is zero. Since the friction force is defined in terms of the coefficient of friction and the normal force, it can be expressed as follows:

ISBN: 978-605-80193-6-2

$$\boldsymbol{F_{friction}(t) = \mu(t) \times N(t)} \quad (2)$$

Therefore, the friction coefficient is expressed as follows:

$$\boldsymbol{\mu(t) = \frac{F_{sürtünme}(t)}{N(t)}} \quad (3)$$

Measured Values Obtained from the Device:

F_z (t) → Total force acting on the needle in the z-direction (measured using the load cell integrated into the device).

N(t) → Normal force (obtained through conversion).

Since the velocity is constant and there is no net acceleration, the friction force can be written as:

$$\boldsymbol{F_{friction}(t) = F_z(t) - M.g} \quad (4)$$

III. MECHANICAL DESIGN AND MANUFACTURING PROCESS

Using the developed interface, the friction force is measured as a function of time and can be recorded graphically on the device.

Following the completion of the supporting frame, the assembly of the linear motion system was initiated. Within this scope, the linear guide rail and linear carriage mechanism were fixed to the main body, enabling the moving platform to perform precise linear motion with low friction. A timing belt and pulley system was integrated to drive the motion mechanism, and the stepper motor connections were established to provide the motion transmission infrastructure. At the same stage, the moving platform, to which the load cell, friction probe, and experimental fixtures would subsequently be attached, was incorporated into the mechanical system.

In the final stage of the mechanical assembly, a belt tensioning system was installed to ensure stable and repeatable operation of the linear motion mechanism. An adjustable mechanism was used to ensure that the timing belt operated at the appropriate pre-tension level. Pulley-belt alignment was performed, and any backlash and deviations that could adversely affect the motion accuracy were eliminated. Thus, the mechanical infrastructure of the device was completed, and the system was made ready for the integration of the electronic control unit, load cell, friction measurement probe, data acquisition system, and other functional components.

As a result of these manufacturing stages, the mechanical platform required for the high-precision measurement of tissue friction force under controlled experimental conditions was successfully developed. Owing to its modular structure, the device provides a suitable infrastructure for subsequent sensor, control, and software integrations and was designed to be easily reconfigured according to different experimental parameters (Figure 2).

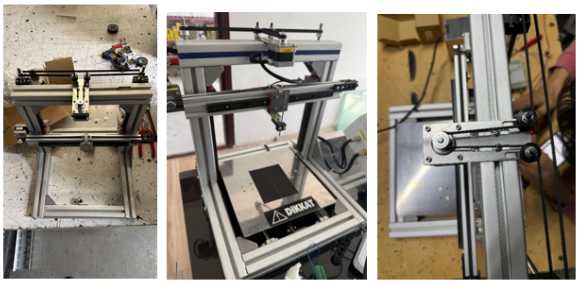


Fig. 2. Main frame assembly stages of the tissue friction force measurement device.

IV. ELECTRONIC CONTROL AND SYSTEM INTEGRATION

Following the completion of the mechanical assembly, the electronic control system of the device was developed. Within this scope, the motor driver, power supply, relay modules, and electrical connection components were installed inside the electrical control panel, thereby establishing the electrical infrastructure of the system. Subsequently, the load cell and probe mechanism used for measuring the friction force were precisely mounted on the moving platform and integrated with the control system. In the final stage, a touchscreen HMI, which serves as the user interface of the device, was integrated into the system, enabling the control of experimental parameters and real-time monitoring of measurement data. Thus, the integration of the mechanical, electronic, and control components was completed, and the device was successfully brought into operation (Figure 3).

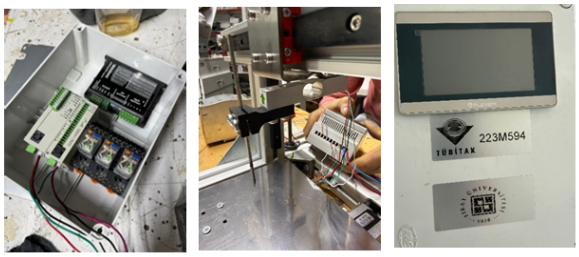


Fig. 3. Electronic control system and display of the device.

V. PILOT EXPERIMENTS

Before commencing the experimental studies, recent literature on medical needle–tissue interactions was reviewed, and the methods used for measuring friction forces were evaluated. Based on the literature review, the technical requirements, experimental parameters, measurement ranges, and performance criteria of the developed measurement system were determined.

To ensure surface cleanliness of the needles to be used in the tissue friction force experiments following pathological procedures, the specimens were cleaned in an ultrasonic cleaning device using an acetic acid solution for 1 hour (Figure 4). Following the cleaning process, the needles were sequentially rinsed with 70% ethyl alcohol and distilled water, and were then dried and prepared for the experiments.

ISBN: 978-605-80193-6-2



Fig. 4. Cleaning of biopsy needles in an ultrasonic cleaning device.

In the experiments conducted using the developed device, the measurement parameters were set as follows: a feed rate of 20 mm/s, a total travel distance of 1000 mm, and a needle insertion speed of 200 mm/s. The tissue specimens used in the experiments were prepared with dimensions of 40 mm³. The effective needle length was set to 40 mm to ensure proper penetration of the needle into the tissue and to prevent deflection effects caused by the needle length.

VI. RESULTS AND DISCUSSION

In this study, the fabrication and electromechanical integration of the mechatronic testing device designed to measure the friction forces arising from the interaction of medical needles with biological tissues were successfully completed.

The system demonstrates high rigidity and low vibration performance owing to its aluminium sigma-profile frame, linear guide-and-carriage system, and belt-and-pulley mechanism. The integrated load cell and electronic control system equipped with an HMI interface enable the precise and repeatable recording of time-dependent tissue friction forces. This novel testing infrastructure is considered a critical experimental tool for biomedical research and medical device design.

Acknowledgment

The authors gratefully acknowledge the Scientific and Technological Research Council of Türkiye (TÜBİTAK) for its financial support of this study under Project No. 223M594.

References

[1] Astley, C. A., & Gibson, T. (2018). Friction forces in medical needle insertion into soft tissues. Journal of Biomechanics, 68, 45-52.

[2] O'Leary, R. P., et al. (2020). Design and calibration of tissue friction measurement systems for biopsy needles. IEEE Transactions on Biomedical Engineering, 67(4), 1102-1110.

[3] Smith, J. M., and Dogan, A. (2022). Linear motion control and force sensing in biomedical tissue testing. Medical Engineering & Physics, 99, 103740.

ISBN: 978-605-80193-6-2

Development and Validation of a Generalized Decision-Support Tool for Engineered Air-Nozzle Retrofits in Compressed-Air Blow-Gun Systems

Orhan Kaya
Department of Mechanical Engineering, Kırşehir Ahi Evran University, Kırşehir 40100, Türkiye
* Corresponding author e-mail: orhan.kaya@ahievran.edu.tr

***Abstract*—Replacing open compressed-air blow guns with engineered air nozzles is one of the most frequently recommended energy-efficiency measures in industry, yet its savings are usually estimated with simplified spreadsheets whose assumptions remain hidden. This study presents a generalized decision-support tool that evaluates such retrofits with a thermodynamically consistent model and, unlike conventional worksheets, automatically checks the plausibility of the input data against the installed compressor. The tool supports different compressor types and control strategies, both reduced-orifice and catalog-rated engineered nozzles, and optional pressure regulation. Applied to a case from the Industrial Assessment Center database, it revealed errors in the pressure basis and in the simultaneity assumption of the original assessment that would have led to unrealistic savings estimates. The results show that the benefit of a nozzle retrofit depends strongly on the selected product, the supply pressure, and the operating hours. The proposed approach provides a transparent and reusable basis for more reliable compressed-air retrofit decisions.**

***Keywords*—Compressed air, engineered air nozzle, venturi nozzle, blow gun, energy assessment, decision-support tool, energy efficiency**

I. INTRODUCTION

Compressed air is an expensive industrial utility because only a limited fraction of the electrical energy supplied to the compressor is ultimately converted into useful work at the point of use [1], [2]. Industrial surveys consistently identify leakage, inappropriate end uses, and excessive operating pressure as major sources of avoidable energy consumption [1], [3], [4]. Open blow guns used for cleaning, drying, and chip removal are a typical inappropriate end use. An unrestricted tube or drilled orifice can discharge large volumes of compressed air at sonic velocity while also producing high noise levels and dead-end pressures that may exceed occupational safety limits [2], [7].

Engineered air nozzles, including venturi-type designs, reduce compressed-air consumption by entraining ambient air through the Coanda effect or by distributing the flow through multiple small passages. A comparable blowing action can therefore be obtained with substantially lower compressed-air demand, depending on nozzle design and operating pressure. Replacing open blow guns with engineered nozzles is consequently a common recommendation in industrial energy assessments and is listed in the U.S. Department of Energy Industrial Assessment Center (IAC) database under assessment recommendation code 2.4236 [8]. Savings are commonly estimated with spreadsheet worksheets that combine an orifice-flow relation [5] with an adiabatic compression-work relation [6]. However, such worksheets are often weakly documented, may conceal inherited assumptions, and are difficult to adapt when a proposed nozzle is specified by catalog air consumption rather than by an equivalent orifice diameter.

This study revisits a blow-gun retrofit recommendation from an industrial energy assessment and develops it into a generalized and validated decision-support tool. The main contributions are:

- A transparent formulation of nozzle flow, compression work, demand, energy use, and cost, extended to include subsonic-flow correction, catalog-rated engineered nozzles, pressure regulation, and a compressor part-load factor.
- A stand-alone Windows application based on the thermodynamically consistent absolute-pressure formulation, with automatic plausibility checks, a built-in case study, and an input guide.
- Application to an IAC case study, showing that the simultaneity assumed in the original worksheet exceeded the compressor capacity, and scenario analyses of nozzle selection, pressure regulation, and operating hours.

II. MATERIALS AND METHODS

A. Case Description

The case study was taken from the Industrial Assessment Center (IAC) database under recommendation code 2.4236 [8]. The assessed facility used 40 hand-held blow guns with 3.18 mm (1/8 in) open orifices to clean production equipment. The guns were supplied by a 22.4 kW (30 hp), single-stage rotary-screw compressor, and the line pressure at the nozzles was 7.93 bar(g) (115 psig). The original worksheet assumed that ten guns operated simultaneously for 125 h per year and represented the recommended air-induced nozzles by a 1.59 mm (1/16 in) orifice. As shown in Section III, ten simultaneous guns are not compatible with the installed compressor; the case study therefore uses three simultaneous guns. The principal input parameters are summarized in Table I. The governing equations are presented in the US customary units used in the original formulation because the empirical constants are defined in those units; the software converts all user inputs and outputs between US customary and SI units.

TABLE I. Input Parameters of the Case Study

Parameter	Value (SI)	Value (US)
Atmospheric pressure	101.3 kPa	14.7 psia

ISBN: 978-605-80193-6-2

Parameter	Value (SI)	Value (US)
Compressor inlet air temperature	30.0 °C	86 °F
Air temperature at nozzle	21.1 °C	70 °F
Compressor discharge pressure	8.62 bar(g)	125 psig
Line pressure at nozzle	7.93 bar(g)	115 psig
Compressor motor (rotary screw, single stage)	22.4 kW	30 hp
Motor / adiabatic efficiency	0.85 / 0.82	0.85 / 0.82
Existing / proposed orifice diameter	3.18 / 1.59 mm	1/8 / 1/16 in
Installed / simultaneously operating nozzles	40 / 3	40 / 3
Annual operating hours	125 h	125 h
Demand coincidence factor / billing months	0.8, 12	0.8, 12
Energy tariff / demand tariff	0.127 USD/kWh / 3.81 USD/kW-month	same
Nozzle cost / installation rate / labor rate	50 USD, 2 per h, 33 USD/h	same

B. Nozzle Flow Model

For a sharp-edged orifice discharging to the atmosphere, the free-air volumetric flow rate Qf (cfm) under choked-flow conditions is given by [5]

$$Q_f = \frac{C_d(T_i+460)C_1C_2\pi D^2 P_l}{4C_3P_a\sqrt{T_l+460}} \quad (1)$$

where Cd = 0.8 is the discharge coefficient; Ti and Tl are the compressor-inlet and nozzle-air temperatures (°F); Pl and Pa are the absolute nozzle-supply and atmospheric pressures (psia), respectively; D is the orifice diameter (in); C1 = 28.37 ft/(s·°R0.5); C2 = 60 s/min; and C3 = 144 in²/ft². Choked flow occurs when Pa/Pl does not exceed the critical pressure ratio

$$r_c = \left(\frac{2}{k+1}\right)^{\frac{k}{k-1}} \quad (2)$$

For air with k = 1.4, the critical pressure ratio is 0.5283. At higher downstream-to-upstream pressure ratios, the flow becomes subsonic. In that region, the result of (1) is multiplied by the ratio of the subsonic mass-flux function to its critical value, with r = Pa/Pl:

$$\varphi(r) = \frac{\sqrt{\frac{2k}{k-1}\left(r^{\frac{2}{k}} - r^{\frac{k+1}{k}}\right)}}{\sqrt{k}\left(\frac{2}{k+1}\right)^{\frac{k+1}{2(k-1)}}} \quad (3)$$

C. Compressor Power, Demand and Energy

The electrical input power required to supply Qf was estimated from the adiabatic work needed to compress the free air from Pa to the absolute discharge pressure Po in N stages [6]:

$$L = \frac{fP_aC_3Q_fNC_4}{E_aE_m}\frac{k}{k-1}\left[\left(\frac{P_o}{P_a}\right)^{\frac{k-1}{kN}} - 1\right] \quad (4)$$

where C4 = 3.03×10−5 hp·min/(ft·lbf), Ea and Em are the adiabatic and motor efficiencies, respectively, and f is a part-load response factor. Setting f = 1 assumes that compressor power varies directly with air demand, consistent with the original worksheet. Values below unity can be used to approximate less favorable load-unload or modulating behavior. The peak demand reduction DR, billed demand DB, annual energy use E, and annual cost C for one nozzle are calculated as

$$DR = C_5LCF, \quad DB = MDR \quad (5)$$

$$E = C_5LH \quad (6)$$

$$C = Ec_e + DBc_d \quad (7)$$

where C5 = 0.746 kW/hp, CF is the demand coincidence factor, M is the number of billing months, H is the annual operating time, and ce and cd are the energy and demand tariffs, respectively. For ns simultaneously operating nozzles, the annual savings are

$$S = n_s\left(C_{ex} - C_{pr}\right) \quad (8)$$

D. Engineered Nozzles, Regulation and Economics

Engineered nozzles are commonly specified by a catalog air consumption Qcat at a rated pressure pcat. When the internal nozzle passages remain choked, the consumption at an actual supply pressure pn can be scaled by the ratio of absolute pressures:

$$Q_{pr} = Q_{cat}\frac{p_n+P_a}{p_{cat}+P_a} \quad (9)$$

If a regulator is installed, pn is taken as the lower of the regulated pressure and the available line pressure. The implementation cost and simple payback period are then calculated as

$$IC = n_rc_n + \frac{n_r}{r_i}c_l + c_o \quad (10)$$

$$SPB = \frac{IC}{S} \quad (11)$$

where nr is the number of nozzles replaced, cn is the unit nozzle cost, ri is the installation rate, cl is the labor rate, and co represents other implementation costs.

E. Pressure Basis

Equations (1), (2), and (4) require absolute pressures, obtained by adding the atmospheric pressure to the measured gauge values. The original worksheet inserted gauge pressures directly into the pressure ratio of (1) and the compression ratio of (4). At the line pressure of the case study this lowers the calculated flow by the factor $P_{l,g}/(P_{l,g} + P_a) = 0.887$, and the error grows as the supply pressure decreases. The software therefore evaluates all cases with absolute pressures only.

F. Software Implementation

The model was implemented in C# as a Windows Forms application for .NET 8 without third-party libraries, allowing distribution as a single self-contained executable. The software architecture is summarized in Fig. 1. Users can enter data in either US customary or SI units, fractional-inch orifice diameters are accepted, and costs can be expressed in US dollars, euros, or Turkish lira. Compressor type is used to assign the adiabatic efficiency automatically. After each calculation, the application reports nozzle-level and facility-level results, implementation cost, simple payback, and diagnostic warnings. The warnings identify subsonic flow, proposed nozzle consumption greater than the

ISBN: 978-605-80193-6-2

existing condition, simultaneous demand exceeding compressor motor input or rated capacity, power reductions greater than 50% of motor input, and potential violation of the 30 psi (2.07 bar) dead-end pressure limit for cleaning air [7]. Four plots are generated: free-air flow and annual cost versus orifice diameter, annual savings versus line pressure, and simple payback versus annual operating hours, which can be entered up to 8,064 h/yr. The IAC case study described above is available as a built-in case, and an input guide lists every input with its meaning and a typical example value. Results can be exported as a multi-sheet spreadsheet, a text report, and high-resolution figures. The program outputs were verified against an independent implementation of (1)-(11).

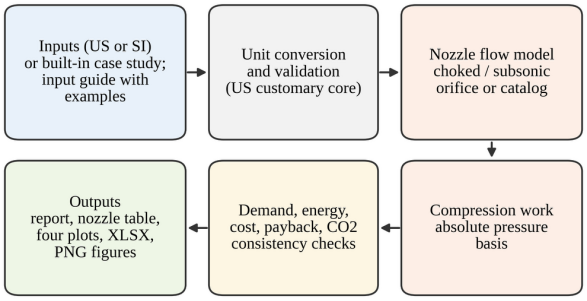


Fig. 1. Structure of the decision-support tool.

III. RESULTS AND DISCUSSION

A. Compressor Capacity Check

With absolute pressures, one 3.18 mm gun operating at 7.93 bar(g) consumes 0.688 m³/min of free air and requires 5.27 kW of compressor electrical input. The installed 22.4 kW motor corresponds to about 26.3 kW of electrical input at the stated motor efficiency of 0.85. Ten simultaneous guns, as assumed in the original worksheet, would require 6.88 m³/min and 52.7 kW, about twice the available input, so the stated line pressure could not be maintained. The software flags this condition, and it also applies the 50% power-reduction rule of the IAC worksheet. With five simultaneous guns the demand reaches 100% of the compressor input and the power reduction 75%; with four guns the demand is acceptable but the reduction still exceeds 50%. Three simultaneous guns are the largest number for which both checks are satisfied, with a demand of 2.06 m³/min and 15.8 kW (60% of the compressor input). This value is used for the case study.

B. Case Study Results

Table II summarizes the case study results, and Fig. 2 illustrates the difference between the existing open-orifice gun and the proposed air-induced nozzle. Halving the orifice diameter reduces the air consumption of each gun from 0.688 to 0.172 m³/min and the compressor input from 5.27 to 1.32 kW. For three simultaneous guns, the compressor power falls by 11.85 kW, the peak demand by 9.48 kW, and the annual electricity use by 1,482 kWh. Because the guns operate only 125 h per year, the demand charge provides USD 434 of the USD 622 annual savings. With an implementation cost of USD 2,660 for 40 nozzles and 20 h of labor, the simple payback is 4.28 years.

TABLE II. Case Study Results (Three Simultaneous Guns)

Quantity	Existing	Proposed
Orifice diameter (mm)	3.18	1.59
Free-air flow per gun (m³/min)	0.688	0.172
Compressor input per gun (kW)	5.27	1.32
Total free-air flow (m³/min)	2.06	0.52
Total compressor input (kW)	15.80	3.95

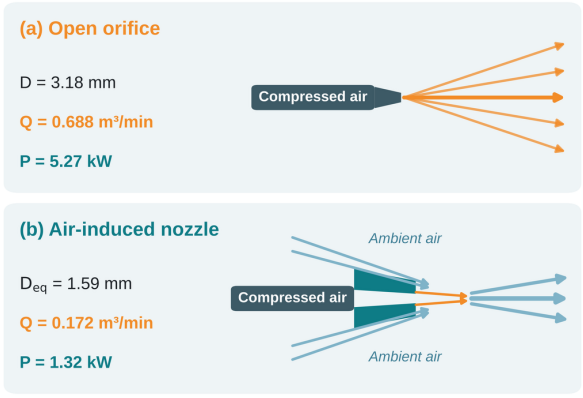


Fig. 2. (a) Existing open-orifice air gun and (b) proposed air-induced nozzle of Table II. D: orifice diameter (D_{eq}: equivalent orifice used in the calculations); Q: free-air flow and P: compressor input per gun at 7.93 bar(g).

TABLE III. Annual Savings and Economics of the Case Study

Quantity	Value
Compressor power reduction (kW)	11.85
Peak demand reduction (kW)	9.48
Billed demand reduction (kW-month/yr)	113.8
Electricity savings (kWh/yr)	1,482
Energy cost savings (USD/yr)	188
Demand cost savings (USD/yr)	434
Total cost savings (USD/yr)	622
Implementation cost (USD)	2,660
Simple payback (years)	4.28

C. Effect of Orifice Diameter and Line Pressure

Fig. 3 shows the free-air flow per nozzle as a function of orifice diameter. Because the flow is proportional to the square of the diameter, halving the diameter lowers the air consumption by 75%. Fig. 4 shows the annual savings of the case study as a function of the line pressure at the nozzle. The savings increase almost linearly with pressure, from USD 142 per year at 1.0 bar(g) to USD 358 at 4.1 bar(g) and USD 622 at 7.93 bar(g), because the choked flow is proportional to the absolute supply pressure.

ISBN: 978-605-80193-6-2

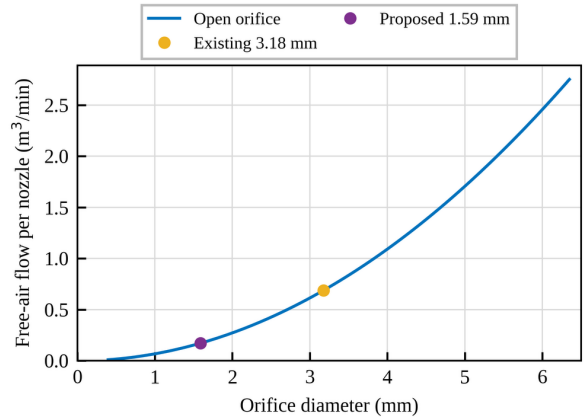


Fig. 3. Free-air flow per nozzle as a function of orifice diameter at 7.93 bar(g).

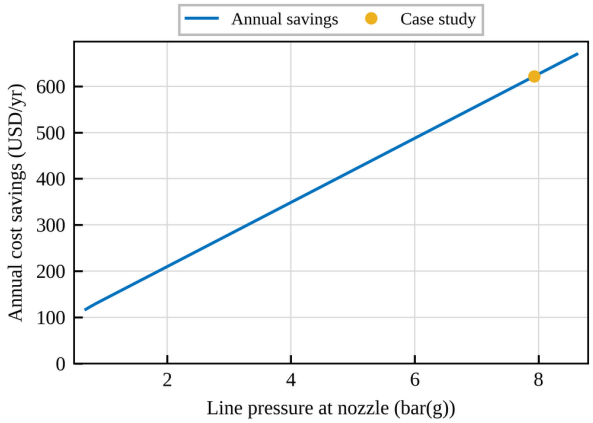


Fig. 4. Annual cost savings of the case study as a function of line pressure at the nozzle.

D. Retrofit Scenarios

Table IV compares three retrofit options for three simultaneous guns. S1 is the reduced orifice of the case study. S2 and S3 use an engineered nozzle with a representative catalog consumption of 0.40 m³/min (14 scfm) at 5.52 bar(g) (80 psig); in S2 it operates at the full line pressure, and in S3 a regulator limits the supply pressure to the catalog-rated value.

TABLE IV. Retrofit Scenarios (Three Simultaneous Guns, 125 h/yr)

Scenario	Air reduction (%)	Energy savings (kWh/yr)	Cost savings (USD/yr)	SPB (yr)
S1 Reduced orifice, 1.59 mm	75.0	1,482	622	4.28
S2 Engineered nozzle at 7.93 bar(g)	21.0	416	174	15.25
S3 Engineered nozzle regulated to 5.52 bar(g)	42.3	837	351	7.58

The engineered nozzle reduced air consumption by 21% at full line pressure and by 42% when regulated, compared with 75% for the 1.59 mm orifice. The reduced-orifice representation therefore implies a much lower air demand than the representative engineered product and can overestimate savings when used as its surrogate. A technically equivalent comparison should also account for blowing force, which requires manufacturer data; the catalog-input mode allows a selected product to be evaluated directly. Pressure regulation doubled the savings of the engineered nozzle.

E. Sensitivity to Operating Hours

The annual operating time of 125 h strongly limits the economic benefit. Fig. 5 shows the simple payback as a function of annual operating hours for S1-S3. At low operating hours the curves flatten because the demand-charge savings do not depend on operating time. At 2,000 h per year the payback falls to 0.77, 2.75, and 1.37 years for S1, S2, and S3, respectively; at 4,000 h per year to 0.41, 1.47, and 0.73 years; and at the upper input limit of 8,064 h per year to 0.21, 0.75, and 0.37 years. Accurate measurement of blow-gun operating time is therefore as important as nozzle selection.

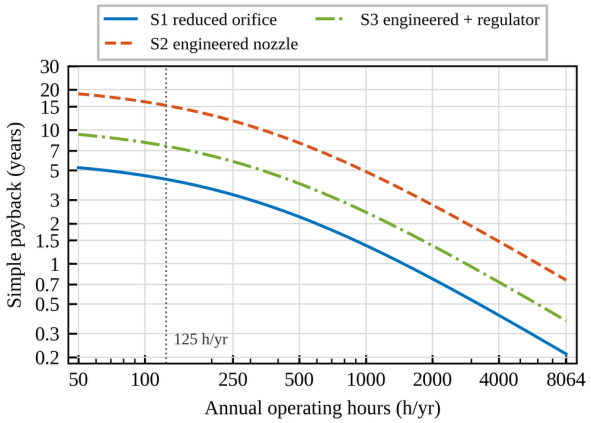


Fig. 5. Simple payback as a function of annual operating hours for scenarios S1-S3.

F. Limitations

Compressor power is related to air demand through a simplified part-load factor, and load-unload cycling, receiver storage, compressor sequencing, and pressure losses in the distribution network are not simulated. Equation (9) assumes choked internal flow when scaling catalog air consumption with pressure, and the discharge coefficient of worn or nonstandard guns may differ from 0.8. The number of simultaneously operating guns was derived from the capacity checks rather than measured. Measured air-flow, pressure, and compressor-power data should be used whenever available.

IV. CONCLUSION

A compressed-air blow-gun retrofit recommendation from the Industrial Assessment Center database was generalized into a stand-alone decision-support tool that evaluates nozzle flow and compression work with absolute pressures and checks the plausibility of the input data. Applied to the case study, the capacity checks showed that the ten simultaneous guns assumed in the original worksheet would require about twice the input of the installed compressor; three simultaneous guns were the largest number satisfying both the capacity and the 50% power-reduction checks. For this case, replacing 3.18 mm open orifices by 1.59 mm orifices reduced air consumption by 75% and yielded annual savings of 1,482 kWh and USD 622, with a simple payback of 4.3 years at 125 h per year. Engineered nozzles with a representative catalog consumption saved 21% of the air at full line pressure and 42% with regulation, showing that representing an engineered nozzle by a reduced orifice can overstate savings. The payback depends strongly on operating hours and falls below one year at 2,000 h per year for the reduced

ISBN: 978-605-80193-6-2

orifice. By making the assumptions explicit and flagging implausible inputs, the tool provides a more transparent basis for blow-gun retrofit assessments. Future development should include equal-thrust nozzle selection, measured demand profiles, and detailed compressor-control models.

Data Availability Statement

The decision-support software (source code and build scripts), the case-study input data, the calculation outputs, and the data underlying the results in Tables II–IV and Figs. 2–4 are archived on Figshare at **https://doi.org/10.6084/m9.figshare.34036674**

References

[1] Saidur, R., Rahim, N. A., & Hasanuzzaman, M. (2010). A review on compressed-air energy use and energy savings. Renewable and Sustainable Energy Reviews, 14(4), 1135–1153. https://doi.org/10.1016/j.rser.2009.11.013

[2] U.S. Department of Energy & Compressed Air Challenge. (2016). Improving compressed air system performance: A sourcebook for industry (3rd ed.). U.S. Department of Energy, Office of Energy Efficiency and Renewable Energy, Advanced Manufacturing Office. https://www.energy.gov/sites/prod/files/2016/03/f30/Improving%20Compressed%20Air%20Sourcebook%20version%203.pdf

[3] Kaya, D., Phelan, P., Chau, D., & Sarac, H. I. (2002). Energy conservation in compressed-air systems. International Journal of Energy Research, 26(9), 837–849. https://doi.org/10.1002/er.823

[4] Dindorf, R. (2012). Estimating potential energy savings in compressed air systems. Procedia Engineering, 39, 204–211. https://doi.org/10.1016/j.proeng.2012.07.026

[5] Shapiro, A. H. (1953). The dynamics and thermodynamics of compressible fluid flow (Vol. 1). Ronald Press.

[6] Rollins, J. P. (Ed.). (1989). Compressed air and gas handbook (5th ed.). Compressed Air and Gas Institute; Prentice Hall.

[7] Occupational Safety and Health Administration. (n.d.). 29 CFR 1910.242: Hand and portable powered tools and equipment, general. U.S. Department of Labor. Retrieved September 30, 2026, from https://www.osha.gov/laws-regs/regulations/standardnumber/1910/1910.242

[8] U.S. Department of Energy. (n.d.). Industrial Assessment Centers (IAC) database [Data set]. Retrieved September 30, 2026, from https://iac.university

ISBN: 978-605-80193-6-2

Accuracy and Computational Cost of Modeling Choices in Computational Fluid Dynamics Simulations of Gas-Solid Cyclone Separators: A Comparative Review

Nurettin ÇAY[1,*] Erman ÇELİK[2]

[1]Fırat University, Graduate School of Natural and Applied Sciences, Department of Mechanical Engineering, Elazığ, Türkiye

[2]Fırat University, Faculty of Technology, Department of Mechanical Engineering, Elazığ, Türkiye

Article note: This paper is derived from the thesis of Nurettin ÇAY entitled "Gaz-Katı Siklon Ayrıştırıcılarda Dalma Borusu Geliştirilmesi ve Performans Çıktılarının Araştırılması" (Development of a Vortex Finder for Gas-Solid Cyclone Separators and Investigation of Its Performance Outputs).

* Corresponding author e-mail: nurettin.cay.9@gmail.com

***Abstract*—Computational fluid dynamics is now routinely used to design gas-solid cyclone separators, but the result of a simulation depends on several modeling decisions that are seldom assessed together. In this paper we compare these decisions using published studies of reverse-flow gas cyclones, and we place special weight on eight studies that simulated the same case, the 0.29 m Stairmand high-efficiency cyclone for which laser Doppler anemometry data are available. The decisions examined are the turbulence model, grid size and near-wall resolution, discretization schemes and time step, steady or transient solution, boundary conditions and extent of the domain, coupling of the particle phase, and the quantities used for validation. Standard eddy-viscosity models do not reproduce the forced vortex in the core. When a curvature correction is added, the realizable k-epsilon model gives mean velocities close to those of the Reynolds stress model in roughly one fifth of the computing time. The Reynolds stress model remains the most reliable Reynolds-averaged option, but only with a transient solution, a pressure-staggering scheme, a high-order convection scheme and a time step below about one thousandth of the gas residence time. Steady solutions gave pressure drop errors of 16 to 20 percent. Large eddy simulation was the only method that resolved the precessing vortex core and the velocity fluctuations, and it costs one to two orders of magnitude more. Leaving the dust outlet out of the domain changed the pressure loss by about 10 percent and the cut size by about 35 percent. Pressure drop on its own was not a reliable validation measure. Based on these findings, a decision matrix is given that relates the level of modeling to the purpose of the simulation.**

***Keywords*—Cyclone separator, computational fluid dynamics, turbulence modeling, large eddy simulation, computational cost, gas-solid flow**

I. INTRODUCTION

Cyclone separators are used in many processes for removing particles from gas streams. They contain no moving parts, are cheap to build and maintain, and can operate at high temperatures, pressures and dust loads [1], [2]. Their performance is usually expressed by the pressure drop and the grade efficiency. Both depend on a strongly swirling three-dimensional flow with an outer free vortex, an inner forced vortex, flow reversal near the cone tip and a precessing vortex core (PVC) [1], [3], [4]. Semi-empirical design models cannot account for these features or for small changes in geometry. For this reason computational fluid dynamics (CFD) is increasingly used in cyclone design, for example when the vortex finder is modified [5], [6].

The outcome of a cyclone simulation depends on a series of modeling decisions, and the literature does not agree on most of them. The standard k-ε model has long been regarded as unsuitable for cyclones [7], [8], and for more than twenty years the Reynolds stress model (RSM) has been the usual choice among Reynolds-averaged Navier–Stokes (RANS) models [9], [10]. Several studies report that large eddy simulation (LES) performs clearly better [4], [7], [11]. Others show that eddy-viscosity models (EVMs) with a curvature correction can match the RSM at much lower cost [10], [12]. Advice on convection schemes also differs between authors [7], [13], [14]. Cost often decides the matter in practice. Parametric and orthogonal-design studies with dozens of geometries often use the standard k-ε model for this reason [15], yet the computing time saved is rarely given.

Earlier reviews describe the available turbulence and particle models in detail [2], [5], but they do not compare accuracy and cost on the same basis. In this paper we compare the main modeling decisions for gas-solid cyclones and, where the published data allow it, express the differences in numbers. Seven aspects are considered: turbulence closure, grid and near-wall resolution, discretization and time step, steady or transient formulation, boundary conditions and domain, particle-phase coupling, and validation. The study is part of a thesis on vortex finder development. In that work a large number of geometries has to be simulated, so the level of modeling must be chosen with care.

II. METHODOLOGY OF THE COMPARATIVE ASSESSMENT

A. Literature Corpus

We selected numerical and experimental studies of reverse-flow gas cyclones published between 2000 and 2026. A study was included if it compared at least two options for the same modeling decision, or if it reported the grid, time step and computing time in enough detail to judge accuracy against cost. Two studies on gas-liquid swirl separators [15], [16] and a semi-analytical study with Lagrangian particle tracking [17] were used only as

 ISBN: 978-605-80193-6-2

supporting evidence. From each paper we recorded the turbulence model, grid size and type, near-wall resolution, discretization schemes, time step, formulation, boundary conditions, domain, particle model and coupling, validation quantity and computing time.

B. Common Benchmark

Results from different groups can only be compared directly when the geometry and operating conditions are the same. The Stairmand high-efficiency cyclone with a barrel diameter of 0.29 m, measured by Hoekstra with laser Doppler anemometry (LDA) at an inlet velocity of 16.1 m/s, serves this purpose [18]. The Reynolds number based on the barrel diameter is about 2.8×10^5 [19]. The data include mean and root-mean-square (RMS) tangential and axial velocities at three heights and the pressure drop. Eight of the studies reviewed here simulated this cyclone. They are listed in Table I, and the other studies in Table II.

TABLE I. Reviewed Simulations of the 0.29 m Stairmand Benchmark Cyclone

Ref.	Closure	Grid / time step	Main finding
[13]	RSM, nine scheme sets	461,272 nodes; 1×10^{-4} s	Steady: Δp +16 to +20%; transient with PRESTO!: within ±2%, at ≈10× CPU time
[11]	URANS-RSM (LG), LES	590,730 cells; 5×10^{-5} s	Mean-velocity error LES 2.94/1.97% vs RSM 3.27/2.48% (tangential/axial)
[12]	SST, SST-CC, SST-CCM	Three hexahedral levels	SST-CCM reaches RSM accuracy, 8% faster than SST-CC
[14], [20]	RSM-LPS, RSM-QPS	245,356 to 756,788 hex.; $y^+ \approx 30$	Medium grid sufficient; third-order convection required
[19]	RSM-QPS	756,788 cells; 1×10^{-4} to 1×10^{-3} s	$\Delta t < \tau_{res}/1470$; 1×10^{-3} s is 9.4× faster but fails validation
[10]	Six EVMs ± CC, two RSMs	756,788 cells; 1×10^{-4} s	RKE-CC ≈ RSM for mean flow; EVM-CC ≈ 5× faster
[21]	k-ε, k-ω, RSM-LPS, LES	0.10 to 2.28 M (tetrahedral)	EVMs overpredict near-wall swirl; only LES resolves wall bursts

C. Accuracy and Cost Measures

Accuracy is taken as reported in each paper. Depending on the study, it is given as the deviation of the mean velocity profiles (least-square error, hit rate or normalized mean square error), the deviation of the pressure drop, or the error in overall or grade efficiency. Cost is given as wall-clock or CPU time. Hardware and parallel settings differ from one study to another, so we compare only cost ratios that were obtained within the same study.

TABLE II. Reviewed Simulations of Other Cyclone Geometries

Ref.	Approach	Grid / time step	Main finding
[9]	RSM, LES	40 k / 640 k; 5×10^{-5} s	LES slightly better; reveals core oscillation
[7]	k-ε, RSM-LRR, LES	0.66 / 1.3 M; 1×10^{-4} s	LES best; 242 h on 40 processors
[4]	URANS-RSM, FV-LES, LB-LES	1.32 / 4.95 M	Steady RANS diverged; RMS and PVC captured only by LES
[22]	RSM	260,230; 1×10^{-4} s	No dust outlet: ≈10% error in Euler number, ≈35% in cut size
[6]	LES	1.17 M; 1×10^{-5} s	Grid independence verified by GCI
[23]	LES + Lagrangian	0.4 / 1.8 M	No dispersion model needed; collection criterion matters
[24]	LB-LES, two-way	$\approx10^7$ particles	Loading causes swirl loss and turbulence damping
[25]	LES, four-way	0.83 / 1.3 M; 1×10^{-5} s	Agglomeration model alters grade efficiency
[26]	RSM + DEM	47,750; 1×10^{-4} s	2 mm particles needed to keep DEM affordable
[27]	Five RANS + DPM	300,697 uniform	Similar Δp; grade-efficiency errors 34 to 48%

III. TURBULENCE CLOSURE

A. Eddy-Viscosity Models and Curvature Correction

Standard EVMs use the Boussinesq assumption of an isotropic eddy viscosity. In a cyclone this leads to an excessive turbulent viscosity, the forced vortex in the core is lost, and the computed rotation tends toward solid-body motion that reaches close to the wall. In the cyclone studied by Shalaby et al. [7], the measured maximum tangential velocity was 1.24 times the inlet velocity at 25% of the radius. The k-ε model gave a maximum of 1.2 times the inlet velocity, shifted toward the wall, and at another height it gave 1.4 times the inlet velocity close to the wall. Savin and Khoperskov [21] also found that k-ε and k-ω overestimate the tangential velocity near the walls, with the largest errors in the axial velocity of the core. For a square cyclone, RNG k-ε overestimated and standard k-ε underestimated the velocities; the grade efficiency error was 34% with the RSM and 48% with standard k-ε [27].

A curvature correction adds a production term that makes the EVM respond to rotation and streamline curvature. Alahmadi and Nowakowski [12] modified the curvature-corrected shear stress transport model (SST-CC) and obtained an accuracy similar to that of the RSM for the benchmark cyclone. Their version needed 8.15% less computing time than the original SST-CC (12 h 14 min compared with 13 h 18 min). Thongnoi et al. [10] tested six EVMs, with and without curvature correction, against RSMs with linear (LPS) and quadratic (QPS) pressure–strain terms on one grid. With suitable near-wall cell sizes and a third-order convection scheme, some EVMs without correction already gave acceptable results in the barrel. The correction improved all of them, and the realizable k-ε model with curvature correction (RKE-CC) predicted the mean flow as well as the RSMs. In single-inlet cylindrical swirl separators the realizable k-ε model was as accurate as LES [16]. These findings show that RSMs are not the only acceptable RANS option for the mean flow. They do not extend to the RMS velocities, which should not be taken from RANS results without caution [10].

ISBN: 978-605-80193-6-2

B. Reynolds Stress Models

The RSM solves a transport equation for each Reynolds stress component and can therefore represent the anisotropy of swirling turbulence [2], [9]. The pressure–strain model has a visible influence on the result. The Launder–Reece–Rodi formulation without wall reflection gave a maximum tangential velocity of 0.95 times the inlet velocity against the measured 1.24 in [7]. Adding the wall-reflection term lowered the least-square error of the tangential profile from 3.31% to 2.74% in [11], and the quadratic formulation gave better mean velocities than the linear one [14]. The additional cost is shown in Fig. 1. With the same grid, time step and hardware, one RSM iteration took 1.04 s and one EVM iteration 0.63 to 0.79 s. The RSMs also needed more iterations, and their total computing time was five to six times that of the corrected EVMs [10]. Because the RSM equations are stiff, an EVM solution is often used as the initial field [10].

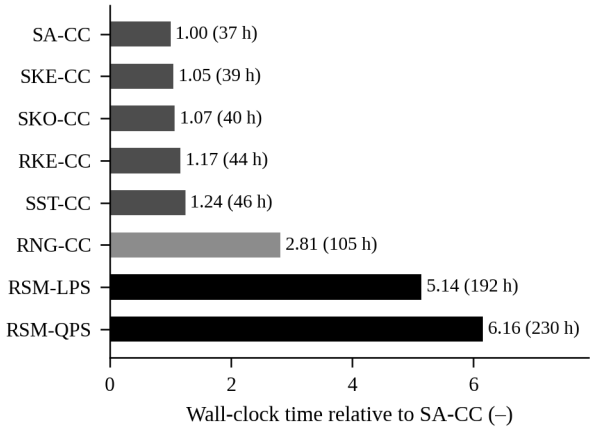


Fig. 1. Total wall-clock time of curvature-corrected EVMs and RSMs for the 0.29 m Stairmand cyclone with the same grid (756,788 cells), time step and hardware, divided by the time of the Spalart–Allmaras model; data from [10].

C. Large Eddy Simulation

In LES the large eddies are computed and only the subgrid scales are modeled. Among the reviewed approaches, only LES reproduced the precession of the vortex core. Slack et al. [9] compared LES on 640,000 cells with an RSM on 40,000 cells; the LES agreed slightly better with the LDA data at five of eight heights and showed core oscillations that are averaged out in a steady RSM solution. Shalaby et al. [7] obtained a maximum tangential velocity of 1.25 times the inlet velocity with LES, against 1.24 in the experiment. Gronald and Derksen [4] found that all approaches gave similar mean axial velocities, but only LES predicted the RMS velocities, and it gave the PVC frequency within 10%. For the benchmark cyclone the mean tangential and axial errors were 2.94% and 1.97% with LES and 3.27% and 2.48% with the RSM. The computed pressure drop was 691.7 Pa with LES and 1027.7 Pa with the RSM, compared with the 773.1 Pa reported as the measured value [11]. LES also showed short bursts in the lower cone and velocity streaks close to the wall, which RANS models cannot represent [21].

The residence time in a cyclone is short and the large eddies dominate the flow, so the choice of subgrid model is less important [9]. Dynamic and wall-adapting (WALE) models still improve the behavior near the wall [21]. Most of the extra cost comes from the time step. One LES time step takes about as long as one RSM step, because LES solves fewer equations but needs three to four times more inner iterations. The LES time step, however, must be 7 times (engineering use) to 100 times (research use) smaller [21]. In the study of Shalaby et al., 35,000 LES time steps took 242 h on 40 processors [7]. Averaging must also cover a long period. Slack et al. noted that only about 2.5 precession periods were included in the upper part of their cyclone, which biased the averaged profiles there [9]. Hybrid RANS–LES methods such as detached eddy simulation might reduce the cost, but none of the reviewed studies tested them against the benchmark data.

IV. GRID RESOLUTION AND NEAR-WALL TREATMENT

As Fig. 2 shows, grid sizes grew from about 4×10^4 cells in early RSM studies to $0.4–0.8\times10^6$ cells for RANS and $0.6–5\times10^6$ cells for LES. For the benchmark cyclone, block-structured hexahedral grids of 245,356, 426,552 and 756,788 cells showed that the medium grid is enough for models from Spalart–Allmaras to the RSM [10], [20]. A grid of 461,272 nodes differed by less than 1% from one of 526,174 nodes [13]. On a grid built from triangulated surfaces, 0.67×10^6 cells were acceptable and about 1.2×10^6 cells were needed for good agreement, while 10^5 cells were not enough in the core [21]. Unstructured tetrahedral grids therefore appear to need more cells than hexahedral grids for the same accuracy. For LES, 590,730 cells gave nearly the same result as 1,100,394 cells [11], 1,174,029 hexahedra were shown to be grid independent with the grid convergence index (GCI) [6], and 830,000 control volumes were sufficient for the turbulence levels [25].

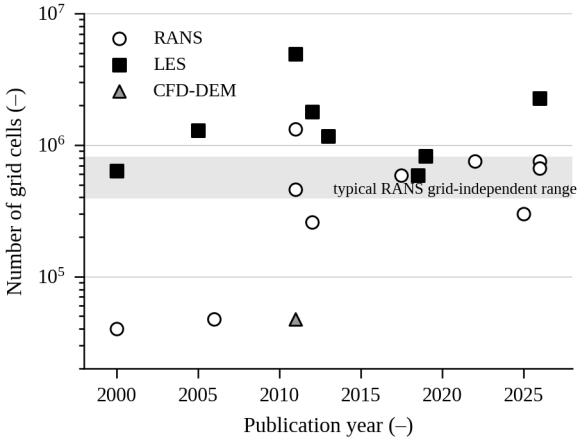


Fig. 2. Number of grid cells in the reviewed cyclone simulations by year of publication and modeling approach. The shaded band is the range in which RANS solutions of the 0.29 m benchmark were reported to be grid independent.

The number of cells is not sufficient on its own. With wall functions the first node next to the wall has to lie in the logarithmic layer. Bumrungthaichaichan [20] estimated the distance of this node separately for the barrel, the cone and the vortex finder, using a target dimensionless wall distance of 30:

$$y_p = y^+ \mu / [\rho\, U\, (\lambda/2)^{1/2}] \quad (1)$$

where y_p is the distance of the first node from the wall, y^+ the target dimensionless wall distance, μ and ρ the viscosity and density of the gas, U the local mean

ISBN: 978-605-80193-6-2

velocity and λ the friction factor taken from a pressure-drop model. The resulting wall distances were 1.27 to 1.66 times the target, and the mean velocities agreed better with the measurements [20]. Models that are integrated down to the wall, such as stress-ω models, need y^+ values near one. This leads to cells with high aspect ratios and to convergence problems [14]. Grid convergence should be checked quantitatively, for instance with the GCI method [28], as was done in [6], [15], [19]. Grid independence shown on very coarse grids should be treated with care. Examples are a difference of less than 5% between 25,900 and 95,350 cells [29], the same grid later used for a CFD–DEM study [26], and a single uniform grid without any convergence test [27]. In such cases numerical diffusion may hide the differences between grids.

V. DISCRETIZATION, TIME STEP AND FORMULATION

A. Spatial Discretization

For pressure interpolation the studies agree. In transient computations, PRESTO! combined with any higher-order convection scheme gave the benchmark pressure drop within ±2%, while schemes without pressure staggering gave values 10% to 41% too low [13]. For convection there is no agreement. Shukla et al. [13] recommended second-order upwinding for momentum and turbulence quantities and first-order upwinding for the Reynolds stresses; this combination needed 12% to 22% less CPU time than the other higher-order sets. Bumrungthaichaichan [14] found that only the third-order QUICK scheme kept the flow pattern correct in six cyclones, especially when the ratio of vortex finder diameter to cone tip diameter is less than one. Shalaby et al. [7] suggested central differencing for the RSM. Part of this disagreement may come from the different grids used. First-order schemes, on the other hand, failed in every comparison.

B. Steady Versus Transient Solution and Time Step

Steady RANS computations of cyclones often do not converge [4], [16], and when they do, the result can be misleading. Shukla et al. [13] found pressure drop errors of +16% to +20% with steady RSM runs and within ±2% with transient runs; the transient runs took about ten times longer (236 h compared with 24 h). The time step has to be a small fraction of the gas residence time [30]. Thongnoi et al. [19] used time steps between 1/7380 and 1/738 of the residence time. The hit rates for the mean axial velocity were 0.75, 0.69, 0.55 and 0.55, and only the two smallest time steps passed the threshold of 0.66. The largest time step made the run 9.42 times faster but failed the validation. From these results they proposed

$$\Delta t < \tau_{res}/1470, \quad \tau_{res} = V/Q \tag{2}$$

where Δt is the time step, τ_{res} the residence time, V the cyclone volume and Q the volume flow rate. The LES studies used time steps between 1×10^{-5} and 1×10^{-4} s [4], [6], [9], [25]. In the benchmark cyclone a semi-periodic flow was reached after about 1.5 s, and time averaging should not begin before this [19]. The LES with particles in [25] covered about 3 s of flow time, or 300,000 time steps.

VI. BOUNDARY CONDITIONS AND COMPUTATIONAL DOMAIN

A uniform inlet velocity with a turbulence intensity of 3.7% to 10% and the hydraulic diameter of the inlet as length scale is common practice [11], [13], [21]. The results are not sensitive to the inlet profile or to random fluctuations imposed at the inlet [4], [7]. Savin and Khoperskov [21] extended the inlet duct by five inlet heights and the outlet pipe by two barrel diameters to reduce the influence of the boundaries. In [12] the standard and the extended outlet geometry gave different flow fields. At the gas outlet most RANS studies used a zero-gradient outflow condition [13], [19]. The dust outlet had the largest effect of all domain choices. Without it, the Euler number changed by about 10% and the cut-off diameter by about 35%, although the maximum tangential velocity hardly changed [22]. If no dust bin is included, particles that reach the underflow have to be counted as collected; counting only escaping particles gives unrealistic grade efficiencies [23]. Transient RSM and LES runs were usually started from a converged solution of a simpler model, which shortens the start-up period [4], [6], [9].

VII. PARTICLE-PHASE MODELING

The particle phase involves a second set of decisions, which are included in the lower part of Table III. In dilute flows, where the particle volume fraction is of the order of 10^{-4}, particles are usually tracked in the computed gas field without feedback on the gas [11], [29]. With RANS a turbulent dispersion model is necessary. Without it, the number of escaping parcels was about one order of magnitude too small and the collection efficiency too high [11]. In RSM computations the continuous random walk (CRW) model gave an efficiency error of 2.55%, compared with 3.41% for the discrete random walk (DRW), and the anisotropy of the velocity fluctuations had to be considered. In LES the two models gave similar errors (1.47% and 1.65%) [11]. LES does not need a tuned dispersion model, but the particle time step (2.5 µs in [23]) and the integration scheme affect the grade efficiency. Lift caused by velocity gradients mainly affects particles of 7.5 to 20 µm, and higher turbulence intensity lowers the efficiency for particles below about 17.5 µm [17].

At higher loadings the coupling between the phases must be included. Mass loadings of 0.05 and 0.1 reduced the swirl and damped the turbulence of the gas; the first effect lowers the efficiency and the second raises it [24]. Two-way coupling was important even at low overall loading, because the particles collect in strands near the wall. Adding particle collisions and agglomeration changed the computed grade efficiency considerably [25]. The discrete element method (DEM) coupled with CFD reproduces the particle strands and the decrease in pressure drop and tangential velocity at solid-to-gas loading ratios up to 2.5. Its cost, however, rises steeply as the particles become smaller. Chu et al. [26] therefore used 2 mm particles, a Young modulus reduced to 10^7 Pa and a grid of 47,750 cells. At present CFD–DEM

ISBN: 978-605-80193-6-2

explains mechanisms rather than predicting the efficiency for fine particles; coarse-graining is the main way to make it affordable [5]. No particle model can make up for errors in the gas flow. In [11], LES predicted the overall collection efficiency better than the RSM mainly because its axial velocity and cut-off diameter were closer to the measurements.

VIII. VALIDATION METRICS AND REPORTING

Agreement in pressure drop alone does not show that a model is correct. In the square cyclone study, all five RANS models gave similar pressure drops and overall efficiencies, but their grade efficiency errors ranged from 34% to 48% [27]. A steady solution with a lower-order scheme matched the measured pressure drop within 3% while its tangential velocity profile was wrong [13]. The way the pressure drop is defined also matters. For the same inlet velocity of 16.1 m/s, the measured pressure drop of the Hoekstra cyclone is given as 477.1 Pa in [13] and as 773.1 Pa in [11]. Pressure drop validations from different studies cannot be compared unless the pressure definition (static or total), the measurement positions and the cyclone configuration are stated.

Velocity profiles are a stricter test. Most advanced models reproduce the tangential profile, so the axial velocity in the core is better at separating good models from poor ones [11], [19]. The measurements have their own uncertainty: LDA errors in the core are 3% to 5% for mean velocities and above 10% for fluctuations, and the uncertainty of the axial velocity can reach 20% of its typical value [21]. Statistical measures such as hit rate, normalized mean square error and fractional bias [19], together with GCI error bars [20], [28], make the comparison less subjective. Only a few of the reviewed papers gave the computing time and the hardware. If these were reported together with the grid type, near-wall distance, schemes, time step and averaging period, accuracy and cost could be compared directly.

IX. ACCURACY–COST SYNTHESIS AND RECOMMENDATIONS

Taking a curvature-corrected EVM on a given grid as reference, the other corrected EVMs cost about 1.05 to 1.24 times as much, RNG k-ε about 2.8 times because of convergence problems, and the RSMs 5 to 6 times (Fig. 1). A transient RSM run costs about ten times more than a steady one [13], and LES adds a further factor of 7 to 100 [21]. Table III uses these ratios to suggest a modeling level for each purpose of the simulation.

TABLE III. Decision Matrix Linking Simulation Objective, Recommended Set-Up and Relative Cost

Objective	Recommended set-up	Relative cost
Geometry screening (many variants)	Transient RANS with RKE-CC or SST-CC; 0.4 to 0.8 M hexahedral cells; $y^+ \approx 30$; PRESTO! and QUICK; $\Delta t < \tau_{res}/1470$	1 (baseline)
Mean flow and pressure drop	Transient RSM with quadratic or wall-reflection pressure–strain term, initialized from the EVM solution; same grid	≈5 to 6
Separation efficiency, dilute flow	RSM + Lagrangian tracking with anisotropic CRW dispersion, or LES + Lagrangian; dust outlet included	≈5 to 6 (RSM); 10 to 100× RSM (LES)
Unsteady physics (PVC, fine particles, erosion)	LES with dynamic Smagorinsky or WALE subgrid model; 1 to 5 M cells; $\Delta t \approx 1\times10^{-5}$ s	10 to 100× RSM
Mass-loading effects, particle strands	Two- or four-way LES–Lagrangian, or CFD–DEM with coarse-grained particles	Highest

For the vortex finder study that motivated this work, Table III suggests a staged approach. Geometric variants can be compared first with a transient curvature-corrected EVM. The most promising designs can then be checked with an RSM started from the EVM solution, and the final design can be run with LES if core precession or fine particles are of interest. The vortex finder dimensions change the axial velocity profile from an inverted W to an inverted V shape and have a strong effect on the pressure drop [6]. The results also become more sensitive to the convection scheme when the ratio of vortex finder diameter to cone tip diameter is below one [14]. The near-wall cell size should therefore be recalculated with (1) for each variant, and the dust outlet should be kept in all models, so that differences between designs are not caused by the numerical set-up.

X. CONCLUSIONS

The main conclusions of this review are as follows:

- Standard EVMs are not suitable for cyclones. With curvature correction, RKE-CC and the modified SST-CC model give mean velocities similar to those of the RSM in about one fifth of the computing time, which makes them useful for comparing many geometries.
- The RSM is a reliable RANS model only when it is solved transiently with PRESTO!, a high-order convection scheme, a suitable pressure–strain model and a time step below 1/1470 of the residence time. Steady RSM solutions gave pressure drop errors of 16% to 20%.
- Only LES resolved the precessing vortex core, the velocity fluctuations and the bursts near the wall. Its cost is one to two orders of magnitude higher, mainly because of the smaller time step.
- The near-wall cell size and the grid type are more important than the number of cells. For RANS of the 0.29 m benchmark, $0.4–0.8\times10^6$ hexahedral cells were enough, and grid convergence should be shown with the GCI.
- Leaving out the dust outlet caused errors of about 10% in pressure loss and 35% in cut size. Pressure drop alone, for which different measured values are quoted in the literature, is not an adequate check of a model.
- Hybrid RANS–LES methods have not yet been tested on the benchmark case, and computing times are rarely reported. Further work on both points

ISBN: 978-605-80193-6-2

would allow the level of modeling to be chosen on firmer grounds.

References

[1] Hoffmann, A. C., & Stein, L. E. (2008). *Gas cyclones and swirl tubes: Principles, design and operation* (2nd ed.). Springer.

[2] Cortés, C., & Gil, A. (2007). Modeling the gas and particle flow inside cyclone separators. *Progress in Energy and Combustion Science, 33*, 409–452.

[3] Hoekstra, A. J., Derksen, J. J., & Van den Akker, H. E. A. (1999). An experimental and numerical study of turbulent swirling flow in gas cyclones. *Chemical Engineering Science, 54*, 2055–2065.

[4] Gronald, G., & Derksen, J. J. (2011). Simulating turbulent swirling flow in a gas cyclone: A comparison of various modeling approaches. *Powder Technology, 205*, 160–171. https://doi.org/10.1016/j.powtec.2010.09.007

[5] El-Emam, M. A., Zhou, L., & Yasser, E. (2024). Numerical investigation of cyclone separators: Physical mechanisms and theoretical algorithms. *ChemBioEng Reviews, 11*(2), 299–338. https://doi.org/10.1002/cben.202300052

[6] Elsayed, K., & Lacor, C. (2013). The effect of cyclone vortex finder dimensions on the flow pattern and performance using LES. *Computers & Fluids, 71*, 224–239. https://doi.org/10.1016/j.compfluid.2012.09.027

[7] Shalaby, H., Pachler, K., Wozniak, K., & Wozniak, G. (2005). Comparative study of the continuous phase flow in a cyclone separator using different turbulence models. *International Journal for Numerical Methods in Fluids, 48*(11), 1175–1197. https://doi.org/10.1002/fld.942

[8] Kaya, F., & Karagöz, İ. (2008). Performance analysis of numerical schemes in highly swirling turbulent flows in cyclones. *Current Science, 94*(10), 1273–1278.

[9] Slack, M. D., Prasad, R. O., Bakker, A., & Boysan, F. (2000). Advances in cyclone modelling using unstructured grids. *Transactions of the Institution of Chemical Engineers, Part A, 78*, 1098–1104.

[10] Thongnoi, P., Chalermsinsuwan, B., Chandra-ambhorn, W., et al. (in press). Reassessment of eddy viscosity turbulence models as alternatives to Reynolds stress models for gas cyclone separator CFD simulations. *Scientific Reports*. https://doi.org/10.1038/s41598-026-61193-3

[11] Jang, K., Lee, G. G., & Huh, K. Y. (2018). Evaluation of the turbulence models for gas flow and particle transport in URANS and LES of a cyclone separator. *Computers & Fluids, 172*, 274–283. https://doi.org/10.1016/j.compfluid.2018.04.032

[12] Alahmadi, Y. H., & Nowakowski, A. F. (2016). Modified shear stress transport model with curvature correction for the prediction of swirling flow in a cyclone separator. *Chemical Engineering Science, 147*, 150–165. https://doi.org/10.1016/j.ces.2016.03.023

[13] Shukla, S. K., Shukla, P., & Ghosh, P. (2011). Evaluation of numerical schemes using different simulation methods for the continuous phase modeling of cyclone separators. *Advanced Powder Technology, 22*(2), 209–219. https://doi.org/10.1016/j.apt.2010.11.009

[14] Bumrungthaichaichan, E. (2023). A note of caution on numerical scheme selection: Evidence from cyclone separator CFD simulations with appropriate near-wall grid sizes. *Powder Technology, 427*, 118713. https://doi.org/10.1016/j.powtec.2023.118713

[15] Chen, D., Zhang, J., Li, X., Lv, R., Li, Y., Kong, X., Lin, T., & Wang, S. (2025). Analysis of structural parameters' influence on separation performance in a gas–liquid cyclone separator. *Processes, 13*(12), 4021. https://doi.org/10.3390/pr13124021

[16] Hreiz, R., Gentric, C., & Midoux, N. (2011). Numerical investigation of swirling flow in cylindrical cyclones. *Chemical Engineering Research and Design, 89*, 2521–2539. https://doi.org/10.1016/j.cherd.2011.05.001

[17] Li, X., Yan, J., Cao, Y., Ni, M., & Cen, K. (2003). Numerical simulation of the effects of turbulence intensity and boundary layer on separation efficiency in a cyclone separator. *Chemical Engineering Journal, 95*, 235–240. https://doi.org/10.1016/S1385-8947(03)00109-8

[18] Hoekstra, A. J. (2000). *Gas flow field and collection efficiency of cyclone separators* [Doctoral dissertation, Delft University of Technology]. Delft, The Netherlands.

[19] Thongnoi, P., Chandra-ambhorn, W., Chalermsinsuwan, B., Wattananusorn, S., Wongpromrat, P., & Bumrungthaichaichan, E. (2024). RANS equation-based gas cyclone separator CFD simulation: An appropriate time step size. *Chemical Engineering Transactions, 113*, 649–654. https://doi.org/10.3303/CET24113109

[20] Bumrungthaichaichan, E. (2022). How can the appropriate near-wall grid size for gas cyclone CFD simulation be estimated? *Powder Technology, 396*, 327–344. https://doi.org/10.1016/j.powtec.2021.10.031

[21] Savin, E. S., & Khoperskov, A. V. (2026). Comparative analysis of turbulent models for gas flow dynamics in cyclone separators. *Technologies, 14*(5), 282. https://doi.org/10.3390/technologies14050282

[22] Elsayed, K., & Lacor, C. (2012). The effect of the dust outlet geometry on the performance and hydrodynamics of gas cyclones. *Computers & Fluids, 68*, 134–147. https://doi.org/10.1016/j.compfluid.2012.07.029

[23] de Souza, F. J., Salvo, R. V., & Martins, D. A. M. (2012). Large eddy simulation of the gas–particle flow in cyclone separators. *Separation and*

ISBN: 978-605-80193-6-2

Purification Technology, 94, 61–70. https://doi.org/10.1016/j.seppur.2012.04.006

[24] Derksen, J. J., Sundaresan, S., & Van den Akker, H. E. A. (2006). Simulation of mass-loading effects in gas–solid cyclone separators. *Powder Technology, 163*, 59–68. https://doi.org/10.1016/j.powtec.2006.01.006

[25] Sgrott, O. L., & Sommerfeld, M. (2019). Influence of inter-particle collisions and agglomeration on cyclone performance and collection efficiency. *The Canadian Journal of Chemical Engineering, 97*. https://doi.org/10.1002/cjce.23371

[26] Chu, K. W., Wang, B., Xu, D. L., Chen, Y. X., & Yu, A. B. (2011). CFD–DEM simulation of the gas–solid flow in a cyclone separator. *Chemical Engineering Science, 66*(5), 834–847. https://doi.org/10.1016/j.ces.2010.11.026

[27] Izzah, C. I. N., Yunardi, Reza, M., Sylvia, N., Malahayati, N., Mulana, F., & Fairweather, M. (2025). Comparative assessment of turbulence models for predicting square cyclone separator performance. *Journal of Advanced Research in Fluid Mechanics and Thermal Sciences, 127*(1), 140–160. https://doi.org/10.37934/arfmts.127.1.140160

[28] Celik, I. B., Ghia, U., Roache, P. J., Freitas, C. J., Coleman, H., & Raad, P. E. (2008). Procedure for estimation and reporting of uncertainty due to discretization in CFD applications. *Journal of Fluids Engineering, 130*(7), 078001. https://doi.org/10.1115/1.2960953

[29] Wang, B., Xu, D. L., Chu, K. W., & Yu, A. B. (2006). Numerical study of gas–solid flow in a cyclone separator. *Applied Mathematical Modelling, 30*, 1326–1342. https://doi.org/10.1016/j.apm.2006.03.011

[30] Chuah, T. G., Gimbun, J., & Choong, T. S. Y. (2006). A CFD study of the effect of cone dimensions on sampling aerocyclones performance and hydrodynamics. *Powder Technology, 162*(2), 126–132.

 ISBN: 978-605-80193-6-2

Optimization of Automotive B Pillar Crash Performance Through Local Carbon Fiber Composite Reinforcement

Ege Uğur[1,2] Samet Deniz Mayda[2] Mahmut Pekedis[1]
[1]Ege University, Department of Mechanical Engineering, Bornova, Izmir, Türkiye
[2]Oyak Renault Automobile Factories, Bursa, Turkey
* Corresponding author e-mail: ege.ugur@renault.com

***Abstract*— Reducing overall vehicle weight while maximizing passenger safety remains a critical objective in the automotive industry in order to meet stringent energy efficiency and emission targets. This requirement is particularly important for improving fuel efficiency. As one of the primary structural components of the vehicle body, the B-pillar plays a crucial role in protecting the occupant survival space by resisting excessive deformation during side-impact collisions. Nowadays, lightweight composite materials possessing high tensile strength and excellent ductility are widely utilized in the automotive industry. In this study, local carbon fiber reinforced polymer (CFRP) prepreg reinforcements were applied to the critical stress concentration regions of the steel B-pillar reinforcement component to enhance side-impact crashworthiness, and the crash performance of the structure was optimized using numerical methods. The developed hybrid structure was evaluated under static/quasi-static lateral impact loading using the finite element method. Both the unreinforced baseline model and the composite-reinforced hybrid model were analysed under identical boundary conditions. Numerical analysis results demonstrated that the maximum total deformation was significantly reduced, indicating a considerable improvement in structural stiffness compared to the baseline configuration. Furthermore, the maximum equivalent stress concentrated around the upper mounting holes decreased substantially in the reinforced structure. In addition, the maximum equivalent plastic strain remained within a safe range. Despite these significant improvements in structural integrity and stiffness, the increase in total mass was limited to a minimal level. The findings reveal that locally applied composite reinforcements in targeted regions can effectively reduce cabin intrusion and critical stress concentrations while introducing only a slight weight penalty. Therefore, this approach presents an effective design methodology for the development of lightweight and high-safety automotive body structures.**

***Keywords*— B-Pillar, carbon fiber reinforced polymer, impact energy absorption, lightweight design**

I. INTRODUCTION

Occupant protection in lateral collisions relies heavily on the internal B-pillar reinforcement steel, which acts as the primary transverse load path resisting catastrophic intrusion into the passenger survival cell. While increasing the steel sheet gauge enhances flexural rigidity, it imposes a significant mass penalty that conflicts with vehicle lightweighting and energy efficiency targets. Conversely, manufacturing the entire component from carbon fiber reinforced polymers (CFRP) remains cost-prohibitive in high-volume production.[1] Targeted metal-composite hybrid structures offer a balanced alternative by applying localized composite patches strictly to high-stress and critical deformation regions. Vehicle safety assessment organizations, such as Euro NCAP, subject vehicles to a series of standardized crash tests and evaluate their performance through a comprehensive rating system. The resulting safety scores and test outcomes serve as key indicators for consumers when comparing vehicle safety levels and making purchasing decisions. Consequently, enhancing side-impact crash performance is critical for improving both occupant protection and the perceived safety of the vehicle, thereby contributing to greater customer confidence and market competitiveness.

II. METHOD

The central B-pillar represents one of the most critical structural load-bearing components in vehicle body-in-white (BIW) architecture, bridging the transverse connection between the roof rail and the floor rocker panel. Maintaining the structural integrity of the B-pillar during a lateral collision is vital to preserving the occupant survival cell and minimizing cabin intrusion.[2] In modern automotive designs, inner steel reinforcements are strategically placed within the pillar assembly based on field crash data and literature to mitigate localized stress and deformation peaks.

In this study, a production-representative automotive B-pillar reinforcement made of a 2.00 mm thick structural steel substrate with a baseline mass of 0.815 kg was investigated. The primary objective is to enhance the lateral crashworthiness and rigidity of the baseline component while imposing a minimal mass penalty. Consequently, critical high-stress regions on the inner surface were locally reinforced with a 1.30 mm thick carbon fiber reinforced polymer (CFRP) prepreg laminate and evaluated under identical boundary conditions using finite element analysis. Following the local CFRP integration, the total component mass increased marginally to 0.941 kg, corresponding to an added weight penalty of only 0.126 kg.

A. Material Constitutive Models and Properties

The baseline production B-pillar reinforcement steel component was analysed using nonlinear finite element formulations under side-impact loading conditions. In accordance with standard vehicle body architecture assembly, fixed support boundary conditions were prescribed at the structural mounting holes. Fig. 1 illustrates the locations of the applied lateral load and fixed support constraints.

ISBN: 978-605-80193-6-2

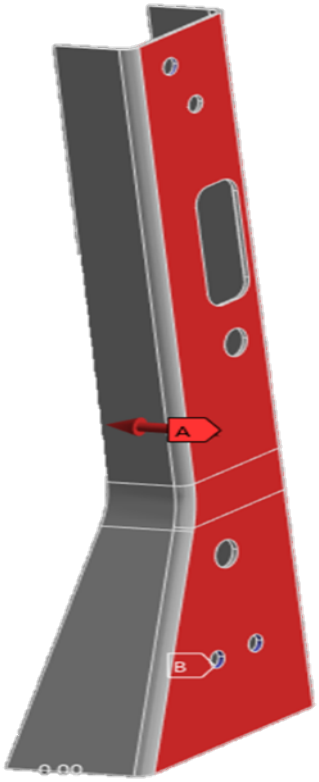


Fig. 1. Fixed and Force Point. (A remark is force, B remark is fixed)

Representative of lateral automotive crash pulses, a dynamic transverse load of 6.00 kN was applied to the outer impact surface over a total duration of 40 ms (t = 40). The transient response was solved across 12 discrete sub-steps, enabling the continuous tracking of localized equivalent von Mises stress evolution and intrusion displacement fields at each loading interval.For comparative crashworthiness evaluation, a hybrid B-pillar reinforcement model locally reinforced with a carbon fiber reinforced polymer (CFRP) prepreg laminate was developed in the ANSYS SpaceClaim. Interfacial compatibility between the steel substrate and composite plies was modeled using a bonded contact formulation, and the hybrid assembly was evaluated under identical boundary and loading conditions. In both analyses, the finite element models were discretized using an element size of 5 mm.

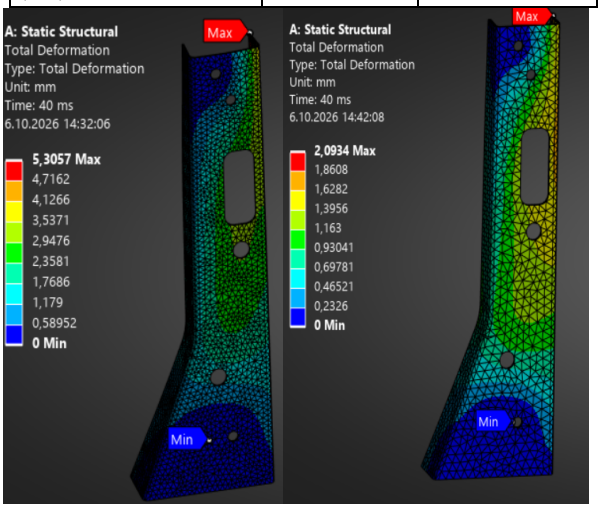
The physical and mechanical constitutive properties assigned to both materials in the numerical simulations are summarized in Table 1.

III. RESULTS AND DISCUSSION

The primary objective of B-pillar crashworthiness is to limit transverse intrusion into the passenger survival cell during lateral impact. The transient deformation response of both configurations under the 6.00 kN dynamic load applied over 40 ms was evaluated across 12 discrete solution sub-steps. Fig. 2 presents the progressive lateral intrusion displacement recorded at each time increment.

Table 1. Material Properties[2]

Parameter / Mechanical Property	Stainless Steel (Baseline)	CFRP Prepreg (Reinforcement)
Constitutive Material Model	Isotropic / Elasto-plastic	Orthotropic / Elastic
Mass Density, ρ (kg/m³)	7850	1550
Longitudinal Modulus, E_{11} (GPa)	200	135
Transverse Modulus, E_{22} (GPa)	200	9.5
In-plane Shear Modulus, G12 (GPa)	76.9	4.8
Major Poisson's Ratio, ν / ν_{12}	0.3	0.32
Tensile Yield Strength, σ_y (MPa)	250	—
Compressive Yield Strength, σ_{yc} (MPa)	250	—
Tensile Ultimate Strength, σ_u (MPa)	460	1850
Nominal Sheet Thickness (mm)	2	1.3



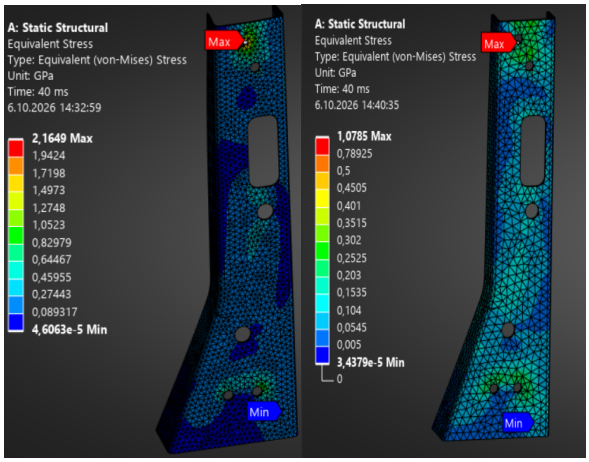
(a) (b)

Fig. 2. Deformation Result. (a is Stainless Steel, b is hybrid)

Conversely, the hybrid B-pillar reinforced with the 1.30 mm CFRP prepreg patch demonstrated a marked enhancement in flexural rigidity. Under the identical 6.00 kN lateral pulse, the peak intrusion decreased to 2.09 mm. This behaviour represents an absolute displacement reduction of 3.22 mm, corresponding to a 60.54% mitigation in cabin intrusion.

Stress concentrations during transverse impact represent critical indicators for potential crack initiation and structural failure, particularly in regions surrounding fixed mounting apertures. Fig. 3 illustrates the comparative equivalent (von Mises) stress contour fields for both models at the peak loading state.

ISBN: 978-605-80193-6-2



(a) (b)

Fig. 3. Equivalent von-Mises Stress Result. (a is Stainless Steel, b is hybrid)

In the baseline 2.00 mm steel component, high stress concentrations were observed around the perimeter of the fixed mounting holes and along the load application zone, reaching a peak equivalent von Mises stress of 2.16 GPa. This stress concentration substantially exceeds the initial yield strength of the material, driving extensive plastic deformation across the metallic web.

With the incorporation of the localized 1.30 mm CFRP prepreg laminate, a substantial load transfer occurred from the ductile steel substrate into the high-strength composite plies. Consequently, the peak equivalent von Mises stress in the steel substrate dropped to 1.08 GPa, achieving an effective stress reduction of 50.18%. Furthermore, the maximum equivalent plastic strain in the metallic structure was successfully restricted to 0.95%. This value remains comfortably within the safe strain hardening boundary of the material, preventing localized ductile tearing and preserving structural integrity throughout the 40 ms impact event.

With a mass penalty of only 0.126 kg (15.51%), the localized CFRP laminate reinforcement provided a 60.54% reduction in cabin intrusion and a 50.18% decrease in structural stress concentrations.

IV. CONCLUSION

This study numerically evaluated the crashworthiness enhancement of an automotive B-pillar reinforcement steel through localized CFRP prepreg reinforcement under dynamic lateral impact loading. The primary conclusions are as follows:

1. Applying a 1.30 mm localized CFRP prepreg laminate onto the 2.00 mm structural steel B-pillar reinforcement reduced the maximum lateral cabin intrusion by 60.54% (from 5.31 mm to 2.09 mm) under a 6.00 kN transverse load applied over 40 ms.

2. The high-modulus composite plies effectively redistributed the structural load path across the B-pillar reinforcement steel, decreasing the peak equivalent von Mises stress around critical mounting holes by 50.18% (from 2.16 GPa to 1.08 GPa) and restricting the maximum plastic strain to a safe threshold of 0.95%.

3. These structural improvements were achieved with a marginal mass increase of only 0.126 kg (15.51%), demonstrating that targeted composite patching on the B-pillar reinforcement steel provides superior specific stiffness and crash protection compared to conventional uniform sheet gauge thickening.

References

[1] Sun, G., Tian, X., Fang, J., Xiao, X., & Li, Q. (2019). Crashworthiness behavior of CFRP/aluminum hybrid structures under transverse loading. International Journal of Mechanical Sciences, 161, 105027.

[2] Belingardi, G., Beyene, A. T., & Koricho, E. G. (2020). Lightweight design of an automotive B-pillar: A comparative study of metallic and composite material solutions. International Journal of Automotive Technology, 21(2), 341–352.

[3] MAT_024 Piecewise Linear Plasticity (Steel) Material Model Documentation, LS-DYNA Keyword User Manual.

ISBN: 978-605-80193-6-2

Designing Soft Body Armor for End of Life: Reliability-Based Selection of Fiber and Areal Density Under Aging

Muhammed Zahid Koç[1*]

[1]Karabük University, Faculty of Engineering, Department of Industrial Engineering, Karabük, Türkiye

* Corresponding author e-mail: mzkoc@karabuk.edu.tr

***Abstract*—Soft body armor is selected and certified on the basis of its performance when new, yet the protection that matters to the wearer is the protection that remains at the end of service. This study asks whether the fiber that gives the lightest armor when new still gives the lightest armor after aging. A transparent model chain was built entirely from published data of the National Institute of Standards and Technology. The Phoenix and Porwal membrane model links yarn strength and failure strain to the ballistic limit, a logistic curve converts the ballistic limit into a perforation probability at a threat velocity, and exponential fits to accelerated aging data describe how yarn properties decay. Fiber and areal density selection was then posed as a chance-constrained, bi-objective design problem. The implementation reproduced published ballistic limit retention values within 0.003 for eleven of fourteen tabulated cases. For a 9 mm threat at 350 meters per second and a perforation risk of one percent, armor made of poly(p-phenylene-2,6-benzobisoxazole) (PBO) needed 22 percent less areal density than aramid armor when new. Along the aging axis, however, the weight versus aging Pareto front switched from PBO to aramid after 2.9 days of accelerated aging in the model-based scenario and after 4.7 days when degradation was anchored to the measured loss of ballistic limit. At the switching point the PBO yarn had lost only 20 to 28 percent of its strength, less than the average loss observed in field-aged PBO armor. Absolute areal densities changed by up to 23 percent under calibration and scatter assumptions, whereas the switching point moved by less than 0.15 days. Fiber selection for soft armor should therefore rest on end-of-life reliability rather than on the new-condition ballistic limit, and coupon-based surveillance can use the switching point as a replacement trigger.**

***Keywords*—Soft body armor, fiber aging, chance-constrained optimization, stress-strength interference, condition-based replacement, multi-objective material selection, ballistic limit**

I. INTRODUCTION

Soft body armor is built from woven or unidirectional layers of high-performance fibers such as aramid, poly(p-phenylene-2,6-benzobisoxazole) (PBO) and ultra-high-molecular-weight polyethylene (UHMWPE) [1]. Armor models are certified when new, and candidate designs are usually compared through their ballistic limit V_{50}, the impact velocity at which half of the impacts perforate the panel [2], [3]. Two features of this practice are at odds with the protection a wearer actually receives. First, the ballistic limit is a median, whereas safety depends on the lower tail of the perforation curve at the threat velocity. Second, high-performance fibers lose strength in service under heat, humidity and mechanical wear [4], [5], [6]. A widely reported field failure of a PBO vest in 2003 showed that this loss can defeat an armor that met its certification when new [7], and it led to a major revision of the NIJ standard in 2008 [8]. Even so, armor is still commonly replaced on a fixed five-year schedule rather than on its measured condition [8].

The National Institute of Standards and Technology (NIST) addressed the degradation issue by relating the tensile properties of yarns extracted from laboratory-aged and field-aged armor to measured ballistic limits through two theoretical models, the dimensionless analysis of Cunniff [9] and the membrane model of Phoenix and Porwal [10], and proposed coupon-based surveillance of fielded armor [7]. That work was diagnostic: it estimated how much an existing armor had degraded. The design question that follows from it has, to the best of the author's knowledge, not been posed explicitly: if the fiber and the areal density are chosen for the end of service rather than for the new condition, does the ranking of fibers change, and when?

Reliability-based design optimization handles this type of question in structural engineering, where the probability of failure of a degrading component must stay below a target over its service life [11], [12]. Viewed through this lens, soft armor is a stress-strength interference problem: the strength, the ballistic limit of the panel, decays with age, whereas the stress, the threat velocity, does not. Choosing the lightest panel whose perforation probability stays below a target is then a chance-constrained optimization problem in the sense of Charnes and Cooper [13], and ranking fibers by a single property index follows the materials-selection logic of engineering design [14], to which the present study adds the dimension of time. This paper transfers that view to textile armor. Its contributions are threefold: (i) a transparent model chain, built entirely from published NIST data, that maps yarn property retention to the perforation probability at a threat velocity; (ii) a weight versus aging Pareto front for two fibers that exposes a switching point t^* beyond which the fiber that is lighter when new is no longer the lighter choice; and (iii) a sensitivity analysis showing that t^* is nearly invariant to the calibration and scatter assumptions that dominate the absolute predictions.

II. BACKGROUND

A. Ballistic Limit Models

Cunniff [9] showed that the ballistic limits of fabric systems collapse onto a master curve when they are normalized by $\Omega^{1/3}$, where Ω is a fiber property group that combines specific toughness and longitudinal wave speed, and plotted against the ratio of armor areal density to projectile areal density. Phoenix and Porwal

ISBN: 978-605-80193-6-2

[10] derived this master curve from a membrane model of the impacted multi-ply system and obtained an explicit expression that contains a model parameter θ, typically between 1 and 2 [7]. Porwal and Phoenix analyzed system effects in multi-layer panels [15] and extended the analysis to two-layer hybrid panels, showing that stacking order affects the ballistic limit [16]. Cheeseman and Bogetti [1] reviewed the material and test parameters that govern ballistic impact into fabrics. The NIST study [7] used both models in ratio form to predict the ballistic limit retention of aged armor made of poly(p-phenylene terephthalamide) (PPTA, aramid) and PBO from extracted-yarn data, and found that both models overestimated the measured loss, that is, they were conservative.

B. Fiber Degradation Mechanisms

The three fiber families degrade by different mechanisms and at different rates. The benzoxazole ring of PBO is susceptible to hydrolysis, particularly under acidic conditions [4]. PBO armor panels exposed to elevated temperature and humidity lost tensile strength and showed marked chemical changes, whereas PBO fibers sealed in dry argon at elevated temperature lost less than 4 % of their strength after 30 weeks, which identified moisture as the key driver [5]; the hydrolytic stability of complete PBO and PPTA armor was examined in [17]. Aramid is a condensation polymer whose amide linkage can also hydrolyze, yet p-aramid yarns retained 86 % to 88 % of their strength after about one year at 70 °C and 76 % relative humidity, and the Cunniff and Phoenix and Porwal models predicted less than 10 % loss of ballistic limit for them [8]. UHMWPE fibers degrade mainly by a thermo-oxidative mechanism [6], and aged flexible unidirectional laminates used in body armor have been characterized by tensile testing [18]. These contrasts in mechanism and rate are the reason why a ranking of fibers in the new condition need not hold after aging.

III. METHODS

A. Ballistic Limit Model

For a fiber with tensile strength σ, failure strain ε, modulus $E = \sigma/\varepsilon$ and density ρ, and for an armor of areal density A_d impacted by a projectile of mass m_p and presented area A_p, the fiber property group and the areal density ratio are

$$\Omega = \frac{\sigma\varepsilon}{2\rho}\left(\frac{E}{\rho}\right)^{1/2}, \quad \Gamma_0 = \frac{A_d A_p}{m_p} \tag{1}$$

The Phoenix and Porwal ballistic limit [7], [10] is

$$V_{50} = \kappa\,\Omega^{1/3}\,\frac{2^{1/3}\,\varepsilon^{1/12}\left(1+\theta^2\Gamma_0\right)}{K_{\max}^{3/4}} \tag{2}$$

$$K_{\max} = \exp\left[-\frac{4\theta^2\Gamma_0(\psi^2-1)}{3(1+\theta^2\Gamma_0)}\right]\psi^{1/2}\left[\frac{z}{\ln(1+z)}\right]^{2/3} \tag{3}$$

$$z = \left(\frac{\psi}{\varepsilon}\right)^{1/2}(\psi-1), \quad \psi = \left(\frac{1+\theta^2\Gamma_0}{2\theta^2\Gamma_0}\right)^{1/2} \tag{4}$$

where ψ is the maximum membrane deflection parameter and κ is an absolute calibration factor introduced in this study. The implementation was verified by recomputing the Cunniff and Phoenix and Porwal retention values that NIST tabulated for 14 lot and aging-time combinations [7]: for 11 combinations both values agreed within 0.003, and the mean absolute differences were 0.004 and 0.008, respectively. The largest deviation, 0.032, occurred for lot A3, whose new yarns showed an unusually wide failure-strain distribution.

B. Calibration and Perforation Probability

With θ = 1.5 the uncalibrated model predicted about half of the measured ballistic limit of the new PBO armor in [7] (A_d = 2.273 kg/m^2, V_{50} = 458 m/s), in line with the NIST recommendation to use the models in relative rather than absolute form. A single factor κ = 2.02 was therefore fitted to this point and applied to both fibers. As a consistency check, the calibrated model gave 435 m/s for the new PPTA armor of [7] (A_d = 2.625 kg/m^2), whereas NIST reported a slightly higher ballistic limit for this armor than for the new PBO armor. The calibrated model therefore slightly underestimates PPTA, which biases every comparison below in favor of PBO.

The probability that an impact at velocity v perforates the panel was described by the logistic model that NIJ Standard-0101.06 recommends for ballistic limit analysis; probit and complementary log-log models are alternatives that may fit some armor systems better [19]:

$$P_f(v) = \frac{1}{1+\exp[-(v-V_{50})/s]} \tag{5}$$

The scale s = 13.5 m/s was back-calculated from the new and aged PBO perforation curves reported in [7], which gave 13.8 and 12.9 m/s. The threat was the 8.0 g, 9 mm full metal jacket projectile of [7] at v_{th} = 350 m/s, the upper end of the fair-hit velocity interval used there.

C. Aging Scenarios

Strength and failure-strain retention of PPTA (lot A4) and PBO (lot B4), both aged at 70 °C and 90 % relative humidity for up to 13 days [7], were fitted by

$$r(t) = r_\infty + (1-r_\infty)\exp(-kt) \tag{6}$$

Single and double exponential models have been applied to aramid aging data, but they fitted those data only loosely and are not recommended for extrapolation [8]; the fits here were therefore used only within the tested range of 0 to 13 days. Table I lists all inputs and fitted parameters. Aging time t is reported in days of this accelerated protocol, and no conversion to field years was attempted. Two scenarios were evaluated. In the model-based scenario the retained yarn properties were inserted into (1) to (4). Because the models proved conservative in [7], an anchored scenario scaled the predicted loss of ballistic limit by the ratio f of measured to predicted loss at the end of aging:

$$V_{50}^{a}(t) = V_{50}(0)\left[1 - f\left(1 - \frac{V_{50}^{m}(t)}{V_{50}(0)}\right)\right] \tag{7}$$

where superscripts a and m denote the anchored and model-based values. For PBO the measured loss was 15 % to 18 % against a predicted 33 % to 35 %, which gives f = 0.49; for PPTA the measured ballistic limit was essentially unchanged, which gives f = 0. A stress test applied f = 0.49 to PPTA as well.

ISBN: 978-605-80193-6-2

TABLE I. Model Inputs and Fitted Parameters

Quantity	Value
Projectile	9 mm FMJ, m_p = 8.0 g, A_p = 6.36×10^{-5} m^2
Threat velocity, v_{th}	350 m/s
PPTA new yarn, σ / ε / ρ	2.80 GPa / 2.84 % / 1440 kg/m^3
PBO new yarn, σ / ε / ρ	4.02 GPa / 2.69 % / 1540 kg/m^3
Calibration point (PBO)	A_d = 2.273 kg/m^2, V_{50} = 458 m/s
Model parameters	θ = 1.5, κ = 2.02, s = 13.5 m/s
PPTA aging fit, σ / ε	r_∞ = 0.80 / 0.84; k = 0.16 / 0.15 d^{-1}
PBO aging fit, σ / ε	r_∞ = 0.53 / 0.54; k = 0.19 / 0.15 d^{-1}
Anchoring ratio, f	PBO 0.49; PPTA 0 (stress test 0.49)

New-yarn properties are the means of two unaged lots per fiber in [7]; fiber densities are typical manufacturer values.

D. Design Problem

For fiber m and aging time t, the minimum areal density that keeps the perforation probability at or below a target α is

$$A_d^*(m,t) = \min\{A_d : P_f(v_{th};\, m, A_d, t) \le \alpha\} \qquad (8)$$

with α = 0.01 as the base case. Constraint (8) is a chance constraint on the interference between the fixed threat velocity and the degrading ballistic limit. Selecting the fiber and the areal density is a bi-objective problem: minimize A_d and maximize the aging time that the panel tolerates. For each fiber the curve $A_d^*(m, t)$ is its Pareto front, the overall front is the lower envelope of the two curves, and the switching point t^* solves A_d^*(PBO, t^*) = A_d^*(PPTA, t^*). All computations used Python with NumPy and SciPy, and roots were found with Brent's method.

IV. RESULTS

A. New Versus Aged Performance

Fig. 1(a) shows the aging data and the fitted retention curves: after 13 days PBO retained about 0.6 of its strength and PPTA about 0.82. Fig. 1(b) shows the resulting risk versus weight curves. When new, PBO met the 1 % target at 1.85 kg/m^2 and PPTA at 2.36 kg/m^2, a 22 % weight advantage for PBO. After 13 days of aging the ranking reversed in both scenarios: 4.90 versus 3.22 kg/m^2 in the model-based scenario and 2.80 versus 2.36 kg/m^2 in the anchored scenario. A PBO panel sized for the new condition saw its ballistic limit fall from 412 to 333 m/s in the anchored scenario, and its perforation probability at the threat velocity rose from 1 % to 78 %.

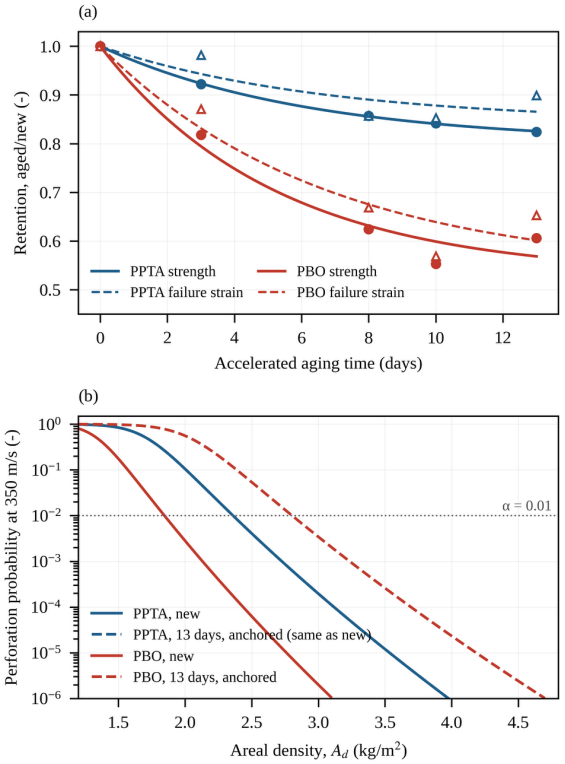


Fig. 1. (a) Retention of yarn strength (filled symbols, solid lines) and failure strain (open symbols, dashed lines) under accelerated aging at 70 °C and 90 % relative humidity, data from [7] with exponential fits. (b) Perforation probability at 350 m/s versus areal density for new armor and after 13 days of aging in the anchored scenario; for PPTA the aged curve coincides with the new one.

B. Weight Versus Aging Front

Fig. 2 plots the minimum areal density against aging time. PBO forms the front only up to t^* = 2.9 days in the model-based scenario and up to t^* = 4.7 days in the anchored scenario; beyond these points PPTA is lighter for the same reliability. In the stress test, where PPTA also lost ballistic limit at the anchoring ratio of PBO, the switch still occurred within the tested range, at 9.7 days. Read on the upper axis of Fig. 2, the switch happened when PBO yarn had retained 0.80 (model-based) or 0.72 (anchored) of its strength. For comparison, the mean strength of yarns extracted from field-aged PBO armor in [7] was 2.64 GPa, about two thirds of the new-yarn strength, so an average fielded PBO vest in that sample had already passed both thresholds.

ISBN: 978-605-80193-6-2

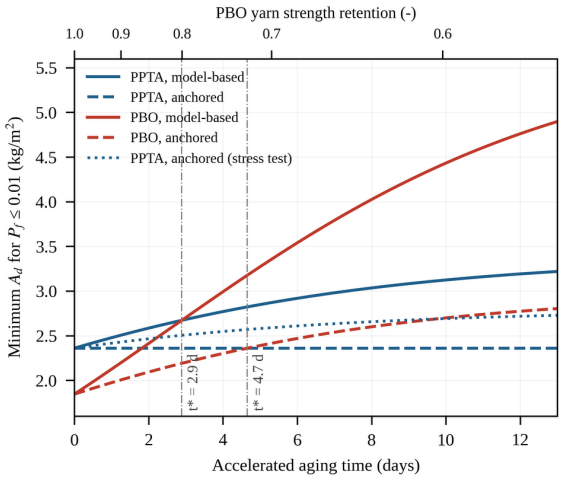


Fig. 2. Weight versus aging front: minimum areal density that keeps the perforation probability at 350 m/s at or below 0.01, as a function of accelerated aging time. Vertical lines mark the switching points t* at which PPTA becomes lighter than PBO. The upper axis gives the corresponding PBO yarn strength retention.

C. Sensitivity

Table II varies the logistic scale s, the model parameter θ, the calibration factor κ and the risk target α. Absolute areal densities moved by up to 23 % across these cases, and the ±10 % change in κ alone shifted them by 17 % to 23 %. The switching point, in contrast, stayed within 2.86 to 2.94 days in the model-based scenario and within 4.61 to 4.75 days in the anchored scenario. The reason is structural: κ scales the ballistic limit of both fibers equally, and s and α only change the ballistic limit that a panel must reach. The switching point therefore depends mainly on the ratio of the fiber property groups and on the aging kinetics, and only weakly on areal density through Γ_0.

TABLE II. Sensitivity of Minimum Areal Density (kg/m²) and Switching Point (days)

Case	PPTA, new	PBO, new	t^*, model	t^*, anchored
Base case	2.36	1.85	2.89	4.65
s = 9 m/s	2.13	1.67	2.88	4.64
s = 18 m/s	2.60	2.03	2.90	4.67
θ = 1.25	2.36	1.85	2.86	4.61
θ = 2.0	2.36	1.84	2.94	4.75
κ − 10 %	2.91	2.27	2.91	4.68
κ + 10 %	1.95	1.53	2.88	4.63
α = 0.001	2.73	2.13	2.90	4.67

V. DISCUSSION

The results support three practical messages. First, a new-condition comparison can rank fibers in the wrong order for the quantity that matters at the end of service: choosing on the ballistic limit when new favors a fiber that becomes heavier for the same end-of-life reliability after moderate aging. Second, the switching point is a relative quantity and inherits the robustness that NIST attributed to the ratio form of the ballistic models [7]; it can therefore be reported with more confidence than any absolute areal density. Third, the switching point translates directly into a surveillance rule: once coupon tests [7] show a PBO yarn strength retention below about 0.72 to 0.80, the panel no longer offers the weight benefit that justified its selection, and replacement or re-certification is warranted. Such a rule turns the fixed five-year replacement interval that remains common practice [8] into a condition-based maintenance policy of the kind used for degrading engineering assets [20]. From a decision-making perspective, t^* acts as a break-even point between an initial weight advantage and a faster loss of strength.

The study has clear limits. The absolute level rests on a single calibration point, and the calibrated model underestimates the new PPTA armor; both effects bias the comparison toward PBO and therefore make the reported reversal conservative. Aging time is expressed in accelerated days because the data provide no validated mapping to field years: strength-decay kinetics fitted to aramid data were found unsuitable for extrapolation [8], and a temperature-only scaling would ignore the role of moisture that dominates PBO degradation [5]. Field-aged yarn strengths in [7] scatter much more than laboratory-aged ones, but this scatter was measured at yarn level and would overstate the variability between panels, so it was not propagated. The models describe single-fiber panels; hybrid panels require the stacking-order analysis of [16]. UHMWPE, which is widely used in current armor, was not included because aging data under the same protocol and a calibration point were not available, although its thermo-oxidative aging has been characterized [6].

VI. CONCLUSION

This paper reframed fiber and areal density selection for soft body armor as a chance-constrained, end-of-life design problem and solved it with a model chain built entirely from published NIST data. PBO gave a 22 % lighter panel than aramid when new, but the weight versus aging front switched to aramid after 2.9 to 4.7 days of accelerated aging, at a PBO strength retention of 0.72 to 0.80, and this switching point was nearly insensitive to the calibration and scatter assumptions. Future work will add UHMWPE, two-layer hybrid stacking and a mapping from accelerated to field aging, and will treat the replacement interval as an explicit decision variable. The analysis code is available from the author upon request.

References

[1] Cheeseman, B. A., & Bogetti, T. A. (2003). Ballistic impact into fabric and compliant composite laminates. *Composite Structures, 61*(1–2), 161–173. https://doi.org/10.1016/S0263-8223(03)00029-1

[2] National Institute of Justice. (2008). *Ballistic resistance of body armor* (NIJ Standard-0101.06). U.S. Department of Justice, Office of Justice Programs, Washington, DC, USA. NCJ 223054. https://nij.ojp.gov/library/publications/ballistic-resistance-body-armor-nij-standard-010106

[3] ASTM International. (2018). *Standard test method for collection of ballistic limit data for ballistic-resistant torso body armor and shoot packs* (ASTM E3110-18). West Conshohocken, PA, USA. https://doi.org/10.1520/E3110-18

[4] Holmes, G. A., Rice, K. D., & Snyder, C. R. (2006). Ballistic fibers: A review of the thermal, ultraviolet and hydrolytic stability of the benzoxazole ring

ISBN: 978-605-80193-6-2

structure. *Journal of Materials Science, 41*, 4105–4116. https://doi.org/10.1007/s10853-005-5597-1

[5] Chin, J., Forster, A., Clerici, C., Sung, L., Oudina, M., & Rice, K. (2007). Temperature and humidity aging of poly(p-phenylene-2,6-benzobisoxazole) fibers: Chemical and physical characterization. *Polymer Degradation and Stability, 92*(7), 1234–1246. https://doi.org/10.1016/j.polymdegradstab.2007.03.030

[6] Forster, A. L., Forster, A. M., Chin, J. W., et al. (2015). Long-term stability of UHMWPE fibers. *Polymer Degradation and Stability, 114*, 45–51. https://doi.org/10.1016/j.polymdegradstab.2015.01.028

[7] Forster, A. L., Leber, D. D., Engelbrecht-Wiggans, A., et al. (2020). Linking theory to practice: Predicting ballistic performance from mechanical properties of aged body armor. *Journal of Research of the National Institute of Standards and Technology, 125*, 125026. https://doi.org/10.6028/jres.125.026

[8] Engelbrecht-Wiggans, A., Burni, F., Guigues, E., et al. (2020). Effects of temperature and humidity on high-strength p-aramid fibers used in body armor. *Textile Research Journal, 90*(21–22), 2428–2440. https://doi.org/10.1177/0040517520918232

[9] Cunniff, P. M. (1999). Dimensionless parameters for optimization of textile-based body armor systems. In *Proceedings of the 18th International Symposium on Ballistics* (Vol. 2, pp. 1303–1310). Technomic Publishing Co.

[10] Phoenix, S. L., & Porwal, P. K. (2003). A new membrane model for the ballistic impact response and V50 performance of multi-ply fibrous systems. *International Journal of Solids and Structures, 40*(24), 6723–6765. https://doi.org/10.1016/S0020-7683(03)00329-9

[11] Aoues, Y., & Chateauneuf, A. (2010). Benchmark study of numerical methods for reliability-based design optimization. *Structural and Multidisciplinary Optimization, 41*(2), 277–294. https://doi.org/10.1007/s00158-009-0412-2

[12] Andrieu-Renaud, C., Sudret, B., & Lemaire, M. (2004). The PHI2 method: A way to compute time-variant reliability. *Reliability Engineering & System Safety, 84*(1), 75–86. https://doi.org/10.1016/j.ress.2003.10.005

[13] Charnes, A., & Cooper, W. W. (1959). Chance-constrained programming. *Management Science, 6*(1), 73–79. https://doi.org/10.1287/mnsc.6.1.73

[14] Ashby, M. F. (2011). *Materials selection in mechanical design* (4th ed.). Butterworth-Heinemann. ISBN 978-1-85617-663-7

[15] Porwal, P. K., & Phoenix, S. L. (2005). Modeling system effects in ballistic impact into multi-layered fibrous materials for soft body armor. *International Journal of Fracture, 135*, 217–249. https://doi.org/10.1007/s10704-005-3993-9

[16] Porwal, P. K., & Phoenix, S. L. (2008). Effects of layer stacking order on the V50 velocity of a two-layered hybrid armor system. *Journal of Mechanics of Materials and Structures, 3*(4), 627–639. https://doi.org/10.2140/jomms.2008.3.627

[17] Forster, A. L., Pintus, P., Messin, G., et al. (2011). Hydrolytic stability of polybenzobisoxazole and polyterephthalamide body armor. *Polymer Degradation and Stability, 96*, 247–254. https://doi.org/10.1016/j.polymdegradstab.2010.10.004

[18] Engelbrecht-Wiggans, A., Burni, F., Krishnamurthy, A., et al. (2020). Tensile testing of aged flexible unidirectional composite laminates for body armor. *Journal of Materials Science, 55*, 1035–1048. https://doi.org/10.1007/s10853-019-04063-w

[19] Mauchant, D., Rice, K. D., Riley, M. A., Leber, D., Samarov, D., & Forster, A. L. (2011). *Analysis of three different regression models to estimate the ballistic performance of new and environmentally conditioned body armor* (NISTIR 7760). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.IR.7760

[20] Jardine, A. K. S., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition-based maintenance. *Mechanical Systems and Signal Processing, 20*(7), 1483–1510. https://doi.org/10.1016/j.ymssp.2005.09.012

 ISBN: 978-605-80193-6-2

Is Biodegradable Greener? Robust Fiber Selection for Single-Use Nonwoven Wipes Under Data and Preference Uncertainty

Muhammed Zahid Koç[1*]

[1]Karabük University, Faculty of Engineering, Department of Industrial Engineering, Karabük, Türkiye

* Corresponding author e-mail: mzkoc@karabuk.edu.tr

***Abstract*—Single-use wet wipes are increasingly marketed as plastic-free or biodegradable, yet fiber choices for these nonwovens are rarely evaluated against production impacts and environmental persistence at the same time. This study ranks seven candidate fibers, namely polypropylene, polyester, polylactic acid, cotton, viscose from integrated and non-integrated production, and lyocell, with stochastic multicriteria acceptability analysis. The decision matrix combines one harmonized life cycle assessment dataset for fossil energy use and net greenhouse gas emissions with aquatic biodegradation evidence for persistence. Data intervals and unknown criterion weights were sampled in 200,000 Monte Carlo runs under three end-of-life scenarios: incineration, leakage to water, and industrial composting. Integrated viscose ranked first in 84 to 100 percent of the runs in every scenario, whereas viscose from non-integrated production fell to fifth place, so the production route mattered more than the fiber family. Polylactic acid tied with cotton when only production impacts counted, but under leakage its holistic acceptability dropped from 0.59 to 0.42 while that of cotton rose from 0.58 to 0.68, because polylactic acid did not degrade in fresh or sea water. Its advantage returned only when industrial composting was assumed. Polypropylene and polyester ranked last in all scenarios. When water use, freshwater ecotoxicity and eutrophication were added for the six fibers with complete data, cotton moved to last place. The biodegradable label is therefore not a reliable proxy for environmental preference: it pays off only when the end-of-life pathway matches the conditions under which the material degrades, and the energy system behind a fiber can outweigh its chemistry.**

***Keywords*—Nonwoven wipes, fiber selection, biodegradability, life cycle assessment, stochastic multicriteria acceptability analysis, Monte Carlo simulation, sustainable procurement, man-made cellulose fibers, polylactic acid**

I. INTRODUCTION

Wet wipes are among the most widely used single-use nonwoven products. Their substrates are made from polypropylene (PP), polyester (PET), polylactic acid (PLA), cotton or man-made cellulose fibers such as viscose and lyocell, often in blends. Because used wipes are frequently flushed or littered, the European Union requires wet wipes that contain plastic to carry a marking on their packaging [1]. Manufacturers have responded with substrates advertised as plastic-free or biodegradable. Whether such claims identify the environmentally preferable fiber is, however, an open question.

Two strands of evidence bear on it but are seldom combined. Life cycle assessment (LCA) quantifies the production impacts of fibers and shows, for example, that man-made cellulose fibers can require far less fossil energy than petrochemical fibers, while cotton carries heavy water and ecotoxicity burdens [2], [3]. Fate studies, on the other hand, show that cellulose microfibers biodegrade in aquatic environments whereas polyester persists [4], and that PLA, despite its compostable label, does not degrade in fresh or sea water [5]. Combining the two requires weighing production impacts against persistence, a judgment on which stakeholders disagree. Even within LCA, the fiber identified as worst changes with the weighting method [6].

Multicriteria methods are well established for environmental supplier and material selection, but a review of green supplier selection found that most applications rely on a single, mostly fuzzy, technique such as the analytic hierarchy process, that criterion weights are often set arbitrarily and subjectively, and that their influence is rarely tested by sensitivity analysis [7]. This paper addresses fiber selection for wipe substrates as a multicriteria decision under both data and preference uncertainty. Instead of eliciting weights, it explores the entire weight space, so that the sensitivity analysis called for in [7] becomes the result itself. Its contributions are: (i) a decision matrix that draws production impacts from one harmonized LCA dataset and persistence from aquatic biodegradation evidence; (ii) a stochastic multicriteria acceptability analysis (SMAA) [8], [9] that requires no weight elicitation; and (iii) three end-of-life scenarios that show when, and only when, biodegradability changes the preferred fiber.

II. METHODS

A. Alternatives, Functional Unit and Data

Seven alternatives were compared: PP, PET, PLA, cotton, viscose from non-integrated production in Asia, viscose from integrated pulp and fiber production in Austria, and lyocell. The functional unit was one tonne of staple fiber, which assumes that equal fiber mass delivers equal wipe function, a screening-level simplification. All production data were taken from the ISO-conformant study of Shen et al. [3], [10] and the associated thesis [11], which assess all fibers with the same system boundary and methods. This consistency is the main reason for relying on a single source rather than on values compiled from heterogeneous studies.

Three criteria, all to be minimized, were available for every alternative. Non-renewable energy use (NREU) is a robust proxy for many environmental impacts [12]. Net global warming potential (GWP) covers production plus incineration of the used product:

ISBN: 978-605-80193-6-2

$$\mathrm{GWP_{net}} = \mathrm{GWP_{gate}} + C_{\mathrm{bio}} - E_{\mathrm{rec}} \quad (1)$$

where $\mathrm{GWP_{gate}}$ is the cradle-to-gate value, which credits biogenic carbon uptake for bio-based fibers, C_{bio} is the carbon embedded in the fiber and released at incineration, and E_{rec} is the emission avoided by energy recovery. The lower bound of each interval assumes energy recovery and the upper bound none [11]. For PLA, which the source covers only as a polymer, the fiber value was built from the polymer data of two producers (0.50 to 1.30 kg CO_2-eq/kg; 31 to 42 MJ/kg), a spinning increment of 0.76 kg CO_2-eq/kg and 14 MJ/kg derived from the PET polymer and fiber data of the same source, the stoichiometric embedded carbon of 1.83 kg CO_2/kg, and the energy recovery credit of the cellulose fibers. The third criterion, persistence, is the fraction of fiber mass expected to remain after release to water. Cotton and rayon microfibers degraded in aquatic tests while polyester hardly degraded [4], and PLA did not degrade in sea or fresh water over one year at 25 °C [5]. Rayon degraded more slowly than cotton [4], so its interval was shifted upward, and lyocell, also a regenerated cellulose, was assigned the same interval. Table I lists the resulting decision matrix. Point values were given an uncertainty of ±10 %.

TABLE I. Decision Matrix (per Tonne of Staple Fiber)

Fiber	NREU (GJ)	Net GWP (t CO_2-eq)	Persistence (-)
PP	79–97	4.25–5.90	1.00
PET	81–95	4.30–6.35	0.94–1.00
PLA	45–56	2.57–3.89	0.95–1.00
Cotton	36–60	3.13–3.70	0.08–0.52
Viscose, non-integrated	55–67	4.74–5.26	0.15–0.60
Viscose, integrated	17–21	0.69–1.20	0.15–0.60
Lyocell	21–42	0.97–2.53	0.15–0.60

Sources: NREU and net GWP from [3], [11], PLA derived as described in the text; persistence from [4], [5].

B. Stochastic Multicriteria Acceptability Analysis

Criterion values were normalized to partial values between 0 (worst) and 1 (best) over the range of all alternatives,

$$u_{ij} = \frac{x_j^{\max} - x_{ij}}{x_j^{\max} - x_j^{\min}} \quad (2)$$

and aggregated with an additive value function

$$v_i(\xi, w) = \sum_{j=1}^{n} w_j \, u_{ij}(\xi), \quad w_j \geq 0, \quad \sum_{j=1}^{n} w_j = 1 \quad (3)$$

where ξ denotes a realization of the uncertain criterion values and w the weight vector. Because no weights were elicited, w was sampled uniformly from the weight simplex, and criterion values were sampled uniformly from their intervals. SMAA-2 [9] reports the rank acceptability index, the share of the joint weight and data space that gives alternative i rank r,

$$b_i^r = \int_{\Xi} f_X(\xi) \int_{w_i^r(\xi)} f_W(w) \, \mathrm{d}w \, \mathrm{d}\xi \quad (4)$$

and summarizes it in the holistic acceptability index with linear meta-weights,

$$a_i^h = \sum_{r=1}^{m} \frac{m-r}{m-1} \, b_i^r \quad (5)$$

which equals 1 for an alternative that always ranks first and 0 for one that always ranks last. Each analysis used 200,000 Monte Carlo samples, implemented in Python with NumPy.

C. End-of-Life Scenarios

Persistence matters only if the fiber reaches the environment. Scenario S1 (managed end of life) assumes collection and incineration and uses NREU and net GWP. Scenario S2 (leakage) adds persistence with the aquatic intervals of Table I, representing flushed or littered wipes. Scenario S3 (matched infrastructure) assumes that compostable fibers, that is PLA and the cellulose fibers, reach industrial composting, where their persistence was set to 0 to 0.10. As a robustness check, scenarios E1 and E2 repeated S1 and S2 for the six fibers with complete data on three further criteria: process and irrigation water, freshwater aquatic ecotoxicity and eutrophication [3]. PLA was excluded from E1 and E2 because these data were not available for it in the same dataset.

III. RESULTS

A. Production Impacts Only

In S1, integrated viscose ranked first in 99.7 % of the samples and lyocell second in 99 % (Fig. 1). PLA and cotton were practically tied, with holistic acceptability indices of 0.585 and 0.581, and shared ranks three and four. Non-integrated viscose ranked fifth in 89 % of the samples, while PP and PET occupied the last two places. The two viscose routes therefore differed more in acceptability (0.999 versus 0.312) than most pairs of different fibers: the energy system behind the fiber mattered more than its chemistry.

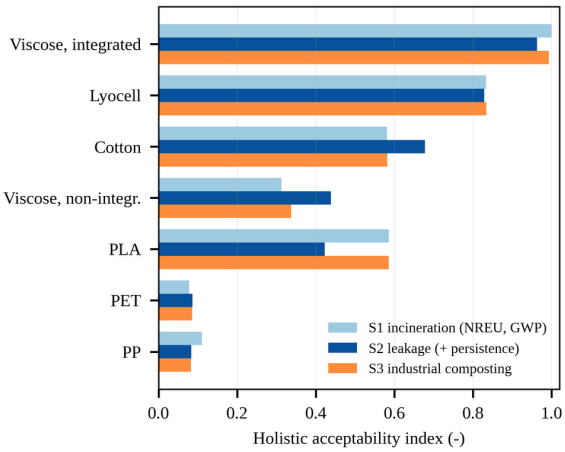


Fig. 1. Holistic acceptability index of the seven fibers under the three end-of-life scenarios. Fibers are ordered by their index in the leakage scenario S2.

B. Leakage and Composting

Adding persistence in S2 reshuffled the middle of the ranking. Cotton rose to third place in 74 % of the samples and its index increased to 0.677, whereas PLA fell to fourth or fifth place with an index of 0.423 (Fig. 2). PLA thus lost its tie with cotton precisely in the scenario in which biodegradability is supposed to help, because the evidence shows no degradation of PLA in water [5]. Non-integrated viscose gained from its degradability and moved level with PLA. Integrated viscose kept the first rank in 84 % of the samples,

ISBN: 978-605-80193-6-2

followed by lyocell with 11 % and cotton with 4 %. PP and PET remained last. In S3, where industrial composting was assumed for the compostable fibers, PLA recovered its index of 0.585 and again tied with cotton (Fig. 1). The value of biodegradability is therefore conditional on an end-of-life pathway that matches the conditions under which the material degrades.

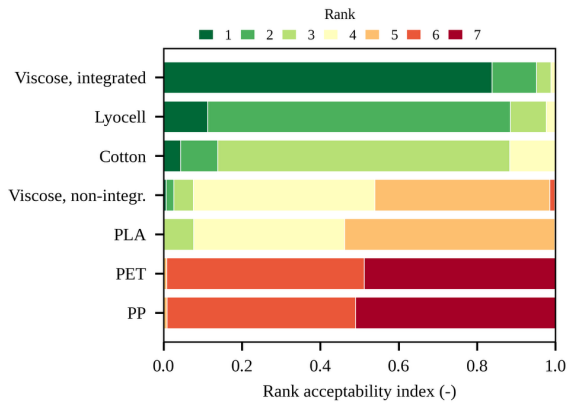


Fig. 2. Rank acceptability indices in the leakage scenario S2: the share of the joint weight and data space in which each fiber attains each rank.

C. Robustness to the Criterion Set

Table II shows the extended analysis. With water use, freshwater ecotoxicity and eutrophication added, cotton fell to last place in 88 % (E1) and 72 % (E2) of the samples, behind PP and PET, because of its irrigation water and the ecotoxicity of its pesticides [3]. The top of the ranking did not change: integrated viscose remained first in 99.8 % (E1) and 88 % (E2) of the samples, and lyocell second. The position of cotton therefore depends on which impacts are counted, whereas the preference for integrated man-made cellulose fibers is robust to both the weights and the criterion set.

TABLE II. Extended Criteria and Holistic Acceptability Indices (PLA Excluded)

Fiber	Water (m^3/t)	Ecotox. (kg 1,4-DB-eq/t)	Eutroph. (kg PO_4-eq/t)	E1	E2
PP	≤ 2	53	1.0	0.34	0.26
PET	≤ 5	58	1.2	0.28	0.25
Cotton	4305–6865	17,310	22	0.05	0.11
Viscose, non-int.	11	160	2.3	0.53	0.59
Viscose, int.	42	74	1.2	1.00	0.97
Lyocell	20	75–85	1.8–1.9	0.80	0.82

IV. DISCUSSION

Three practical messages follow. First, procurement specifications for wipe substrates should name the production route rather than the fiber family: viscose from integrated production with renewable process energy and viscose from fossil-fueled non-integrated production sit at opposite ends of the ranking. Fiber selection thus becomes, in part, a supplier selection problem, in which the energy system of the producer is a decisive attribute. Second, a biodegradable or compostable label is not a reliable proxy for environmental preference. For PLA, the benefit materializes only in industrial composting, which flushed or littered wipes never reach; cellulose fibers, by contrast, keep their advantage under leakage. Third, SMAA identifies where the decision is robust and where it hinges on preferences: the first rank of integrated viscose survives every weight vector, scenario and criterion set examined, whereas the order of PLA, cotton and non-integrated viscose depends on how persistence and water are valued. From a decision-analysis perspective, these ties mark where additional information is worth most: persistence data for regenerated cellulose fibers and complete LCA data for PLA would sharpen the ranking more than further refinement of the already dominant alternatives.

The study has limits. The mass-based functional unit ignores differences in fiber density and substrate basis weight. The production data date from the 2000s, and the toxicity results of cotton rest on US practice and on characterization factors with known uncertainty [3]. Persistence was represented by a single index from laboratory tests, and lyocell was assumed to behave like rayon. Blends, binders and landfill methane were not modeled, and PLA could not be assessed on water, ecotoxicity and eutrophication. Because most commercial wipes are blends, extending the model to blend ratios is the natural next step, in line with the materials-selection logic of engineering design [13].

V. CONCLUSION

Ranking seven candidate fibers for single-use wipe substrates with SMAA showed that integrated man-made cellulose fibers are the robust first choice, that petrochemical fibers rank last, and that the production route of a fiber can matter more than its chemistry. Biodegradability changed the ranking only where the end-of-life pathway allowed it to act: PLA lost its tie with cotton under leakage and regained it only under industrial composting. Environmental claims for wipes should therefore be tied to the actual fate of the product rather than to the label of the material. The analysis code is available from the author upon request.

References

[1] European Parliament and Council. (2019). Directive (EU) 2019/904 of 5 June 2019 on the reduction of the impact of certain plastic products on the environment. *Official Journal of the European Union, L 155*, 1–19. http://data.europa.eu/eli/dir/2019/904/oj

[2] Shen, L., & Patel, M. K. (2008). Life cycle assessment of polysaccharide materials: A review. *Journal of Polymers and the Environment, 16*(2), 154–167. https://doi.org/10.1007/s10924-008-0092-9

[3] Shen, L., Worrell, E., & Patel, M. K. (2010). Environmental impact assessment of man-made cellulose fibres. *Resources, Conservation and Recycling, 55*(2), 260–274. https://doi.org/10.1016/j.resconrec.2010.10.001

[4] Zambrano, M. C., Pawlak, J. J., Daystar, J., Ankeny, M., Cheng, J. J., & Venditti, R. A. (2019). Microfibers generated from the laundering of cotton, rayon and polyester based fabrics and their aquatic biodegradation. *Marine Pollution Bulletin, 142*,

ISBN: 978-605-80193-6-2

394–407. https://doi.org/10.1016/j.marpolbul.2019.02.062

[5] Bagheri, A. R., Laforsch, C., Greiner, A., & Agarwal, S. (2017). Fate of so-called biodegradable polymers in seawater and freshwater. *Global Challenges, 1*(4), 1700048. https://doi.org/10.1002/gch2.201700048

[6] Shen, L., & Patel, M. K. (2010). LCA single score analysis of man-made cellulose fibres. *Lenzinger Berichte, 88*, 60–66. https://dspace.library.uu.nl/handle/1874/203523

[7] Govindan, K., Rajendran, S., Sarkis, J., & Murugesan, P. (2015). Multi criteria decision making approaches for green supplier evaluation and selection: A literature review. *Journal of Cleaner Production, 98*, 66–83. https://doi.org/10.1016/j.jclepro.2013.06.046

[8] Lahdelma, R., Hokkanen, J., & Salminen, P. (1998). SMAA – Stochastic multiobjective acceptability analysis. *European Journal of Operational Research, 106*(1), 137–143. https://doi.org/10.1016/S0377-2217(97)00163-X

[9] Lahdelma, R., & Salminen, P. (2001). SMAA-2: Stochastic multicriteria acceptability analysis for group decision making. *Operations Research, 49*(3), 444–454. https://doi.org/10.1287/opre.49.3.444.11220

[10] International Organization for Standardization. (2006). *ISO 14044:2006 Environmental management – Life cycle assessment – Requirements and guidelines*. Geneva, Switzerland.

[11] Shen, L. (2011). *Bio-based and recycled polymers for cleaner production: An assessment of plastics and fibres* (Doctoral dissertation). Utrecht University, Utrecht, The Netherlands. https://dspace.library.uu.nl/handle/1874/202398

[12] Huijbregts, M. A. J., Rombouts, L. J. A., Hellweg, S., Frischknecht, R., Hendriks, A. J., van de Meent, D., et al. (2006). Is cumulative fossil energy demand a useful indicator for the environmental performance of products? *Environmental Science & Technology, 40*(3), 641–648. https://doi.org/10.1021/es051689g

[13] Ashby, M. F. (2011). *Materials selection in mechanical design* (4th ed.). Butterworth-Heinemann. ISBN 978-1-85617-663-7

 ISBN: 978-605-80193-6-2

Integrating Environmental Performance into Hotel Pricing and Financial Outcomes: A Comparative Review of Quantitative and Analytical Studies Published from 2021 to 2026

Sarper SERİNTÜRK[1] Pırıl TEKİN[1*] Tolunay GÖÇKEN[1]

[1]Adana Alparslan Türkeş Science and Technology University, Industrial Engineering Department, Adana, Türkiye

* Corresponding author e-mail: ptekin@atu.edu.tr

***Abstract*—This comparative review examines how selected hotel studies published from 2021 to 2026 connect sustainability-related measures with pricing and economic outcomes. The objective is to identify which components occur within the same research design and which studies address environmental and economic objectives through a common analytical decision structure. The selected evidence consists of eleven quantitative or analytical studies and one teaching case that is discussed separately. An exploratory Sustainable Revenue Management Integration Score records five binary components covering environmental or broader sustainability measures, economic outcomes, direct pricing or capacity outcomes or decisions, demand or customer behavior, and joint environmental and economic decision modeling. Across the eleven scored studies, the mean score is 2.91 out of five. Four studies directly examine pricing or capacity, six include demand or customer behavior, and only one meets the joint decision criterion. Certification, aggregate environmental, social, and governance measures, perceived practices, and physical resource indicators represent distinct evidence and require separate interpretation. Stated willingness to pay also differs from realized room prices and bookings. Results describe this comparison set. The score describes component coverage rather than study quality, effect size, or operational integration. A promising research direction is an explainable daily pricing framework that combines demand and financial outcomes with measured resource use per occupied room-night.**

***Keywords*—Hotel revenue management, sustainability, dynamic pricing, environmental performance, optimization**

I. INTRODUCTION

Hotel revenue management uses demand and room capacity to inform pricing. Sustainability research examines environmental relationships with economic outcomes. Such relationships do not themselves provide daily pricing rules.

Reference [1] reports approximately 6% higher revenue in computer experiments relative to historical fixed pricing, without environmental indicators. Other studies examine willingness to pay [2, 3], reputation and competitiveness [4, 5], financial outcomes [6, 7], and certification [8, 9, 10].

Building on [11], this review asks which environmental, economic, pricing, and demand components the selected studies include, and which address environmental and economic objectives jointly. It distinguishes observed prices, stated intentions, and inferred premiums.

II. METHOD

A. Search and Selection

This structured comparative review covers a selected evidence set. The search took place on September 2026 across publisher platforms, open academic indexes, and bibliographic services. Hotel and hospitality terms were combined with pricing, revenue management, financial performance, environmental performance, resource use, carbon, environmental, social, and governance (ESG) indicators, certification, and optimization. Reference lists and forward citations were examined.

Eligible studies were peer-reviewed, hotel-focused, and published from 2021 to the search date. They quantitatively or analytically examined sustainability-related economic outcomes or provided a pricing comparator. Tourism-wide, conceptual, duplicate, and bibliographically unverifiable records were excluded from scoring. Eleven studies are scored. Reference [12] provides a contextual teaching case. Database-specific searches, screening counts, and independent-coder agreement are not reported, limiting reproducibility.

B. Coding and Interpretation

Five binary components describe each scored study. E denotes a physical environmental indicator, certification, structured practice scale, or aggregate environmental, social, and governance (ESG) measure containing an environmental dimension. F denotes an economic outcome or intention, including revenue, profit, costs, financial-market outcomes, or willingness to pay. P denotes a directly analyzed realized room-price or capacity-allocation outcome, or an explicit pricing or capacity decision. D denotes demand, occupancy as a demand proxy, bookings, or customer behavioral intentions. O denotes joint environmental-economic optimization or explicit trade-offs within a common analytical decision structure. Narrative recommendations do not qualify.

The exploratory Sustainable Revenue Management Integration Score (SRMIS) equals E + F + P + D + O, from zero to five. Equal weights provide a transparent count, not calibrated importance weights. No validated integration thresholds are assumed. Five means coverage of all components. Component profiles, especially E and O, accompany totals: a study can score three without environmental measurement. Stated willingness to pay, inferred premiums, and descriptive room stock used only as an input or control do not qualify for P.

ISBN: 978-605-80193-6-2

III. RESULTS

A. Research Designs and Findings

Table I summarizes the research designs, measures, and reported findings of the eleven scored studies.

TABLE I. Research Designs, Measures, and Findings of the Scored Studies

Study	Method and data	Measures and outcomes	Reported finding and interpretation
Bandalouski et al. [1]	Dynamic optimization and computer experiments using hotel data	No environmental indicator; revenue and room prices	Approximately 6% higher average revenue than historical fixed-price revenue in model-based experiments.
Park et al. [2]	Customer survey and structural modeling	Perceived water conservation and waste reduction; willingness to pay and revisit intention	Personal and social norms help explain stated willingness to pay more and revisit intentions.
Becerra-Vicario et al. [4]	Empirical relationship model	Environmental performance, reputation, and financial performance	The study examines reputation within a moderated mediation framework linking environmental and financial performance.
Perramon et al. [5]	Empirical analysis of 148 hotel managers in Catalonia	Environmental practices, service quality, competitiveness, and financial performance	The analysis examines joint relationships of service quality and environmental practices with business performance.
Chen et al. [6]	Autoregressive jump intensity trend model	Aggregate ESG commitment; stock returns and volatility	The analysis links ESG commitment with financial resilience during pandemic-related market disruption.
Bianco et al. [8]	Competitive-set analysis and random-effects models; 251 certified Florida hotels	Sustainability certification; ADR, occupancy, and RevPAR	Certified hotels show higher performance relative to competitive sets. ADR is a direct price outcome and occupancy is a demand proxy.
Yang et al. [3]	Customer survey and PLS-SEM	Perceived green image and practices; satisfaction, loyalty, and willingness to pay	The study examines relationships between environmental perceptions, customer responses, and stated premium-payment willingness.
He et al. [7]	Two-way fixed-effects regression	Aggregate ESG and its dimensions; cost of debt	Higher overall ESG performance is associated with lower debt costs. Overall ESG results do not isolate an environmental effect.
Lin et al. [13]	Analytical pricing and emission-reduction model	CO2 reduction and customer participation; profit, operating costs, and prices	The model identifies conditions supporting lower emissions and higher profit within a common decision structure.
Arbelo et al. [9]	Bayesian stochastic frontier model; 79 Canary Islands hotels, 2010-2021	Environmental certification; operating revenue and revenue efficiency	Certification-related revenue-efficiency effects vary across hotels. Premium-pricing potential is inferred rather than directly estimated as a room-rate outcome.
Abdelrazek et al. [10]	Two-way fixed-effects panel; 55,753 booking transactions, 2012-2018	Third-party eco-certification; sales, realized room prices, and booking duration	Certified periods show approximately 43% higher sales in two comparable hotels, alongside higher prices and longer booking durations.

Notes: ADR = average daily rate; RevPAR = revenue per available room; PLS-SEM = partial least squares structural equation modeling. Publication years differ from underlying data periods. The findings are study-specific and are not pooled causal effects.

Reference [12], first published online in 2024, provides a contextual teaching case on environmental measurement, benchmarking, and revenue management. It is discussed separately and excluded from the quantitative score distribution.

Table II reports the binary component codes and their study-level totals.

TABLE II. Exploratory Component Coverage Scores for Eleven Studies

Study	E	F	P	D	O	SRMIS
Bandalouski et al. [1]	0	1	1	1	0	3
Park et al. [2]	1	1	0	1	0	3
Becerra-Vicario et al. [4]	1	1	0	0	0	2
Perramon et al. [5]	1	1	0	0	0	2
Chen et al. [6]	1	1	0	0	0	2
Bianco et al. [8]	1	1	1	1	0	4
Yang et al. [3]	1	1	0	1	0	3
He et al. [7]	1	1	0	0	0	2
Lin et al. [13]	1	1	1	1	1	5
Arbelo et al. [9]	1	1	0	0	0	2
Abdelrazek et al. [10]	1	1	1	1	0	4

E = environmental or broader sustainability measure; F = economic outcome or intention; P = direct room-price or capacity outcome or decision; D = demand or customer behavior; O = joint environmental-economic analytical decision structure. All components are binary. Scores describe coverage, not quality or effectiveness.

ISBN: 978-605-80193-6-2

B. Component Coverage

The eleven scored studies yield a total SRMIS of 32 and a mean of 2.91. Five studies score two, three score three, two score four, and one scores five. Ten include an E component, all eleven include F, four include P, six include D, and one includes O. These counts describe the selected set, not the prevalence of integration across the entire literature. Excluding the pricing-only comparator, [1], gives a mean of 2.90 across ten studies containing E.

Two coding decisions clarify the boundary between economic evidence and operational pricing. Reference [8] receives P = 1 because ADR is an explained price outcome, and D = 1 because occupancy measures realized room utilization as a demand proxy. Reference [9] receives P = 0 because its dependent outcome is operating revenue within a revenue-efficiency framework. Room stock is an explanatory input, while premium-pricing potential is an interpretation. If this indirect price evidence were instead counted as P, the mean would become 3.00; the number of studies meeting O would remain one.

C. Thematic Comparison

The pricing comparator and sustainability-related financial studies address different decision levels. Reference [1] optimizes prices without environmental indicators. References [4], [5], [6], and [7] connect environmental practices or broader ESG measures with economic outcomes, without specifying daily room-pricing decisions. Aggregate ESG findings include social and governance information and cannot be interpreted as environmental effects alone.

Certification studies provide observed or inferred economic evidence. Reference [8] directly examines ADR and occupancy relative to competitors. Reference [9] estimates heterogeneous revenue-efficiency relationships, supporting potential premium pricing indirectly. Reference [10] uses realized transactions and reports approximately 43% higher sales during certified periods. This is a context-specific association from two hotels, not a general expected uplift or a pooled effect. Fixed-effects adjustment does not, by itself, eliminate every source of time-varying confounding or selection.

References [2] and [3] clarify customer intentions and perceived environmental value. Their stated willingness-to-pay outcomes differ from observed prices and bookings. Reference [12] discusses measurement, benchmarking, and communication through a teaching case. Its discussion questions motivate managerial action but do not provide an estimated pricing model, which justifies its exclusion from SRMIS calculations.

Reference [13] is the only scored study that meets the common environmental-economic decision criterion. Its analytical model considers seasonal demand, customer involvement, CO_2 reduction, emission costs, and room operating-cost savings. The review identifies a further modeling opportunity: testing separate energy, water, and waste measurements and constraints at the occupied-room-night level. This proposal extends the comparison toward operational resource accounting without claiming that the reviewed model lacks every possible resource-related consideration.

IV. DISCUSSION

The component profiles show partial overlap between sustainability-related economic evidence and operational pricing within this selected set. Distinguishing direct price outcomes, customer intentions, inferred premiums, and joint decisions clarifies the research gap. The mean score describes coverage and does not demonstrate operational effectiveness.

Certification, perceived practices, and aggregate ESG scores differ from physical resource intensity. Hotel indicators should distinguish energy, water, waste, and CO_2 per occupied room-night, alongside total consumption. Both are needed because efficiency improvements can coexist with higher total consumption when occupancy rises.

A future model could maximize sustainable net contribution: room revenue less variable operating and explicitly defined environmental costs. To avoid double counting, environmental costs should include only charges or monetized externalities outside operating costs. Resource limits should reflect technical or policy constraints. Demand, price elasticity, and room capacity could enter stochastic, robust, or multi-objective optimization. Evaluation should compare conventional pricing with financial results and both total and per-room-night environmental outcomes. This is a proposed research direction, not a model estimated or validated here.

A. Limitations

The selected corpus and incomplete 2026 window limit coverage. Unreported database-specific search logs, screening counts, and independent-coder agreement limit reproducibility and prevent claims of exhaustive retrieval or demonstrated coding reliability. Binary coding and equal weights simplify heterogeneous designs. No effect-size pooling or formal risk-of-bias assessment is undertaken. Differences in samples, measurement, and identification limit causal comparisons. Future reviews should report reproducible searches, independently checked coding with source locations, and design-specific evidence appraisal.

V. CONCLUSION

Across eleven scored quantitative or analytical hotel studies and one separately discussed teaching case, sustainability-related measures connect with economic outcomes through several mechanisms. Four scored studies directly examine pricing or capacity outcomes or decisions, and only [13] meets the joint environmental-economic decision criterion. The exploratory SRMIS makes component coverage visible but does not establish research quality or operational effectiveness. The clearest research opportunity within this evidence set is an explainable pricing framework that connects demand and financial outcomes with measured resource use per occupied room-night.

ISBN: 978-605-80193-6-2

References

[1] Bandalouski, A. M., Egorova, N. G., Kovalyov, M. Y., Pesch, E., & Tarim, S. A. (2021). Dynamic pricing with demand disaggregation for hotel revenue management. Journal of Heuristics, 27(5), 869-885. https://doi.org/10.1007/s10732-021-09480-2

[2] Park, J., Park, Y., Yoo, J. L., & Yu, J. (2021). Can hotel companies' water conservation management and waste reduction measures influence hotel customers' willingness to pay more and intention to revisit? International Journal of Environmental Research and Public Health, 18(17), 9054. https://doi.org/10.3390/ijerph18179054

[3] Yang, J., Reza, M. N. H., Al Mamun, A., Masud, M. M., & Che Abdul Rahman, M. R. (2024). Predicting guest satisfaction and willingness to pay premium prices for green hotels. Humanities and Social Sciences Communications, 11, 1487. https://doi.org/10.1057/s41599-024-04024-y

[4] Becerra-Vicario, R., Ruiz-Palomo, D., Fernández-Miguélez, S. M., & Gutiérrez-Ruiz, A. M. (2022). Examining the effects of hotel reputation in the relationship between environmental performance and hotel financial performance. Journal of Hospitality and Tourism Management, 53, 10-20. https://doi.org/10.1016/j.jhtm.2022.08.012

[5] Perramon, J., Oliveras-Villanueva, M., & Llach, J. (2022). Impact of service quality and environmental practices on hotel companies: An empirical approach. International Journal of Hospitality Management, 107, 103307. https://doi.org/10.1016/j.ijhm.2022.103307

[6] Chen, C.-D., Su, C.-H. J., & Chen, M.-H. (2022). Are ESG-committed hotels financially resilient to the COVID-19 pandemic? An autoregressive jump intensity trend model. Tourism Management, 93, 104581. https://doi.org/10.1016/j.tourman.2022.104581

[7] He, Y., Qi, R., So, K. K. F., & Li, Y. (2024). Which ESG dimensions matter in the hotel industry? Evidence from the cost of debt. International Journal of Hospitality Management, 122, 103866. https://doi.org/10.1016/j.ijhm.2024.103866

[8] Bianco, S., Bernard, S., & Singal, M. (2023). The impact of sustainability certifications on performance and competitive action in hotels. International Journal of Hospitality Management, 108, 103379. https://doi.org/10.1016/j.ijhm.2022.103379

[9] Arbelo, A., Arbelo-Pérez, M., De Vera, V., & Bilgihan, A. (2025). Green premiums: Assessing the revenue impact of eco-certification in the hospitality sector. International Journal of Contemporary Hospitality Management, 37(13), 64-83. https://doi.org/10.1108/IJCHM-04-2024-0534

[10] Abdelrazek, M. A., Dose, D., Rabea, A., & Agag, G. (2026). Understanding the relationship between eco-certification and hotel performance: A longitudinal analysis of two comparable hotels. International Journal of Hospitality Management, 135, 104619. https://doi.org/10.1016/j.ijhm.2026.104619

[11] Vrionis, I. A., & Varelas, S. (2026). Implementing sustainable habits in dynamic pricing through revenue management strategies in the hotel industry: A literature review. In A. Kavoura, U. Gretzel, & V. Vrana (Eds.), Strategic Innovative Marketing and Tourism (pp. 761-768). Springer. https://doi.org/10.1007/978-3-032-12968-0_83

[12] Kruesi, M. A., & Remy, D. (2025). Green is the new black in the hotel industry: Optimizing revenue through promoting sustainability. Journal of Hospitality & Tourism Cases, 13(3), 134-141. https://doi.org/10.1177/21649987241270463

[13] Lin, Y., Chiu, C.-H., & Guo, X. (2025). Achieving win-win situation: Strategy decisions on hotel CO2 emission reduction with customer involvement. Annals of Operations Research, 346(3), 2093-2139. https://doi.org/10.1007/s10479-025-06493-z

 ISBN: 978-605-80193-6-2

Determination of the Heating Energy Demand of a Public Building According to the TS 825:2008 and TS 825:2024 Standards: A Comparative Assessment for Four Different Provinces

Enes KAVUN[1,*] Ertaç HÜRDOĞAN[2]

[1]Osmaniye Korkut Ata University, Energy Education-Etude Application and Research Center, Osmaniye, Türkiye

[2]Osmaniye Korkut Ata University, Faculty of Engineering and Natural Sciences, Department of Energy Systems Engineering, Osmaniye, Türkiye

Article note: This paper is derived from the doctoral dissertation of Enes KAVUN.

* Corresponding author e-mail: eneskavun@osmaniye.edu.tr

***Abstract*—The TS 825 standard, which sets out the thermal insulation rules for buildings in Türkiye, was revised in 2024 and has been mandatory for new buildings since 1 April 2025. With the revision, the provinces were reclassified according to their heating and cooling degree-day values, the number of climate zones was increased from four to six, and the recommended thermal transmittance (U) values and the limit values of energy demand were lowered. In this study, the Rectorate Building of Osmaniye Korkut Ata University was selected as the sample building, and its heating energy demand was examined comparatively according to the TS 825:2008 and TS 825:2024 standards, assuming that the building is located in the provinces of Osmaniye, Giresun, Çankırı and Erzincan. For both standards, the U values of the building components, the required insulation thicknesses, the transmission heat loss coefficient of the building envelope and the annual heating energy demand were calculated. It was determined that the uninsulated existing building envelope does not meet the U values recommended by either standard, and that the largest shares of the transmission heat loss belong to the roof with 37.9% and the windows with 27.5%. It was found that compliance with TS 825:2024 requires the application of 6.1–7.2 cm of XPS on the external walls, and 12.3–15.0 cm on the roof and 9.8–11.7 cm on the ground floor of stone wool. It was observed that the transition from the TS 825:2008 to the TS 825:2024 U values reduced the heat loss of the opaque envelope (external walls, roof and ground floor) by 42.6% in Osmaniye, whose climate zone changed, and by 13.4% in Erzincan, whose zone remained unchanged. When the calculation method and climatic data were kept constant, it was calculated that meeting the TS 825:2024 U values reduced the annual heating energy demand by 66.2–70.1% compared to the existing condition and by 14.3–25.1% compared to the TS 825:2008 U values. The results indicate that the effect of the new standard largely depends on the reclassification of the provinces, and it was determined that it would be appropriate to take the current standard as the basis for the retrofits to be carried out in existing public buildings.**

***Keywords*—TS 825, thermal insulation, climate zones, public building, heating energy demand**

I. INTRODUCTION

The building sector holds an important place in the fight against climate change due to its high energy consumption and the associated greenhouse gas emissions. The recast Energy Performance of Buildings Directive of the European Union, revised in 2024, aims for new buildings to be constructed as zero-emission buildings from 2030 onwards and for public buildings to take a leading role in this transformation [1]. In Türkiye, the main regulations on energy efficiency in buildings are set out by the Regulation on Thermal Insulation in Buildings, the Regulation on Energy Performance of Buildings and the TS 825 "Thermal Insulation Requirements for Buildings" standard referred to by these regulations [2], [3], [4].

The thermal performance of the building envelope is one of the most important parameters determining the heating and cooling energy demand. Transmission heat losses decrease as the thermal transmittance coefficients (U) of the external walls, roof, floor and windows are reduced. For this reason, the U values recommended for building components and the corresponding insulation thicknesses form the basis of thermal insulation standards. The TS 825 standard, first published in Türkiye in 1998, was updated in 2008 and revised again in October 2024 [4], [5].

With TS 825:2024, the heating energy calculation method was harmonised with TS EN ISO 52016-1, the calculation of the net cooling energy demand was added to the standard, the number of climate zones was increased from four to six, and the recommended U values and the limit values of energy demand were lowered [6], [5]. The new standard has been mandatory for new buildings since 1 April 2025. With the reclassification of the provinces according to their heating and cooling degree-day values, the same building faces different insulation requirements depending on the province in which it is located. Therefore, it was considered beneficial to examine the building-scale effect of the change in the standard separately for different climate zones.

In this study, the Rectorate Building of Osmaniye Korkut Ata University was selected as the sample building, and the building was evaluated according to the two standards by assuming that it is located in the provinces of Osmaniye, Giresun, Çankırı and Erzincan, which represent different degree-day zones according to TS 825:2008 [7]. For the comparison to be consistent, the evaluation was limited to the heating energy demand. Within the scope of the study, the difference

ISBN: 978-605-80193-6-2

between the two standards was revealed on a real public building instead of generic wall sections, the effect of the reclassification of the provinces was examined by comparing provinces whose climate zone changed and did not change, and the insulation thicknesses required for the compliance of the existing building envelope with the new standard were determined.

II. LITERATURE REVIEW

Determining the insulation thickness to be applied to the building envelope for the different climate zones of Türkiye has long been studied in the literature. In most of these studies, the optimum insulation thickness was determined from economic and environmental perspectives based on the climate zones and degree-day values defined in TS 825. Çomaklı and Yüksel [8] determined the optimum insulation thickness for external walls under cold climate conditions by life-cycle cost analysis and showed that significant energy savings can be achieved with insulation. Bolattürk [11] calculated the optimum insulation thicknesses, energy savings and payback periods according to different fuel types for provinces representing the climate zones of TS 825. Sisman et al. [9] determined the optimum insulation thicknesses of external walls and roofs (ceilings) for the different degree-day regions of Türkiye. Uçar and Balo [10], on the other hand, examined the effect of fuel type on the optimum thickness of selected insulation materials in four different climate zones.

In some studies, cooling load and environmental impacts were also taken into account, in addition to the heating load, when determining the insulation thickness. Bolattürk [13] used heating and cooling degree-hour values together for the warmest climate zone of Türkiye and showed that the cooling load should also be taken into account when determining the insulation thickness in hot climates. Dombaycı [14] determined that fuel consumption and the associated emissions decreased with the application of the optimum insulation thickness. Özel [17] compared the thermal performance and optimum insulation thicknesses of walls made of different building materials. Bektaş Ekici et al. [12] found that the optimum insulation thickness for stone, brick and concrete walls varied between 0.2 and 18.6 cm depending on the insulation material, fuel type and climate zone. Kaynaklı [15] evaluated the methods and findings related to the determination of the optimum insulation thickness in a review study, while Kürekci [16] calculated the optimum external wall insulation thicknesses on a provincial basis using the heating and cooling degree-day values of the 81 provincial centres of Türkiye.

Following the publication of TS 825:2024, studies comparing the two standards have also begun to appear. Ünal [18] compared the U values recommended by the two standards and the corresponding insulation thicknesses on a provincial basis and stated that the external wall U values decreased by up to 43% in the provinces whose climate zone changed. However, this evaluation was carried out on generic wall and ceiling sections and does not include results on heat loss and heating energy demand at the building scale. It was found that the studies in the literature are mainly based on individual building components and on previous versions of TS 825. No study examining the effect of TS 825:2024 on a real building in terms of heating energy demand in different climate zones was encountered.

III. METHODOLOGY

The Rectorate Building located on the Karacaoğlan Campus of Osmaniye Korkut Ata University was selected as the sample building in the study. The heat-losing vertical surface area of the building is 1732.5 m², of which 816.5 m² is external wall and 916 m² is window area. The distribution of the window areas over the façades is 349.6 m² south, 198.72 m² north, 202.4 m² east and 165.6 m² west. The window-to-vertical external surface ratio is 52.9%, which remains below the 60% limit defined in TS 825:2024 for buildings with a high proportion of glazing. The roof and ground floor areas are 1440 m² each. The gross volume of the building is 16,200 m³ (32 m × 45 m × 11.25 m), its A/V ratio is 0.285 m^{-1}, and its usable floor area according to TS 825:2008 ($A_n = 0.32 \cdot V_{gross}$) is 5184 m². In its current condition, the building envelope has no thermal insulation [7]. The model of the building created in the SketchUp software is given in Fig. 1.

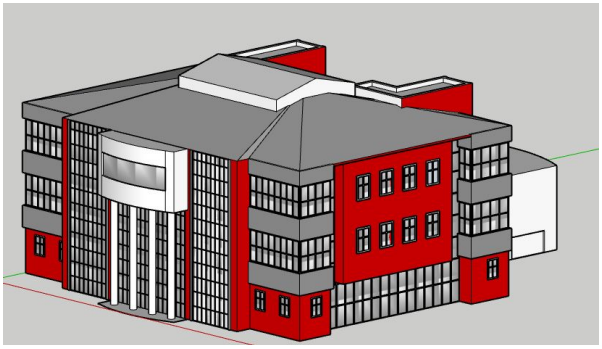


Fig. 1. SketchUp/TRNSYS3d model of the reference building [7]

The layer properties of the building components and the calculated U values are given in Table I. In the calculations, the internal and external surface thermal resistances were taken as 0.13 and 0.04 m²K/W, respectively, for the external wall and the roof, and the internal surface thermal resistance was taken as 0.13 m²K/W for the ground floor. The thermal transmittance coefficient of the windows is 2.8 W/m²K [7].

ISBN: 978-605-80193-6-2

TABLE I. Existing building components of the reference building [7]

Building component	Layers (thickness, cm / λ, W/mK)	Area (m²)	Existing U (W/m²K)
External wall	Fine plaster (2 / 0.87) Rough plaster (2 / 1.4) Brick (13 / 0.45) Rough plaster (3 / 1.4) Gypsum plaster (2 / 0.87)	816.5	1.850
Roof	Reinforced concrete (12 / 2.1) Rough plaster (1.5 / 1.4) Gypsum plaster (2 / 0.87) Stone cladding (3 / 1.3)	1440	3.834
Ground floor	Levelling concrete (5 / 1.4) Reinforced concrete (15 / 1.74) Crushed stone (10 / 0.7)	1440	2.393
Window	Double glazing	916	2.8

The building was evaluated in the provinces of Osmaniye (Zone 1), Giresun (Zone 2), Çankırı (Zone 3) and Erzincan (Zone 4), each of which represents one of the four degree-day zones according to TS 825:2008. With TS 825:2024, the number of climate zones was increased to six, and the provinces were reclassified according to their heating and cooling degree-day values [5]. In this classification, the central districts of Osmaniye, Giresun and Çankırı move to Zones 2, 3 and 4, respectively, whereas the central district of Erzincan remains in Zone 4 in both standards [19]. Thus, provinces whose zone changed and did not change could be compared. Since TS 825:2024 defines different climate zones for districts within provincial boundaries (the Şebinkarahisar district of Giresun, the Çerkeş district of Çankırı and the Tercan district of Erzincan are in Zone 5), the central district was taken as the basis for all provinces.

The climate zones of the selected provinces in the two standards and the recommended U values are given in Table II. As seen in Table II, the provinces divided into four different U value groups in TS 825:2008 are reduced to two groups in TS 825:2024: Osmaniye and Giresun, and Çankırı and Erzincan have the same recommended U values. The recommended U value for windows was reduced from 2.4 W/m²K to 1.8 W/m²K in all provinces.

TABLE II. Climate zones of the selected provinces and recommended U values (W/m²K) [4], [5], [19]

Province (central district)	Zone 2008	Zone 2024	U wall 2008 - 2024	U ceiling 2008 - 2024	U floor 2008 - 2024	U window 2008 - 2024
Osmaniye	1	2	0.70 - 0.40	0.45 - 0.30	0.70 - 0.35	2.4 - 1.8
Giresun	2	3	0.60 - 0.40	0.40 - 0.30	0.60 - 0.35	2.4 - 1.8
Çankırı	3	4	0.50 - 0.35	0.30 - 0.25	0.45 - 0.30	2.4 - 1.8
Erzincan	4	4	0.40 - 0.35	0.25 - 0.25	0.40 - 0.30	2.4 - 1.8

In TS 825:2008, the annual heating energy demand of a building is calculated with the monthly heat balance method using (1) [4]:

$$Q_{year} = \sum Q_{month} = \sum \left[H(\theta_i - \theta_e) - \eta_{month}\left(\phi_{i,month} + \phi_{g,month}\right) \right] \cdot t \quad (1)$$

Here, H is the specific heat loss of the building (W/K), θ_i is the indoor temperature (°C), θ_e is the monthly mean outdoor temperature (°C), η is the gain utilisation factor, ϕ_i and ϕ_g are the monthly mean internal and solar gains (W), respectively, and t is the duration of the month (s). The specific heat loss is the sum of the transmission and ventilation heat losses, as given in (2):

$$H = H_T + H_V = \sum A \cdot U + \sum l \cdot U_l + 0.33 \cdot n_h \cdot V_h \quad (2)$$

In the calculations, the indoor temperature of 19 °C recommended in TS 825 for administrative buildings was used, and the monthly mean outdoor temperatures and solar radiation values were taken from the zone data in TS 825:2008 [4]. The ventilated volume was taken as $V_h = 0.8 \cdot V_{gross}$ according to TS 825 [4], and the air change rate was assumed to be 1 h^{-1} [7]. The monthly mean internal gains were calculated using the value of 5 W/m² given in TS 825 for residences, schools and normally equipped buildings and the usable floor area [4]. For the transparent surfaces, the shading factor was taken as 0.6 and the solar energy transmittance as $g = F_w \cdot g_\perp = 0.8 \times 0.75 = 0.6$. The gain utilisation factor was calculated with the equation $\eta = 1 - e^{-1/GLR}$, and it was assumed that no thermal bridges or condensation occur [4], [7]. The correction factors defined in TS 825 were applied for the ceiling and the ground floor. Accordingly, the transmission heat loss coefficient, the ventilation heat loss coefficient and the monthly mean internal gains of the existing building were calculated as 9331.4 W/K, 4276.8 W/K and 25,920 W, respectively.

In TS 825:2024, the net heating energy demand is calculated with a monthly method harmonised with TS EN ISO 52016-1, the indoor set-point temperature is taken as 20 °C, and the climatic data are derived from typical meteorological year data [6], [5]. In this study, in order to separate the effect of the change in the recommended U values from the changes in the calculation method and climatic data, all scenarios were calculated using the TS 825:2008 calculation method and climatic data. Thus, the difference between the scenarios results solely from the thermal transmittance of the building envelope.

Three scenarios were defined for each province. In the first scenario (S1), the uninsulated existing building was considered. In the second scenario (S2), the building components were insulated to meet the U values recommended for the TS 825:2008 zone of the province, and the window U value was taken as 2.4 W/m²K. In the third scenario (S3), the building components were insulated to meet the U values recommended for the TS 825:2024 zone of the province, and the window U value was taken as 1.8 W/m²K. XPS (λ = 0.031 W/mK) was used as the insulation material

ISBN: 978-605-80193-6-2

on the external walls, and stone wool (λ = 0.040 W/mK) on the roof and the ground floor [7].

IV. RESULTS AND DISCUSSION

A. U Values and Required Insulation Thicknesses

In the existing condition, the U value of the external wall is 1.850 W/m²K, that of the roof is 3.834 W/m²K and that of the ground floor is 2.393 W/m²K (Table I). These values are higher than all of the values recommended by both standards for the four provinces. The insulation thicknesses required to meet the recommended U values are given in Table III. When Table III is examined, it is seen that, except for the roof insulation in Erzincan, the required insulation thicknesses increase with TS 825:2024 in all provinces and for all building components. The largest increase occurred in Osmaniye, which moved from Zone 1 to Zone 2. In Osmaniye, the required thickness increased from 2.8 cm to 6.1 cm on the external wall, from 7.8 cm to 12.3 cm on the roof and from 4.0 cm to 9.8 cm on the ground floor.

TABLE III. Insulation thicknesses meeting the recommended U values (cm; XPS on the external wall, λ = 0.031 W/mK; stone wool on the roof and ground floor, λ = 0.040 W/mK)

Province	External wall 2008	External wall 2024	Roof 2008	Roof 2024	Ground floor 2008	Ground floor 2024
Osmaniye	2.8	6.1	7.8	12.3	4.0	9.8
Giresun	3.5	6.1	9.0	12.3	5.0	9.8
Çankırı	4.5	7.2	12.3	15.0	7.2	11.7
Erzincan	6.1	7.2	15.0	15.0	8.3	11.7

In Erzincan, whose zone did not change, the increase in external wall insulation remained limited to 1.1 cm, and since the U value recommended for the roof did not change, the roof insulation thickness remained constant at 15.0 cm. Since the U values recommended for the ground floor were lowered in all zones, an additional thickness of 3.3–5.7 cm was required for the ground floor. These results show that the effect of the new standard on insulation thickness is greater in provinces whose climate zone changed, and it was determined that the trend reported by Ünal [18] for generic sections is also valid for the sample building.

B. Transmission Heat Loss of the Building Envelope

In order to determine the contribution of the building components to the heat loss, the transmission heat loss coefficient of the building envelope (H_T) was calculated for each scenario, and the results are given in Table IV. In the existing condition, the H_T value is 9331 W/K, of which 37.9% originates from the roof, 27.5% from the windows, 18.5% from the ground floor and 16.2% from the external walls. When the windows are excluded, the share of the roof in the heat loss of the opaque envelope (external walls, roof and ground floor) is 52.2%. While the H_T value decreases by 60.5–67.4% when the U values recommended in TS 825:2008 are met, this ratio increases to 73.2–74.5% with the TS 825:2024 values.

TABLE IV. Transmission heat loss coefficient of the building envelope, H_T (W/K)

Province	Existing	TS 825:2008 U values	TS 825:2024 U values	2008 reduction (%)	2024 reduction (%)	2008 - 2024 reduction (%)
Osmaniye	9331	3689	2504	60.5	73.2	32.1
Giresun	9331	3489	2504	62.6	73.2	28.2
Çankırı	9331	3207	2381	65.6	74.5	25.8
Erzincan	9331	3043	2381	67.4	74.5	21.8

The difference between the values recommended by the two standards in the total envelope heat loss is 32.1% in Osmaniye and 21.8% in Erzincan. When the windows are excluded and only the opaque envelope is considered, this difference increases to 42.6% in Osmaniye, whereas it decreases to 13.4% in Erzincan. Accordingly, the effect of TS 825:2024 largely results from the reclassification of the provinces, while in provinces whose zone did not change, a significant part of the difference comes from the decrease in the window U value (from 2.4 W/m²K to 1.8 W/m²K). It was found that, in the sample building with a window-to-vertical external surface ratio of 52.9%, the effect of this improvement in the windows on the total heat loss is pronounced.

C. Annual Heating Energy Demand

The annual heating energy demands calculated for the three scenarios in each province are given in Table V. The annual heating energy demand of the uninsulated existing building (S1) was found to be 189,391.7, 400,966.5, 574,323.9 and 775,983.9 kWh for Osmaniye, Giresun, Çankırı and Erzincan, respectively [7]. These values correspond to 36.5, 77.3, 110.8 and 149.7 kWh/m² per unit usable floor area. The heating demand is concentrated in January, February, March, November and December, and no heating demand occurs in Osmaniye during the May–September period.

When the building envelope meets the U values recommended in TS 825:2008 (S2), the annual heating energy demand decreases by 59.8–61.2% compared to the existing condition, whereas this ratio increases to 66.2–70.1% when the TS 825:2024 values are met (S3). The additional reduction obtained by moving from S2 to S3 is 25.1% in Osmaniye, 20.7% in Giresun, 17.8% in Çankırı and 14.3% in Erzincan. The greater additional reduction in the provinces whose climate zone changed also supports, in terms of energy demand, the result obtained for transmission heat loss in Section IV-B.

ISBN: 978-605-80193-6-2

TABLE V. Annual heating energy demand according to the scenarios

Province	S1 (kWh)	S1 (kWh/m²)	S2 (kWh)	S2 (kWh/m²)	S3 (kWh)	S3 (kWh/m²)
Osmaniye	189,391.7	36.5	75,439.2	14.6	56,538.6	10.9
Giresun	400,966.5	77.3	161,329.9	31.1	127,907.5	24.7
Çankırı	574,323.9	110.8	222,864.8	43.0	183,124.5	35.3
Erzincan	775,983.9	149.7	306,469.8	59.1	262,663.0	50.7

In TS 825:2024, the energy limit is given as a single value in which the net heating and net cooling demands are evaluated together, and it is taken as 70 kWh/m²·year for office buildings in every zone [5]. Even when only the heating energy demand of the existing building is considered, this limit is exceeded in Giresun, Çankırı and Erzincan. Although the heating demand in the S3 scenario decreases to the range of 10.9–50.7 kWh/m², it was determined that it would not be correct to state that this scenario meets the limit value without taking the cooling demand into account. Moreover, since this comparison was made with the TS 825:2008 data and 19 °C instead of the TS 825:2024 climatic data and 20 °C indoor temperature, it should be considered only as an indicator.

D. Compliance of the Existing Building with the New Standard

Since the existing building is uninsulated, it was determined that 6.1–7.2 cm of XPS on the external wall, and 12.3–15.0 cm on the roof and 9.8–11.7 cm on the ground floor of stone wool need to be applied for compliance with TS 825:2024 (Table III). With this retrofit, it was found that the annual heating energy demand decreases by 66.2–70.1% compared to the existing condition and by 14.3–25.1% compared to a retrofit at the TS 825:2008 level (Table V). It was determined that, among the building components, the largest share in the transmission heat loss belongs to the roof with 37.9%, followed by the windows with 27.5%, the ground floor with 18.5% and the external walls with 16.2%. It was found that the total share of the vertical façade consisting of the external walls and windows is 43.7% and that most of this loss originates from the windows. It was observed that the U value of the uninsulated reinforced concrete roof (3.834 W/m²K) is more than twice that of the external wall (1.850 W/m²K), and it was evaluated that priority in the retrofit should be given to roof insulation and, on the façade, to the windows. It was calculated that, for a retrofit remaining at the TS 825:2008 level, the annual heating energy demand in Osmaniye is approximately 33% higher than at the TS 825:2024 level. In line with these results, it was concluded that it would be appropriate to take the current standard as the basis for the energy efficiency retrofits to be carried out in existing public buildings.

It was determined that the U values defined by TS 825:2024 are lower than those of the previous standard. However, it is seen that improving the building envelope alone is not sufficient to achieve the zero-emission building targets of the European Union, and that high-efficiency heating systems and on-site renewable energy generation are also defined as part of these targets [1]. Accordingly, it was evaluated that it would be appropriate to plan envelope retrofits in public buildings together with these systems. In addition, it was found in the literature that the optimum insulation thickness varies over a wide range depending on the fuel type, building material and climatic conditions [12], [16]. Therefore, it was concluded that insulation thicknesses above the standard values should also be examined from an economic perspective.

V. CONCLUSIONS

In this study, the Rectorate Building of Osmaniye Korkut Ata University was examined according to the TS 825:2008 and TS 825:2024 standards by assuming that it is located in the provinces of Osmaniye, Giresun, Çankırı and Erzincan. With TS 825:2024, the recommended U values decreased in all provinces, and the insulation thicknesses required for the sample building increased by 1.1–3.3 cm on the external wall, 0–4.4 cm on the roof and 3.3–5.7 cm on the ground floor. It was found that the effect of the new standard largely depends on the reclassification of the provinces. It was observed that the transition to the TS 825:2024 U values reduced the heat loss of the opaque envelope (external walls, roof and ground floor) by 42.6% in Osmaniye, whose climate zone changed, and by 13.4% in Erzincan, whose zone remained unchanged. Since 37.9% of the transmission heat loss in the uninsulated existing building originates from the roof and 27.5% from the windows, it was determined that it would be appropriate to give priority to roof insulation in the retrofit. The annual heating energy demand of the existing building per unit usable floor area is 36.5 kWh/m² in Osmaniye and 149.7 kWh/m² in Erzincan. In Giresun, Çankırı and Erzincan, the heating demand alone exceeds the limit of 70 kWh/m²·year given by TS 825:2024 for the sum of heating and cooling in office buildings. It was determined that meeting the TS 825:2024 U values reduces the heating energy demand by 66.2–70.1% compared to the existing condition and by 14.3–25.1% compared to the TS 825:2008 U values. The largest additional reduction was obtained in Osmaniye, whose climate zone changed, and the smallest in Erzincan, whose zone remained unchanged.

In conclusion, it was found that TS 825:2024 reduces the heating energy demand of existing public buildings more than the previous standard, and it was concluded that it would be appropriate to take the current standard as the basis for the retrofits to be carried out in these buildings. In future studies, it is planned to repeat the analysis with the calculation method and climatic data of TS 825:2024, to include the cooling energy demand in the evaluation, and to examine the retrofit scenarios from economic and environmental perspectives.

This study was supported by the Scientific Research Projects Unit of Osmaniye Korkut Ata University under project number OKÜBAP-2024-PT2-017.

ISBN: 978-605-80193-6-2

References

[1] European Parliament and Council of the European Union. (2024). Directive (EU) 2024/1275 of the European Parliament and of the Council of 24 April 2024 on the energy performance of buildings (recast). Official Journal of the European Union, L 2024/1275.

[2] Regulation on Thermal Insulation in Buildings [Binalarda Isı Yalıtımı Yönetmeliği]. (2008). Official Gazette of the Republic of Türkiye, No. 27019, 09.10.2008.

[3] Regulation on Energy Performance of Buildings [Binalarda Enerji Performansı Yönetmeliği]. (2008). Official Gazette of the Republic of Türkiye, No. 27075, 05.12.2008.

[4] TSE. (2008). TS 825 Binalarda ısı yalıtım kuralları [Thermal insulation requirements for buildings]. Turkish Standards Institution, Ankara.

[5] TSE. (2024). TS 825 Binalarda ısı yalıtım kuralları [Thermal insulation requirements for buildings]. Turkish Standards Institution, Ankara.

[6] TSE. (2017). TS EN ISO 52016-1 Binaların enerji performansı – Isıtma ve soğutma için enerji ihtiyacı, iç sıcaklıklar ile duyulur ve gizli ısı yükleri – Bölüm 1: Hesaplama prosedürleri [Energy performance of buildings – Energy needs for heating and cooling, internal temperatures and sensible and latent heat loads – Part 1: Calculation procedures]. Turkish Standards Institution, Ankara.

[7] Kavun, E. (2026). Binalarda Enerji Verimliliği Alanının Bibliyometrik Analizi ve Örnek Bir Kamu Binasının Enerji Performansının Değerlendirilmesi [Bibliometric analysis of the field of energy efficiency in buildings and evaluation of the energy performance of a sample public building] (Doctoral dissertation). Osmaniye Korkut Ata University, Graduate Education Institute, Osmaniye.

[8] Çomaklı, K., & Yüksel, B. (2003). Optimum insulation thickness of external walls for energy saving. Applied Thermal Engineering, 23(4), 473–479. https://doi.org/10.1016/S1359-4311(02)00209-0

[9] Sisman, N., Kahya, E., Aras, N., & Aras, H. (2007). Determination of optimum insulation thicknesses of the external walls and roof (ceiling) for Turkey's different degree-day regions. Energy Policy, 35(10), 5151–5155. https://doi.org/10.1016/j.enpol.2007.04.037

[10] Uçar, A., & Balo, F. (2009). Effect of fuel type on the optimum thickness of selected insulation materials for the four different climatic regions of Turkey. Applied Energy, 86(5), 730–736. https://doi.org/10.1016/j.apenergy.2008.09.015

[11] Bolattürk, A. (2006). Determination of optimum insulation thickness for building walls with respect to various fuels and climate zones in Turkey. Applied Thermal Engineering, 26(11–12), 1301–1309. https://doi.org/10.1016/j.applthermaleng.2005.10.019

[12] Bektaş Ekici, B., Aytaç Gülten, A., & Aksoy, U. T. (2012). A study on the optimum insulation thicknesses of various types of external walls with respect to different materials, fuels and climate zones in Turkey. Applied Energy, 92, 211–217. https://doi.org/10.1016/j.apenergy.2011.10.008

[13] Bolattürk, A. (2008). Optimum insulation thicknesses for building walls with respect to cooling and heating degree-hours in the warmest zone of Turkey. Building and Environment, 43(6), 1055–1064. https://doi.org/10.1016/j.buildenv.2007.02.014

[14] Dombaycı, Ö. A. (2007). The environmental impact of optimum insulation thickness for external walls of buildings. Building and Environment, 42(11), 3855–3859. https://doi.org/10.1016/j.buildenv.2006.10.054

[15] Kaynaklı, Ö. (2012). A review of the economical and optimum thermal insulation thickness for building applications. Renewable and Sustainable Energy Reviews, 16(1), 415–425. https://doi.org/10.1016/j.rser.2011.08.006

[16] Kürekci, N. A. (2016). Determination of optimum insulation thickness for building walls by using heating and cooling degree-day values of all Turkey's provincial centers. Energy and Buildings, 118, 197–213. https://doi.org/10.1016/j.enbuild.2016.03.004

[17] Özel, M. (2011). Thermal performance and optimum insulation thickness of building walls with different structure materials. Applied Thermal Engineering, 31(17–18), 3854–3863. https://doi.org/10.1016/j.applthermaleng.2011.07.033

[18] Ünal, Ö. (2025). TS 825 2008 ve 2024 standartlarına göre U değerleri ve yalıtım kalınlığı karşılaştırması [Comparison of U values and insulation thickness according to the TS 825 2008 and 2024 standards]. 16th National Plumbing Engineering Congress (TESKON 2025), İzmir.

[19] İZODER. (2025). TS 825:2024 il bilgi notları (Osmaniye, Giresun, Çankırı, Erzincan) [TS 825:2024 provincial information notes (Osmaniye, Giresun, Çankırı, Erzincan)]. İZODER Heat, Water, Sound and Fire Insulators Association. https://www.izoder.org.tr/ts825-harita (Accessed: 27.09.2026)

 ISBN: 978-605-80193-6-2

The Effect of Different Thermal Insulation Materials on the Heating-Related Carbon Emissions of a Building

Enes KAVUN[1,*] Ertaç HÜRDOĞAN[2]

[1]Osmaniye Korkut Ata University, Energy Education-Etude Application and Research Center, Osmaniye, Türkiye

[2]Osmaniye Korkut Ata University, Faculty of Engineering and Natural Sciences, Department of Energy Systems Engineering, Osmaniye, Türkiye

Article note: This paper is derived from the doctoral dissertation of Enes KAVUN.

* Corresponding author e-mail: eneskavun@osmaniye.edu.tr

Abstract*—The energy consumed for heating in buildings constitutes a significant part of carbon emissions. Thermal insulation applied to the building envelope is one of the main measures for reducing the heating energy demand and the associated emissions. In this study, the Rectorate Building of Osmaniye Korkut Ata University was selected as the sample building, and the effect of applying extruded polystyrene (XPS), expanded polystyrene (EPS), stone wool and glass wool to the external walls at thicknesses of 3, 4 and 5 cm on the heating-related carbon emissions was investigated. The roof, ground floor and windows were defined according to the U values recommended for the 2nd degree-day zone in which Osmaniye is located according to TS 825:2024, and 13 cm and 10 cm of stone wool were applied to the roof and the ground floor, respectively. The annual heating energy demand was calculated with the monthly heat balance method, and the emissions were determined using the emission conversion factor for natural gas. The annual heating-related emissions of the uninsulated building were calculated as 49.2 t CO_2e. It was found that, with the insulation applications, the annual emissions decreased to 14.8–15.8 t CO_2e depending on the material and thickness, corresponding to a reduction of 67.8–69.9%. It was determined that, for the same thickness, the lowest emissions were obtained with XPS, which has the lowest thermal conductivity, and the highest with stone wool. The difference between the materials was observed to be 0.37 t CO_2e per year at a thickness of 3 cm and 0.30 t CO_2e per year at 5 cm. It was determined that increasing the thickness from 3 cm to 5 cm reduced the emissions by a further 0.64–0.71 t CO_2e per year depending on the material, and that the insulation thickness was more effective than the material type in reducing the emissions.

***Keywords*—Carbon emissions, thermal insulation, insulation materials, heating energy, TS 825**

I. INTRODUCTION

The building sector holds an important place in the fight against climate change due to its high energy consumption and the associated greenhouse gas emissions. The recast Energy Performance of Buildings Directive of the European Union, revised in 2024, aims to reduce emissions from buildings and for public buildings to take a leading role in this transformation [1]. Thermal insulation applied to the building envelope is one of the main measures for reducing energy consumption for heating.

In Türkiye, the rules on thermal insulation to be applied in buildings are set out by the TS 825 standard, and the recommended thermal transmittance (U) values were lowered with TS 825:2024, which became mandatory on 1 April 2025 [2], [3]. In practice, different insulation materials can be used at the same thickness, and since the thermal conductivities of the materials differ, different U values and, consequently, different heating energy demands arise at the same thickness [4], [5].

In this study, the effect of applying XPS, EPS, stone wool and glass wool at thicknesses of 3, 4 and 5 cm to the external walls of the Rectorate Building of Osmaniye Korkut Ata University on the heating-related carbon emissions was investigated [6]. The roof, ground floor and windows were defined according to the U values recommended for Osmaniye in TS 825:2024. In the study, the annual and monthly emissions were calculated according to the insulation material and thickness, the materials were compared at the same thickness, and the resulting emissions were evaluated together with the case in which all building components meet the TS 825:2024 values.

II. LITERATURE REVIEW

Thermal insulation is one of the main methods that reduce the heating energy demand and the associated emissions by decreasing the heat losses through the building envelope. Most of the studies on insulation thickness in Türkiye have focused on the economic optimum. Çomaklı and Yüksel [8] determined the optimum insulation thickness for external walls under cold climate conditions by life-cycle cost analysis. Bolattürk [12] calculated the optimum insulation thicknesses and payback periods for different fuel types and climate zones, while Şişman et al. [9] determined the optimum insulation thicknesses of external walls and roofs (ceilings) for different degree-day regions. Dombaycı [13] examined the effect of the optimum insulation thickness on fuel consumption and emissions. Uçar and Balo [10] determined the optimum insulation thicknesses for different fuel types in four climate zones of Türkiye. Bektaş Ekici et al. [7] found that the optimum insulation thickness varied between 0.2 and 18.6 cm for different wall types, insulation materials and fuel types. Kürekci [15] calculated the optimum thicknesses based on degree-day values for 81 provincial centres, while Kaynaklı [14] evaluated the methods and findings in this field in a review study. Ozalp et al. [11] evaluated a sample building with different building elements from energy and environmental perspectives.

ISBN: 978-605-80193-6-2

Schiavoni et al. [4] reviewed the thermal, physical and environmental properties of the insulation materials used in the building sector. Grazieschi et al. [5], on the other hand, examined the embodied energy and carbon values of common insulation materials and, within this scope, also provided the thermal conductivity ranges for EPS, XPS, stone wool and glass wool. In these studies, the thermal conductivities of the materials are seen to vary in the range of 0.031–0.045 W/mK. No study in the literature was encountered that examines the effect of different insulation materials at the same thickness on the heating-related emissions of a public building in Türkiye and relates the results to the TS 825:2024 values.

III. METHODOLOGY

A. Sample Building

The sample building is an uninsulated public building with a footprint area of 1440 m², a gross volume of 16,200 m³ and a usable floor area of 5184 m² (Fig. 1). The building has 816.5 m² of external wall, 1440 m² of roof, 1440 m² of ground floor and 916 m² of window area. The existing U values of the external wall, roof and ground floor are 1.850, 3.834 and 2.393 W/m²K, respectively, and the U value of the windows is 2.8 W/m²K. The building geometry and the layer properties of the building components were taken from Kavun [6]. The U values of the building components were calculated with the equation $U = 1/(Rsi + \Sigma R + Rse)$ by summing the thermal resistances of the layers ($R = d/\lambda$) and the internal and external surface thermal resistances [2].

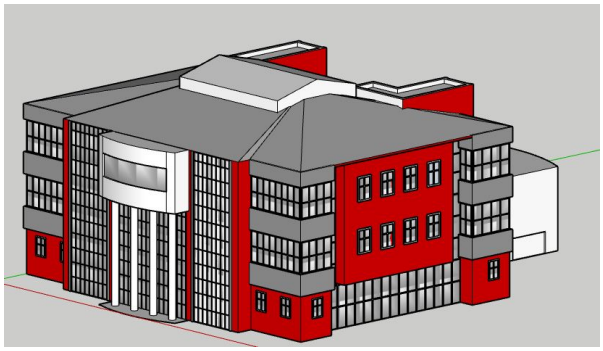


Fig. 1. SketchUp/TRNSYS3d model of the reference building [6]

B. Scenarios

According to TS 825:2024, Osmaniye is located in the 2nd degree-day zone. The U values recommended for this zone are 0.40 for the external wall, 0.30 for the roof, 0.35 for the ground floor and 1.8 W/m²K for the windows [2], [3]. In the study, a total of 12 insulation scenarios, in which XPS, EPS, stone wool and glass wool were applied to the external walls at thicknesses of 3, 4 and 5 cm, were examined together with the uninsulated existing building. In all insulation scenarios, 13 cm and 10 cm of stone wool were applied to the roof and the ground floor, respectively, so as to meet the U values recommended in TS 825:2024. The thicknesses required for these values were calculated as 12.3 cm for the roof and 9.8 cm for the ground floor and were rounded up to the next whole value, taking into account the board thicknesses used in practice. With these thicknesses, the U value of the roof was calculated as 0.285 W/m²K and that of the ground floor as 0.343 W/m²K. The U value of the windows was taken as 1.8 W/m²K, while the solar energy transmittance was kept the same as in the existing condition for the consistency of the comparison. Thus, it was ensured that the difference between the scenarios results solely from the insulation material and thickness used on the external wall. The thermal conductivity values of XPS and stone wool were taken as 0.031 and 0.040 W/mK, respectively [6]. For EPS, 0.035 W/mK, the lowest thermal conductivity group given in TS 825:2024, was used [2], and for glass wool, 0.038 W/mK, the midpoint of the range given by Grazieschi et al. [5], was used. The thermal conductivities of the materials and the U values of the external wall according to the scenarios are given in Table I.

TABLE I. Thermal conductivities of the insulation materials and U values of the external wall

Material	λ (W/mK)	U wall, 3 cm (W/m²K)	U wall, 4 cm (W/m²K)	U wall, 5 cm (W/m²K)
XPS	0.031	0.663	0.546	0.464
EPS	0.035	0.715	0.594	0.508
Stone wool	0.040	0.775	0.649	0.558
Glass wool	0.038	0.752	0.628	0.539

C. Heating Energy Demand

With TS 825:2024, changes were made such as the separate calculation of the net heating and net cooling energy demands and the increase in the number of climate zones to six [2]. In this study, only the heating energy demand and the associated emissions were evaluated. The building components were defined according to TS 825:2024, while the annual heating energy demand was calculated with the monthly heat balance method given in TS 825:2008 and the climatic data of the 1st degree-day zone in which Osmaniye is located according to that standard, so that the results could be compared with previous calculations for the same building, and the indoor temperature was taken as 19 °C [6]. The usable floor area of the building was calculated as 5184 m² with the equation $A_n = 0.32 \cdot V_{gross}$, and the monthly mean internal gains were taken as 25,920 W based on a value of 5 W/m² [16]. The ventilated volume was taken as 12,960 m³ with the equation $V_h = 0.8 \cdot V_{gross}$, and the air change rate was taken as 1 h^{-1} [6]. The transmission heat loss coefficient, the ventilation heat loss coefficient and the total specific heat loss of the existing building were calculated as 9331.4 W/K, 4276.8 W/K and 13,608.2 W/K, respectively. For the transparent surfaces, the shading factor was taken as 0.6 and the solar energy transmittance as $g = F_w \cdot g_\perp = 0.8 \times 0.75 = 0.6$, and the gain utilisation factor was calculated with the equation $\eta = 1 - e^{(-1/GLR)}$ [16], [6]. With these values, the annual heating energy demand of the uninsulated building was calculated as 189,391.7 kWh (36.5 kWh/m²).

D. Carbon Emissions

A natural gas-fired boiler was considered for heating the building, and the combustion efficiency was taken as 0.90. The annual heating-related carbon emissions were calculated with (1) [11], [6]:

$$E_{op} = \frac{Q_{year} \cdot FSEG}{\eta_k} \quad (1)$$

ISBN: 978-605-80193-6-2

Here, E_{op} is the annual heating-related carbon emission (kg CO_2e/year), Q_{year} is the annual heating energy demand (kWh/year), FSEG is the emission conversion factor for natural gas (0.234 kg CO_2e/kWh), and η_k is the combustion efficiency.

IV. RESULTS AND DISCUSSION

A. Heating-Related Emissions of the Uninsulated Building

The annual heating energy demand of the uninsulated building was calculated as 189,391.7 kWh and its annual heating-related emissions as 49.2 t CO_2e. The monthly emissions of the uninsulated building and of the building with 3, 4 and 5 cm of XPS, EPS, stone wool and glass wool applied to its external walls are given in Figs. 2–5, respectively. It was determined that the emissions are concentrated in January, February, March, November and December, with the highest value occurring in January.

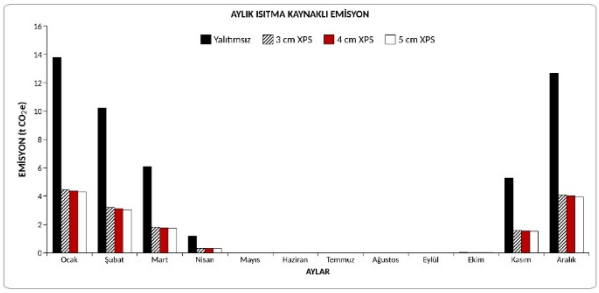


Fig. 2. Monthly heating-related emissions of the uninsulated building and of the building with 3, 4 and 5 cm XPS applied to its external walls

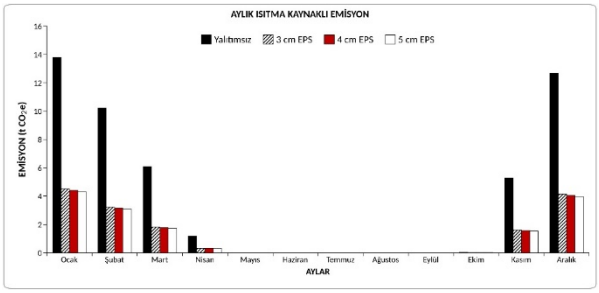


Fig. 3. Monthly heating-related emissions of the uninsulated building and of the building with 3, 4 and 5 cm EPS applied to its external walls

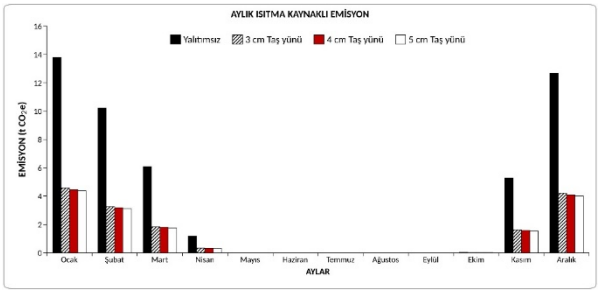


Fig. 4. Monthly heating-related emissions of the uninsulated building and of the building with 3, 4 and 5 cm stone wool applied to its external walls

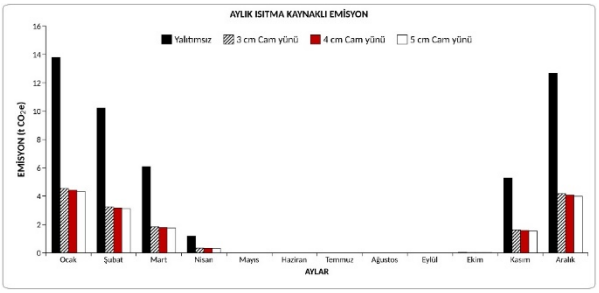


Fig. 5. Monthly heating-related emissions of the uninsulated building and of the building with 3, 4 and 5 cm glass wool applied to its external walls

It was observed that the emissions decreased in all months with the insulation applications. It was found that the emissions in January, which were 13.8 t CO_2e in the uninsulated building, decreased at thicknesses of 3, 4 and 5 cm to 4.46, 4.36 and 4.28 with XPS; 4.51, 4.40 and 4.32 with EPS; 4.55, 4.43 and 4.35 with glass wool; and 4.57, 4.45 and 4.37 t CO_2e with stone wool, respectively. It was observed that the monthly emissions showed a similar distribution for all four materials and that the emissions decreased slightly further in all months as the thickness increased. It was determined that the monthly difference between the materials at the same thickness remained at around 0.10 t CO_2e even in January, when it was highest.

B. Emissions According to Insulation Material and Thickness

The annual heating energy demand and emissions according to the insulation material and thickness are given in Table II and Fig. 6. It was found that, with the insulation applications, the annual emissions decreased to 14.83–15.84 t CO_2e depending on the material and thickness, corresponding to a reduction of 67.8–69.9% compared to the uninsulated condition. It was understood that most of this reduction results from the improvements to the roof, ground floor and windows. It was calculated that, when the external wall is left uninsulated and only the roof, ground floor and windows are brought to the TS 825:2024 values, the annual missions decrease to 19.53 t CO_2e, and that external wall insulation further reduces this value by 3.69–4.70 t CO_2e per year depending on the material and thickness.

TABLE II. Annual heating energy demand and carbon emissions according to the scenarios

Scenario	Heating energy demand (kWh/year)	Carbon emissions (t CO_2e/year)	Reduction compared to uninsulated case (%)
Uninsulated	189,391.7	49.24	–
XPS, 3 cm	59,511.3	15.47	68.6
XPS, 4 cm	58,053.6	15.09	69.3
XPS, 5 cm	57,041.2	14.83	69.9
EPS, 3 cm	60,170.5	15.64	68.2
EPS, 4 cm	58,649.0	15.25	69.0
EPS, 5 cm	57,578.2	14.97	69.6
Stone wool, 3 cm	60,920.3	15.84	67.8
Stone wool, 4 cm	59,337.3	15.43	68.7
Stone wool, 5 cm	58,206.5	15.13	69.3
Glass wool, 3 cm	60,629.6	15.76	68.0
Glass wool, 4 cm	59,069.0	15.36	68.8
Glass wool, 5 cm	57,960.6	15.07	69.4

It was determined that, for the same thickness, the lowest emissions were obtained with XPS, which has the lowest thermal conductivity, and the highest with stone wool, which has the highest thermal conductivity. EPS

ISBN: 978-605-80193-6-2

and glass wool lie between these two materials. The difference between the materials was observed to be 0.37 t CO_2e per year at a thickness of 3 cm and 0.30 t CO_2e per year at 5 cm. The additional reduction obtained by increasing the thickness from 3 cm to 5 cm was found to be 0.64–0.71 t CO_2e per year depending on the material. It was understood that the effect of increasing the thickness on the emissions is greater than the effect of changing the material at the same thickness. In addition, the emissions obtained with 4 cm of XPS (15.09 t CO_2e) were observed to be close to those obtained with 5 cm of stone wool (15.13 t CO_2e).

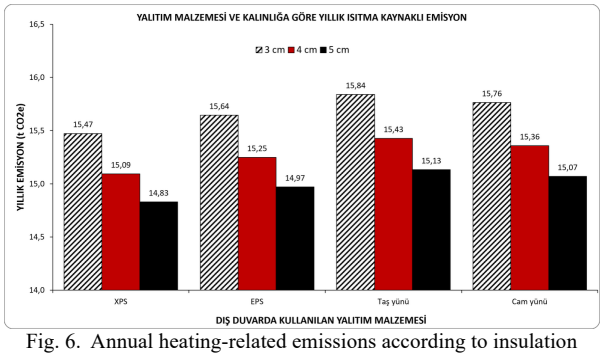


Fig. 6. Annual heating-related emissions according to insulation material and thickness

C. Comparison with the TS 825:2024 Value for the External Wall

If the external wall is also brought to the value of 0.40 W/m²K recommended in TS 825:2024, all building components meet the standard values, and the annual emissions are calculated as 14.62 t CO_2e. In the examined scenarios, the U value of the external wall lies in the range of 0.464–0.775 W/m²K depending on the material and thickness (Table I), and it was determined that the resulting annual emissions are 0.21–1.21 t CO_2e higher than this value. This difference was observed to be 0.85–1.21 t CO_2e at a thickness of 3 cm and 0.21–0.51 t CO_2e at 5 cm. Accordingly, it was understood that, as the thickness increases, the emissions approach the case in which all components meet the TS 825:2024 values, and that the emission level closest to this case is reached with 5 cm of XPS.

V. CONCLUSIONS

In this study, the effect of applying XPS, EPS, stone wool and glass wool at thicknesses of 3, 4 and 5 cm to the external walls of the Rectorate Building of Osmaniye Korkut Ata University, whose roof, ground floor and windows were brought to the TS 825:2024 values, on the heating-related carbon emissions was investigated. It was found that the annual emissions, which were 49.24 t CO_2e for the uninsulated building, decreased with the insulation applications to 14.83–15.84 t CO_2e depending on the material and thickness, corresponding to a reduction of 67.8–69.9%. It was determined that, compared to the case in which only the roof, ground floor and windows were improved, external wall insulation further reduced the annual emissions by 3.69–4.05 t CO_2e at a thickness of 3 cm, 4.10–4.43 t CO_2e at 4 cm and 4.39–4.70 t CO_2e at 5 cm. Accordingly, it was observed that external wall insulation provides a significant improvement in emissions even at a thickness of 3 cm.

It was determined that, for the same thickness, the lowest emissions were obtained with XPS, which has the lowest thermal conductivity, and the highest with stone wool, and that the difference between the materials remained at 0.30–0.37 t CO_2e per year. It was observed that increasing the thickness from 3 cm to 5 cm further reduced the emissions by 0.64–0.71 t CO_2e per year depending on the material, and that the emissions obtained with 4 cm of XPS were close to those obtained with 5 cm of stone wool. Based on the findings, it was concluded that the insulation thickness is more effective than the material type in reducing the emissions and that the same emission level can be achieved with thinner insulation using materials with lower thermal conductivity. In future studies, it is planned to repeat the calculations with the alculation method and climatic data of TS 825:2024 and to evaluate them for different climate zones.

This study was supported by the Scientific Research Projects Unit of Osmaniye Korkut Ata University under project number OKÜBAP-2024-PT2-017.

References

[1] European Parliament and Council of the European Union. (2024). Directive (EU) 2024/1275 of the European Parliament and of the Council of 24 April 2024 on the energy performance of buildings (recast). Official Journal of the European Union, L 2024/1275.

[2] TSE. (2024). TS 825 Binalarda ısı yalıtım kuralları [Thermal insulation requirements for buildings]. Turkish Standards Institution, Ankara.

[3] İZODER. (2025). TS 825:2024 Osmaniye il bilgi notu [TS 825:2024 provincial information note for Osmaniye]. İZODER Heat, Water, Sound and Fire Insulators Association. https://www.izoder.org.tr/ts825/osmaniye.pdf (Accessed: 27.09.2026)

[4] Schiavoni, S., D'Alessandro, F., Bianchi, F., & Asdrubali, F. (2016). Insulation materials for the building sector: A review and comparative analysis. Renewable and Sustainable Energy Reviews, 62, 988–1011. https://doi.org/10.1016/j.rser.2016.05.045

[5] Grazieschi, G., Asdrubali, F., & Thomas, G. (2021). Embodied energy and carbon of building insulating materials: A critical review. Cleaner Environmental Systems, 2, 100032. https://doi.org/10.1016/j.cesys.2021.100032

[6] Kavun, E. (2026). Binalarda Enerji Verimliliği Alanının Bibliyometrik Analizi ve Örnek Bir Kamu Binasının Enerji Performansının Değerlendirilmesi [Bibliometric analysis of the field of energy efficiency in buildings and evaluation of the energy performance of a sample public building] (Doctoral dissertation). Osmaniye Korkut Ata University, Graduate Education Institute, Osmaniye.

[7] Bektaş Ekici, B., Aytaç Gülten, A., & Aksoy, U. T. (2012). A study on the optimum insulation thicknesses of various types of external walls with

ISBN: 978-605-80193-6-2

respect to different materials, fuels and climate zones in Turkey. Applied Energy, 92, 211–217. https://doi.org/10.1016/j.apenergy.2011.10.008

[8] Çomaklı, K., & Yüksel, B. (2003). Optimum insulation thickness of external walls for energy saving. Applied Thermal Engineering, 23(4), 473–479. https://doi.org/10.1016/S1359-4311(02)00209-0

[9] Şişman, N., Kahya, E., Aras, N., & Aras, H. (2007). Determination of optimum insulation thicknesses of the external walls and roof (ceiling) for Turkey's different degree-day regions. Energy Policy, 35(10), 5151–5155. https://doi.org/10.1016/j.enpol.2007.04.037

[10] Uçar, A., & Balo, F. (2009). Effect of fuel type on the optimum thickness of selected insulation materials for the four different climatic regions of Turkey. Applied Energy, 86(5), 730–736. https://doi.org/10.1016/j.apenergy.2008.09.015

[11] Ozalp, C., Saydam, D. B., Çerçi, K. N., Hürdoğan, E., & Moran, H. (2019). Evaluation of a sample building with different type building elements in an energetic and environmental perspective. Renewable and Sustainable Energy Reviews, 115, 109386. https://doi.org/10.1016/j.rser.2019.109386

[12] Bolattürk, A. (2006). Determination of optimum insulation thickness for building walls with respect to various fuels and climate zones in Turkey. Applied Thermal Engineering, 26(11–12), 1301–1309. https://doi.org/10.1016/j.applthermaleng.2005.10.019

[13] Dombaycı, Ö. A. (2007). The environmental impact of optimum insulation thickness for external walls of buildings. Building and Environment, 42(11), 3855–3859. https://doi.org/10.1016/j.buildenv.2006.10.054

[14] Kaynaklı, Ö. (2012). A review of the economical and optimum thermal insulation thickness for building applications. Renewable and Sustainable Energy Reviews, 16(1), 415–425. https://doi.org/10.1016/j.rser.2011.08.006

[15] Kürekci, N. A. (2016). Determination of optimum insulation thickness for building walls by using heating and cooling degree-day values of all Turkey's provincial centers. Energy and Buildings, 118, 197–213. https://doi.org/10.1016/j.enbuild.2016.03.004

[16] TSE. (2008). TS 825 Binalarda ısı yalıtım kuralları [Thermal insulation requirements for buildings]. Turkish Standards Institution, Ankara.

ISBN: 978-605-80193-6-2

Feasibility of LPG Conversion for Passenger Cars in Türkiye: A Total Cost of Ownership Comparison of Gasoline and LPG Operation

Dr. Cemre POLAT[1*] Prof. Dr. Coşkun ÖZALP[2]

[1]Osmaniye Korkut Ata University, Düziçi Vocational School, Department of Motor Vehicles and Transportation Technologies, Osmaniye, Türkiye

[2]Osmaniye Korkut Ata University, Faculty of Engineering and Natural Sciences, Department of Energy Systems Engineering, Osmaniye, Türkiye

* Corresponding author e-mail: cemrepolat@osmaniye.edu.tr

***Abstract*—Approximately one third of the passenger cars registered in Türkiye run on LPG; however, only 1.1% of the new cars registered in 2025 were LPG-fuelled. It is seen from this picture that the LPG fleet is formed largely through aftermarket conversion and that the conversion decision retains its importance for the individual consumer. In this study, the feasibility of aftermarket LPG conversion was examined using a differential Total Cost of Ownership (TCO) model that compares the gasoline and LPG operation of the same vehicle. The model covers two separate scenarios for naturally aspirated port fuel injection (MPI) and turbocharged direct injection (DI) engines, the share of gasoline in LPG mode taken from the peer-reviewed literature, the volumetric consumption penalty of LPG, conversion costs compiled from market data, and September 2026 fuel prices. At an annual mileage of 15,000 km, the conversion of a vehicle with an MPI engine pays back at approximately 13.6 thousand km (10.9 months) and the five-year net present value reaches 113,406 TL. For the vehicle with a DI engine, the payback extends to 28.1 thousand km (22.5 months) and the five-year net present value remains at 63,972 TL. The minimum annual mileage required for the conversion to pay for itself over five years is 3,365 km in the MPI scenario and 6,955 km in the DI scenario. As a result of the one-at-a-time sensitivity analysis, it was determined that both scenarios produce a positive net present value across all parameter ranges examined; however, in the DI scenario the gasoline share and the kit cost in particular narrow the outcome considerably. The LPG-to-gasoline price ratio at which the conversion ceases to be economical over a five-year horizon was calculated as 0.72 for MPI and 0.62 for DI.**

***Keywords*— LPG conversion, autogas, total cost of ownership, payback period, gasoline share, direct injection.**

I. INTRODUCTION

Türkiye is among the countries in which the use of autogas (LPG) in passenger cars is most widespread in the world. Owing to the high special consumption tax and fuel taxes, the share of LPG-fuelled vehicles within the passenger car fleet exceeded 40% in the mid-2010s, and Özdemir and Tosun (2015) characterised this ratio as the highest LPG share in the world. According to TurkStat data, the share of LPG-fuelled passenger cars was 14.7% in 2004, reached its highest level of 41.5% in 2013, and as of December 2025 accounted for 30.1% of 17,373,581 passenger cars (TurkStat, 2026). Since this ratio is fairly close to the shares of diesel (32.6%) and gasoline (30.9%) passenger cars, it is seen that the vehicle fleet is in practice divided almost equally among three fuel types.

Of the 1,120,427 passenger cars registered in 2025, only 1.1% were registered as LPG-fuelled (TurkStat, 2026). From this difference between the high share within the fleet and the low share in new registrations, it is seen that the LPG fleet is formed largely not at the factory but through the aftermarket conversion of gasoline vehicles. In the investor presentation of one of the sector's LPG distributors, it is stated that approximately 160 thousand conversions are carried out annually in Türkiye and that the average conversion investment is recovered at approximately 15,000 km (Aygaz, 2016). Therefore, the LPG conversion decision continues to be a consumer decision that directly affects individual transport costs and national fuel demand in Türkiye.

The technical conditions that determine the economics of the conversion have changed in recent years; in the passenger car market, naturally aspirated port fuel injection (MPI) engines have increasingly begun to be replaced by turbocharged direct injection (DI) engines. In DI engines in particular, gasoline continues to be injected in order to cool the gasoline injectors and the combustion chamber even after switching to LPG mode (Grzelak and Taubert, 2021). Since the conversion systems suitable for these engines are considerably more expensive, it has become difficult to evaluate the conversion decision independently of the engine architecture.

In this context, the question “At how many kilometres does an aftermarket LPG conversion pay for itself in a passenger car driven 15,000 km per year, and to what extent does it reduce the total cost over a five-year ownership period?” constituted the main motivation of the study. C-segment sedans were taken as the reference vehicle class in the analysis, and there are two justifications for this choice. According to ODMD (2026) data, in 2025 the C segment accounted for 55.3% of the passenger car market with 599,790 units, while the sedan body type ranked second after SUVs with 247,005 units and a 22.8% share. In addition, this class is one of the vehicle classes to which LPG conversion has traditionally been applied most frequently. The model is structured so that it can also be applied to other vehicle classes by changing the fuel consumption and conversion cost inputs.

 ISBN: 978-605-80193-6-2

The contribution of the study to the literature is gathered under four points: (i) basing the gasoline share in LPG mode on peer-reviewed measurement data, (ii) treating MPI and DI engines as separate scenarios, (iii) establishing a differential TCO model containing only the cost items that differ, and (iv) presenting the results through criteria that directly guide the decision-maker, such as payback distance, break-even annual mileage and break-even price ratio. In order to preserve a brand- and model-independent academic approach, no reference was made to any automobile manufacturer or model in the study; the evaluations were carried out at the level of engine architecture.

II. THEORETICAL FRAMEWORK / LITERATURE REVIEW

A. LPG-Fuelled Passenger Car Use and Cost Analyses in Türkiye

One of the pioneering studies examining the cost structure of car ownership in Türkiye within the Total Cost of Ownership (TCO) framework was carried out by Özdemir and Tosun (2015). The researchers compared gasoline, diesel and LPG powertrains across five segments from A to E and for different annual mileage values. As a result of the study, the researchers determined that the one-off taxes at the purchase stage can rise to as much as 189% of the price excluding tax. It is also emphasised in the same study that fuel taxes place Türkiye among the countries with the highest fuel prices in the world. The researchers evaluated this tax structure as the main reason for the high LPG share within the vehicle fleet; Karamangil (2007), who examined the development of the sector through the case of Bursa, also revealed that the spread occurred largely with mechanically controlled first-generation systems and that the consumer motive was fundamentally the difference in pump prices.

Among the studies that directly address the conversion decision, the work carried out by Atik, Koruvatan, Yılmaz and Köse (2015) stands out. The researchers examined the consumer's fuel system conversion decision with a cost-benefit approach. As a result of the study, the researchers revealed the payback period of the conversion investment, the effect of increases in gasoline prices on this period, and how the conversion changes the consumer's point of indifference, using volume-profit analysis and gain matrix methods. In the international literature, in the study carried out by Letmathe and Suares (2017), the researchers applied a consumer-oriented TCO model to different powertrains in the German market and tested the sensitivity of the results to annual mileage profiles with Monte Carlo simulations. This situation reveals that the conversion decision should be evaluated depending on the intensity of use. However, in the existing Turkish studies the gasoline consumed in LPG mode is generally neglected and the market conditions prior to the spread of DI engines are reflected.

B. Fuel Properties of LPG and the Volumetric Consumption Difference

In the study carried out by Çelik, Aktaş and Özdalyan (2006), the researchers stated that the autogas used in Türkiye consists of a mixture of approximately 30% propane and 70% butane, and compared the basic properties of the fuels (specific gravity, lower heating value, octane number, etc.). Accordingly, the specific gravities of propane and butane are 0.51 and 0.58 kg/L and their lower heating values are 46.50 and 45.46 MJ/kg, respectively; the specific gravity of gasoline is 0.73–0.78 kg/L and its lower heating value is 44.03 MJ/kg. The research octane number is 111 for propane, 103 for butane and 95 for gasoline. When calculated with these values, the volumetric energy density of autogas with a 30%/70% mixture is approximately 25.6 MJ/L, whereas that of gasoline is approximately 33.2 MJ/L. When pure energy equivalence is taken as the basis, approximately 1.30 times, that is 30% more, LPG by volume is required for the same distance.

In practice, however, the actual volumetric consumption increase remains below this theoretical value. The high octane number of LPG and its formation of a more homogeneous mixture in the gas phase improve combustion efficiency. As a result, the increase calculated from energy equivalence is realised at a lower level under field conditions. The conversion system manufacturer Prins reports that consumption with LPG is approximately 15–25% higher than with gasoline depending on the propane-butane ratio (Prins Autogassystemen, n.d.-a); Masi (2012), who carried out experimental measurements on a passenger car fitted with an aftermarket conversion kit, also determined that the deterioration in performance was largely dependent on the setting of the vaporiser. In this study, a coefficient of 1.20 was adopted for the base scenario; the values of 1.30, representing the theoretical upper bound, and 1.15 were included in the sensitivity analysis. Çelik et al. (2006) additionally determined, in the measurements they carried out under real road conditions, that CO and HC emissions in LPG operation were generally lower than in gasoline operation.

C. The Gasoline-to-LPG Switching Condition and the Gasoline Share

Vehicles fitted with an aftermarket LPG system are always started on gasoline, and the switch to LPG occurs only after a certain temperature threshold has been exceeded. The physical reason for this threshold is that LPG draws the latent heat of vaporisation required for the transition from the liquid phase to the gas phase in the pressure reducer (regulator) from the engine coolant. If the switch is made before the coolant has warmed up sufficiently, the regulator cools excessively and freezes, and the engine runs irregularly. In practice the switching threshold is commonly set at around 40 °C, and in the control unit handbook of a conversion system manufacturer it is stated that the switch is always a function of the engine coolant temperature and takes place in two separate stages, according to a threshold of 40 °C in a cold vehicle and 50 °C in a warm vehicle (BRC, n.d.). In the experimental study carried out by Ceviz, Kaleli and Güner (2015), the researchers determined that the LPG temperature at the regulator outlet should be kept within the 30–40 °C band. Below this band, complete transition to the gas phase cannot be achieved, while above it the charge temperature rises and reduces volumetric efficiency and engine power. Therefore, the switching threshold on the order of 40 °C

ISBN: 978-605-80193-6-2

is consistent both with manufacturer practice and with experimental findings.

The time required to reach this threshold determines how much of the annual mileage will be driven on gasoline. The most comprehensive study directly measuring this subject belongs to Grzelak and Taubert (2021). The researchers tested a C-segment passenger car with a 1.6 L indirect gasoline injection engine and a sequential LPG system in the gas phase on a chassis dynamometer in a low-temperature chamber over the WLTC cycle; the principal results obtained are summarised in Table 1.

Table 1. LPG Switching Time and Gasoline Share by Engine Starting Temperature

Parameter	−10 °C	0 °C	+23 °C
Time to switch to LPG (s)	215	170	74
Distance covered on gasoline (m)	1,100	790	396
Mass share of gasoline in WLTC (%)	7.63	5.71	2.3
Gasoline share in an 8 km single trip (%)	20.9	16.2	7.7
Gasoline share in a 15 km single trip (%)	12.2	9.3	3.8

Source: Grzelak and Taubert (2021).

When Table 1 is examined, it is seen that the gasoline share is strongly dependent on two variables: the engine starting temperature and the length of a single trip. While it takes 74 seconds at +23 °C for the coolant to reach the switching threshold on the order of 40 °C, at −10 °C this extends to 215 seconds. The researchers calculated the ratio of the distance driven on gasoline to the total distance as 5.1–11% for 8 km trips and 2.7–5.8% for 15 km trips in two scenarios assuming commuting use only. It is emphasised by the researchers that these values are of an upper-bound nature because they do not include longer trips. Since winter temperatures in most of Türkiye are higher than in Warsaw, the +23 °C and 0 °C columns of the table are more representative of Turkish conditions. Within this framework, the base gasoline share for the MPI scenario was set at 5% and the sensitivity range at 3–12%.

In DI engines, however, the situation is different: Raslavičius et al. (2014), who classified LPG supply systems under five generations, reveal that gas-phase and liquid-phase architectures differ markedly from one another; Grzelak and Taubert (2021), on the other hand, state that in vehicles with direct gasoline injection the gasoline injectors continue to operate for cooling purposes even after switching to LPG mode and that a different test methodology is required for these vehicles. Although Ahn et al. (2020), who compared LPG with gasoline in turbocharged DI engines, determined reductions of 9% in carbon dioxide and 76% in particle number, no quantitative gasoline share data for DI engines could be found in the peer-reviewed literature, and for this reason manufacturer technical documents were consulted. One manufacturer reports that with its own control algorithm gasoline consumption remains at approximately 5–10%, whereas other systems on the market inject an additional 20–30% of gasoline (Alex LPG, 2022). The same manufacturer states that in more complex engines this ratio may rise to around 15% (Autogas ALEX, n.d.). Another manufacturer foresees gasoline consumption below 5% for its DI system (Prins Autogassystemen, n.d.-b). On the basis of these data, the base gasoline share in the DI scenario was taken as 15% and the sensitivity range as 5–30%. The fact that the DI values rely on manufacturer declarations is among the limitations of the study.

D. Regulatory, Safety and Environmental Dimensions

In Türkiye, LPG conversion is regulated as a modification procedure. Pursuant to the Regulation on the Manufacture, Modification and Assembly of Vehicles, published in the Official Gazette dated 21 October 2004 and numbered 25620, vehicles using LPG must meet all the requirements of UNECE Regulation R67.01 (Official Gazette, 2004); vehicles whose tank period has expired must present a Fuel System Assembly Certificate at inspection (TÜVTÜRK, n.d.). Since the service life of LPG tanks is generally applied as 10 years (Aygaz, n.d.; MMO, 2015), the five-year analysis horizon does not require tank replacement and the tank renewal cost is not included in the model. From an environmental standpoint, measurements carried out on Euro 6 gas-fuelled vehicles determined that CO_2 emissions in gas mode were 8–23% lower in the cold-start WLTC and 6–22% lower (14% on average) in real driving emissions (RDE) tests compared with gasoline mode (Giechaskiel et al., 2022). In the same study, CO and particle number emissions were also found to be lower in gas mode, whereas it was noted that ammonia (NH_3) emissions should be monitored; Kontses et al. (2020), who compared different powertrains under real driving conditions, also revealed that LPG provides a reduction in particle number of up to 93% compared with gasoline, but that this advantage disappears in older retrofitted vehicles.

As is also seen in the literature review, the studies addressing the economics of LPG conversion in Türkiye reflect the market conditions prior to the spread of DI engines and generally leave the gasoline consumed in LPG mode out of account. It is considered that this shortcoming would make the conversion decision appear more optimistic than it is for current engine architectures. In view of this situation, the necessity of carrying out a feasibility assessment in which the gasoline share is based on peer-reviewed measurement data and in which MPI and DI engine architectures are treated as separate scenarios constituted the motivation of this study. In addition, it is aimed with this study to put forward decision criteria that can directly guide the individual consumer and the policymaker.

III. METHOD

A. The Differential Total Cost of Ownership Approach

Since the gasoline and LPG operation of the same passenger car are compared in the study, as a natural consequence of this design a portion of the cost items remains exactly the same in the two scenarios: the purchase price, motor vehicle tax (since the engine displacement does not change), compulsory traffic insurance, basic periodic maintenance, vehicle inspection and motorway toll expenses. Since these

 ISBN: 978-605-80193-6-2

items enter the total cost of the two scenarios in equal amounts, they do not affect the difference between them, and therefore neither the payback period nor the net present value; they cancel each other out mathematically in the comparison. For this reason, instead of an absolute TCO calculation, the analysis was established as a differential (incremental) TCO model covering only the items that differ between the two scenarios. The items entering the model are the conversion investment, the fuel cost difference and the periodic maintenance specific to the LPG system. Since this approach frees the result from the magnitude of irrelevant items, it makes the economic rationale of the conversion decision directly visible; Langshaw et al. (2020), who established the TCO design for alternative fuels through investment, fuel and maintenance items, also applied a similar narrowing of scope.

The fuel cost per kilometre in gasoline operation was calculated with Equation (1) and in LPG operation with Equation (2):

$$c_B = (C / 100) \times P_B \quad (1)$$

$$c_L = (C / 100) \times [p \times P_B + (1 - p) \times k \times P_{LPG}] \quad (2)$$

Here, C denotes the combined consumption in gasoline mode (L/100 km), P_B and P_{LPG} the unit prices of gasoline and autogas (TL/L), p the share of gasoline within the fuel consumed in LPG mode, and k the volumetric consumption coefficient of LPG. Combined consumption was calculated as $C = 0.55\ C_{urb} + 0.45\ C_{ext}$ with weights of 55% urban and 45% extra-urban driving. The net saving per kilometre, the payback distance, the net present value (NPV) and the break-even annual mileage were obtained from Equations (3)–(6), respectively:

$$s = c_B - c_L - m \quad (3)$$

$$D^* = I_0 / s \quad (4)$$

$$NPV = -I_0 + \Sigma_{t=1\ldots5}\ (s \times L) / (1 + i)^t \quad (5)$$

$$L^* = I_0 / (s \times AF) \quad (6)$$

Here, m denotes the maintenance cost specific to the LPG system (TL/km), I_0 the conversion investment (TL), L the annual mileage, i the real discount rate, and $AF = \Sigma\ 1/(1 + i)^t$ the five-year annuity factor (4.046 for i = 7.5%). The conversion investment was taken as a cash flow occurring at the beginning (t = 0), and the annual net savings as cash flows occurring at year ends.

B. Scenarios and Assumptions

Of the two engine scenarios defined in the study, the MPI scenario represents a vehicle with a 1.4–1.6 L naturally aspirated port fuel injection engine converted with a gas-phase sequential LPG system, while the DI scenario represents a vehicle with a 1.0–1.5 L turbocharged direct injection engine converted with a DI-compatible gas-phase LPG system. In both scenarios the reference vehicle class is C-segment sedans; the fuel consumption values were based on field observations rather than synthetic test cycle values, so as to reflect the real-world use of this class under Turkish traffic conditions. The uncertainty relating to these values was addressed through the sensitivity analysis. For fuel prices, the Istanbul European Side pump prices of a distribution company dated September 2026 were used: unleaded gasoline 80.40 TL/L and autogas 34.99 TL/L. With these values the LPG-to-gasoline price ratio is at the level of 0.44. All assumptions are presented in Table 2.

Table 2. Model Assumptions

Parameter	MPI	DI	Source / Note
Annual mileage	15,000 km	15,000 km	Base scenario
Urban / extra-urban ratio	55 / 45	55 / 45	Assumption
Analysis period / real discount rate	5 years / 7.5%	5 years / 7.5%	Assumption
Gasoline consumption, urban (L/100 km)	8.5	8.0	Real-world assumption
Gasoline consumption, extra-urban (L/100 km)	5.5	5.3	Real-world assumption
Combined gasoline consumption (L/100 km)	**7.15**	**6.79**	**Calculated**
LPG volumetric consumption coefficient (k)	1.20	1.20	Section II-B
Gasoline share in LPG mode (p)	5%	15%	Section II-C
Kit + installation cost	27,500 TL	50,000 TL	Table 3
Official procedures (project, modification inspection, registration)	5,300 TL	5,300 TL	TÜVTÜRK (2026); assumption
LPG-specific maintenance (every 10,000 km)	2,000 TL	2,500 TL	Assumption
Gasoline price (September 2026)	80.40 TL/L	80.40 TL/L	Petrol Ofisi (2026)
Autogas price (September 2026)	34.99 TL/L	34.99 TL/L	Petrol Ofisi (2026)
Contribution of conversion to second-hand value	0 TL	0 TL	Conservative assumption

C. Determination of the Conversion Cost

The conversion costs were compiled from market sources for 2026 and from the recommended installation prices of conversion system distributors (Table 3). The DI kit prices given in euros were converted to Turkish lira at a rate of approximately 53 TL/€. Because the price difference between the sources is high, the values in the middle of the ranges were used in the base scenario, and the lower and upper bounds were included in the sensitivity analysis. Systems injecting LPG in the liquid phase were not included in the base scenario because of their costs, approximately two and a half times those of gas-phase DI systems, and were evaluated as a separate scenario.

Table 3. Sources of Conversion Cost (Kit + Installation, VAT Included)

System / scope	Price	Source
4-cylinder sequential system (MPI)	17,500–23,500 TL	Teknobilgi (2025)
Average conversion for 4-cylinder engines	25,000–40,000 TL	Otomerkezi (2026)
Average kit cost	35,000 TL	Araclo (2026)
DI-compatible gas-phase conversions	28,000–45,000 TL and above	Teknobilgi (2025)

ISBN: 978-605-80193-6-2

System / scope	Price	Source
DI gas-phase kits (4 cylinders)	710–1,350 € (≈37,600–71,600 TL)	Oto Tasarruf (2026)
Liquid-phase DI system (4 cylinders)	2,500 € (≈132,500 TL)	Oto Tasarruf (2026)
Base scenario: MPI / DI	**27,500 / 50,000 TL**	**This study**

D. Evaluation Criteria

Four criteria were used in evaluating the conversion decision, the first of which is the payback distance (D*). This criterion shows at how many kilometres the conversion investment is recovered and is converted into months using the annual mileage in the base scenario. The five-year net present value (NPV) expresses the net gain created by the conversion over the ownership period, taking into account the time value of money. The break-even annual mileage (L*) shows the lowest annual use that makes the NPV zero over a five-year period. Finally, the break-even price ratio is the LPG-to-gasoline price ratio that makes the NPV zero at 15,000 km per year while the gasoline price is held constant, and it measures the robustness of the decision against changes in price policy.

IV. FINDINGS

A. Annual Fuel Consumption and Fuel Cost

The fuel consumption and cost components for September 2026 prices and an annual mileage of 15,000 km are given in Table 4. In the MPI scenario, LPG operation reduces the annual fuel cost from 86,229 TL to 47,092 TL and provides a saving of 45.4%. In the DI scenario, the same saving ratio remains at the level of 40.6%. There are two reasons for this difference: the amount of gasoline consumed in LPG mode rises to approximately three times that of the MPI scenario, while the lower consumption of the turbocharged DI engine in gasoline mode reduces the base to which the saving is applied.

Table 4. Annual Fuel Consumption and Cost (15,000 km/year, September 2026 Prices)

Item	MPI	DI
Gasoline operation: annual gasoline (L)	1,072.5	1,017.8
Gasoline operation: annual fuel cost (TL)	86,229	81,827
LPG operation: annual gasoline (L)	53.6	152.7
LPG operation: annual LPG (L)	1,222.7	1,038.1
LPG operation: annual fuel cost (TL)	47,092	48,597
Annual fuel saving (TL)	**39,137**	**33,230**
Reduction in fuel cost	45.4%	40.6%
LPG-specific annual maintenance (TL)	−3,000	−3,750
Annual net saving (TL)	**36,137**	**29,480**
Net saving per kilometre (TL/km)	2.41	1.97

B. Payback Period and Five-Year Net Present Value

The decision criteria for both engine scenarios are summarised in Table 5. In the MPI scenario, the conversion investment of 32,800 TL is recovered at 13,615 km, that is, in approximately 10.9 months at an annual mileage of 15,000 km. It was determined that the five-year NPV reaches 113,406 TL, approximately 3.5 times the investment. It is seen that an annual use of 3,365 km is sufficient for the conversion to pay for itself over a five-year period. In the DI scenario, recovering the investment of 55,300 TL takes 28,138 km (22.5 months), the five-year NPV is 63,972 TL, and the break-even annual mileage rises to 6,955 km.

Table 5. Criteria Relating to the Conversion Decision (15,000 km/year, September 2026 Prices)

Criterion	MPI	DI
Conversion investment (TL)	32,800	55,300
Annual net saving (TL)	36,137	29,480
Payback distance (km)	**13,615**	**28,138**
Payback period (months)	10.9	22.5
Five-year net present value (TL)	**113,406**	**63,972**
Five-year nominal net gain (TL)	147,885	92,099
Break-even annual mileage (km)	3,365	6,955
Break-even LPG / gasoline price ratio	0.72	0.62

The cumulative net cash flow as a function of the distance covered after conversion is given in Figure 1. The points at which the curves intersect the horizontal axis represent the payback distances, while the slopes represent the net saving per kilometre. In the MPI scenario, at the end of five years of use (75,000 km) the cumulative net gain reaches 147,885 TL, whereas in the DI scenario it remains at the level of 92,099 TL.

ISBN: 978-605-80193-6-2

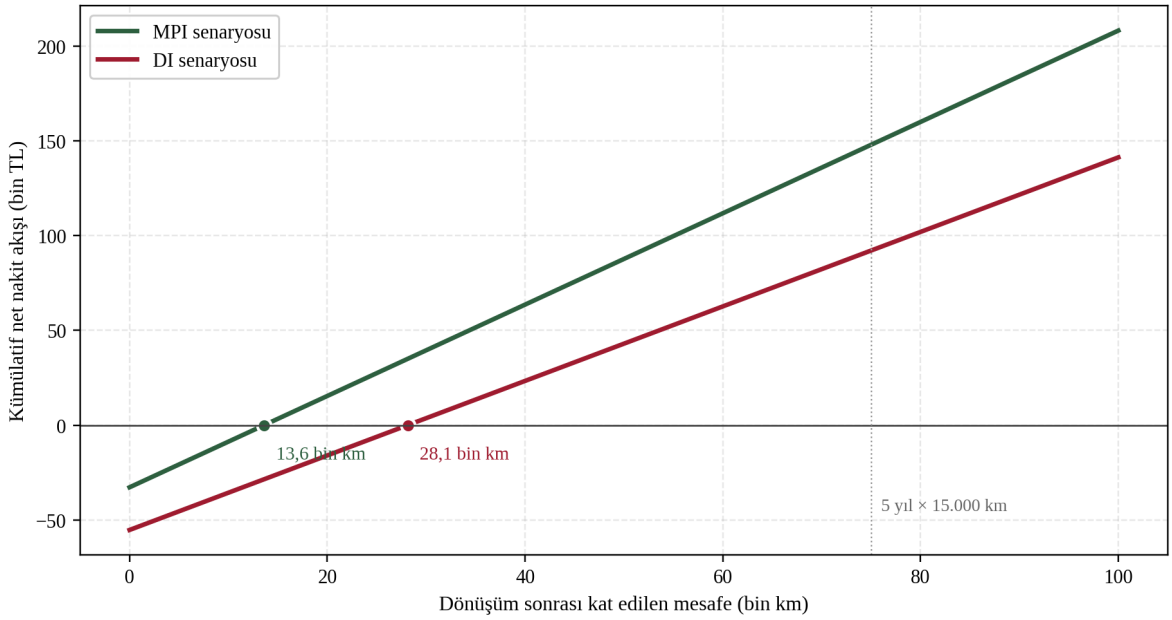


Fig. 1. Cumulative Net Cash Flow as a Function of Distance Covered After Conversion

C. Sensitivity to Annual Mileage

Since the economics of the conversion is directly dependent on the intensity of annual use, the payback periods and five-year NPV values in the range of 10,000–40,000 km per year are given in Table 6 and Figure 2. It is seen that both scenarios produce a positive NPV at all mileage values examined. However, at 10,000 km per year the payback period of the DI scenario rises to 33.8 months, that is, above three years. At intensive use of 25,000 km per year and above, on the other hand, the payback period falls below one and a half years even in the DI scenario.

Table 6. Payback Period and Five-Year NPV by Annual Mileage

Annual km	MPI: payback	MPI: 5-year NPV	DI: payback	DI: 5-year NPV
10,000	16.3 months	64,671 TL	33.8 months	24,214 TL
15,000	10.9 months	113,406 TL	22.5 months	63,972 TL
25,000	6.5 months	210,877 TL	13.5 months	143,486 TL
40,000	4.1 months	357,083 TL	8.4 months	262,758 TL

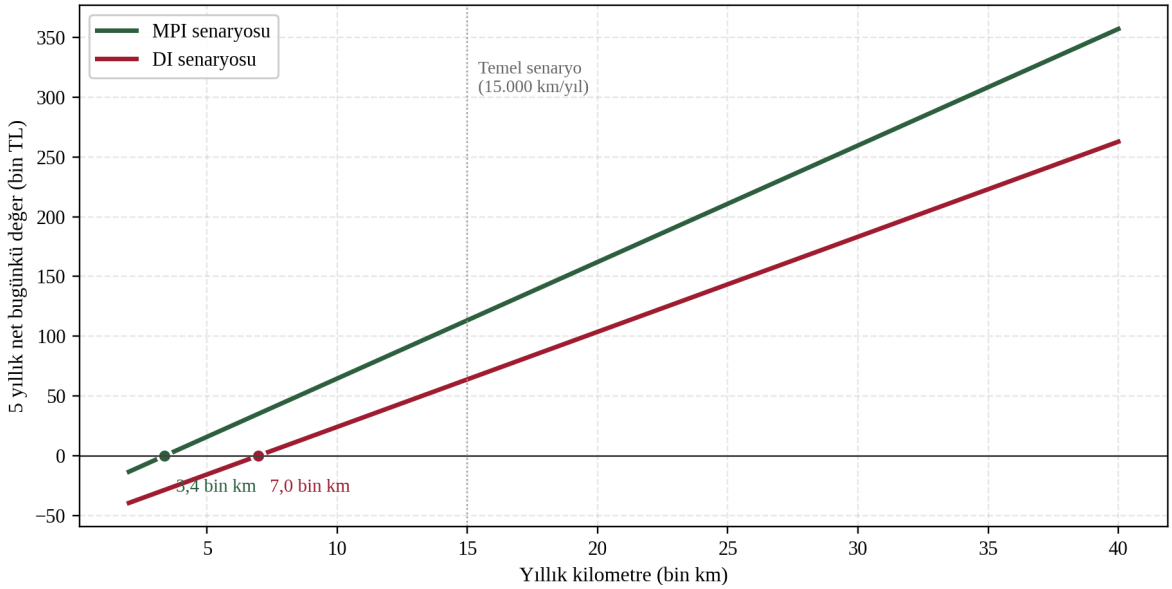


Fig. 2. Five-Year Net Present Value as a Function of Annual Mileage

ISBN: 978-605-80193-6-2

D. Parameter Sensitivity

The NPV and payback periods obtained by drawing each parameter in the base scenario to its lower and upper bounds while the others are held constant are given in Table 7. In the MPI scenario the most influential parameters are the kit cost and the LPG-to-gasoline price ratio. It was determined that the NPV does not fall below 87,000 TL at any of the parameter values examined. An increase in the gasoline share from 3% to 12% reduces the NPV by only approximately 15,000 TL. This situation reveals that the MPI scenario is robust against use dominated by short trips. In the DI scenario, on the other hand, the effect of the parameters is more pronounced: an increase in the gasoline share to 30% reduces the NPV to 40,246 TL, an increase in the kit cost to 72,000 TL reduces it to 41,972 TL, and a rise in the price ratio to 0.50 reduces it to 42,089 TL. It is considered that if these three parameters were to materialise unfavourably at the same time, the economic rationale of the DI scenario could disappear. This analysis, carried out on a one-at-a-time basis, does not cover the combined effect in question.

Table 7. One-at-a-Time Sensitivity Analysis (15,000 km/year, September 2026)

Parameter	Value	MPI: NPV / payback	DI: NPV / payback
Base scenario	—	**113,406 TL / 10.9 mo**	**63,972 TL / 22.5 mo**
Gasoline share in LPG mode	3% (MPI) / 5% (DI)	116,740 TL / 10.6 mo	79,789 TL / 19.9 mo
Gasoline share in LPG mode	12% (MPI) / 30% (DI)	101,739 TL / 11.8 mo	40,246 TL / 28.1 mo
LPG volumetric consumption coefficient	1.15	120,618 TL / 10.4 mo	70,095 TL / 21.4 mo
LPG volumetric consumption coefficient	1.30	98,982 TL / 12.1 mo	51,725 TL / 25.1 mo
Kit + installation cost	17,500 TL (MPI) / 28,000 TL (DI)	123,406 TL / 7.6 mo	85,972 TL / 13.6 mo
Kit + installation cost	40,000 TL (MPI) / 72,000 TL (DI)	100,906 TL / 15.0 mo	41,972 TL / 31.5 mo
LPG maintenance cost	×0.5	119,475 TL / 10.5 mo	71,558 TL / 21.2 mo
LPG maintenance cost	×1.5	107,337 TL / 11.4 mo	56,386 TL / 24.0 mo
LPG / gasoline price ratio	0.40	127,405 TL / 9.9 mo	75,858 TL / 20.5 mo
LPG / gasoline price ratio	0.50	87,634 TL / 13.2 mo	42,089 TL / 27.6 mo

In the scenario additionally evaluated for liquid-phase DI systems, a kit cost of approximately 132,500 TL and, in line with manufacturer declarations, a gasoline share of 5% were assumed. In this scenario the total investment rises to 137,800 TL, the payback distance extends to 61,906 km (49.5 months), and the five-year NPV remains just below the break-even point at −2,711 TL. It was determined that, for liquid-phase systems to be economical over a five-year horizon, annual use must exceed 15,301 km.

E. Break-Even Price Ratio

When the gasoline price is held constant at the September 2026 level (80.40 TL/L), the LPG-to-gasoline price ratio that makes the five-year NPV zero at an annual mileage of 15,000 km was calculated as 0.720 for the MPI scenario and 0.625 for the DI scenario. These values mean that the conversion would lose its economic rationale if the autogas price were to rise to the levels of 57.92 TL/L and 50.22 TL/L, respectively. The ratio realised in September 2026 (0.44) is markedly below both break-even values. This situation reveals that the current price structure leaves a wide safety margin in favour of the conversion.

V. DISCUSSION

From the findings obtained, it is seen that the economics of LPG conversion differs markedly according to the engine architecture. In a vehicle with a naturally aspirated MPI engine, the conversion pays for itself within one year at an average use such as 15,000 km per year and produces a high positive NPV across all parameter ranges examined. The payback distance of 13,615 km obtained largely coincides with the expectation of approximately 15,000 km long voiced in the sector (Aygaz, 2016). In DI engines, on the other hand, the payback distance doubles.

There are three sources of the difference in the DI scenario: the approximately twofold higher cost of DI-compatible kits, the consumption of gasoline in LPG mode as well for the cooling of the gasoline injectors, and the fact that turbocharged DI engines already have lower consumption in gasoline mode. These three factors reduce the rate of decrease in fuel cost from the level of 45% in the MPI scenario to the level of 41%. As a result of the sensitivity analysis, it was determined that in the DI scenario the decision is strongly dependent on the gasoline share performance of the system to be used and on the kit cost. Therefore, in these engines the conversion decision must be made on a vehicle and system basis rather than as a general rule. For liquid-phase systems, it may be said that they can be meaningful over a five-year horizon only for fleet and commercial users driving more than 15,301 km per year.

The findings relating to the gasoline share support the conclusion of Grzelak and Taubert (2021) that the length of a single trip is decisive. Since the vehicle runs on gasoline throughout the time required for the coolant to reach the switching threshold on the order of 40 °C, the gasoline share increases in short-distance urban use. In contrast, in the MPI scenario an increase in the gasoline share from 3% to 12% reduces the NPV by only about 13%. Since the temperate climate of Türkiye shortens the cold-start period, it is considered that a lower gasoline share may be in question compared with Warsaw conditions. On the other hand, in the literature the values relating to the volumetric consumption increase of LPG are distributed over a wide range such as 1.15–1.30. It is considered that the parameter in question needs to be clarified with field measurements under Turkish conditions.

Qualitative factors not included in the analysis, such as the warranty effect, the additional comprehensive insurance premium and luggage volume, are also

ISBN: 978-605-80193-6-2

important in the decision process. In vehicles still under warranty, a conversion carried out outside an authorised service may affect the engine warranty depending on the manufacturer's warranty policy. The LPG equipment must be declared to the comprehensive insurance policy and this declaration may lead to an increase in the premium; although the additional comprehensive insurance premium is an item that differs between the two scenarios, it was not included in the model because reliable data could not be found. Placing a toroidal tank in the spare wheel well may restrict luggage use. Since there are no systematic data on the value of LPG-fuelled vehicles in the second-hand market, a conservative assumption was adopted whereby the conversion investment is not reflected at all in the second-hand price; a favourable reflection would strengthen the findings further in favour of the conversion. From an environmental standpoint, since CO_2 emissions in gas mode are 6–23% lower than with gasoline (Giechaskiel et al., 2022), it is considered that the conversion provides a climate benefit in addition to its individual economic benefit.

As Özdemir and Tosun (2015) emphasised, the prevalence of LPG in Türkiye is largely a consequence of the tax structure. From the break-even price ratio findings obtained in this study, it is seen that if the LPG-to-gasoline price ratio were to exceed approximately 0.72 for MPI vehicles and approximately 0.62 for DI vehicles, the conversion would lose its economic rationale over a five-year horizon. Although the ratio of 0.44 in September 2026 is below these thresholds, a change to be made in the special consumption tax applied to autogas is a policy instrument that could directly affect conversion trends nationwide.

VI. CONCLUSION

In this study, the feasibility of aftermarket LPG conversion was examined with a differential TCO model comparing the gasoline and LPG operation of the same vehicle. The fact that the model bases the gasoline consumed in LPG mode on peer-reviewed measurement data and treats the MPI and DI engine architectures as separate scenarios constitutes the distinguishing aspect of the study. As a result of the study, it was determined that, with an annual mileage of 15,000 km and September 2026 prices, the conversion of a vehicle with an MPI engine pays for itself in approximately 10.9 months and the five-year NPV reaches 113,406 TL. For the vehicle with a DI engine, it was determined that the payback period extends to 22.5 months and the five-year NPV remains at the level of 63,972 TL.

When evaluated in terms of the user profile, it was determined that for owners of vehicles with naturally aspirated MPI engines the conversion is economical over a five-year horizon at any use above approximately 3,500 km per year. For owners of vehicles with turbocharged DI engines, on the other hand, the break-even threshold rises to approximately 7,000 km per year. Although the conversion is economical above this threshold, the decision needs to be supported by a system providing a low gasoline share and by a reasonable kit cost. For liquid-phase DI systems, it may be said that they should be evaluated only for fleet and commercial users driving more than 15,301 km per year.

The limitations of the study are as follows: the consumption values in gasoline mode and the maintenance costs specific to LPG are based on assumptions. The gasoline share data for DI engines were taken from manufacturer declarations rather than from the peer-reviewed literature. The conversion costs are distributed over a wide range; the additional comprehensive insurance premium, the warranty effect and the second-hand value difference, on the other hand, were not included in the model. The fuel prices are based on a single reference date and the analysis was carried out under the assumption of constant real prices. Since the sensitivity analysis is one-at-a-time, it does not cover cases in which the parameters move together.

The first of the three prominent suggestions for future studies is the determination of the gasoline share through field measurements in the different climatic regions of Türkiye by making use of the diagnostic records of LPG control units. The second is the measurement of the gasoline share for DI engines with a peer-reviewed methodology and the comparison of different system architectures. The third is the extension of the model so as to also cover the additional comprehensive insurance premium, the warranty effect and the second-hand value difference, and its testing with combined scenarios or with a Monte Carlo simulation of the kind applied by Langshaw et al. (2020) for heavy goods vehicles.

References

Ahn, S., Cho, J., Baek, S., Kim, K., Ko, J., Myung, C.-L., & Park, S. (2020). Comparative investigation of gaseous emissions and particle emission characteristics from turbo-charged direct injection (DI) engine with gasoline and LPG fuel depending on engine control parameters. International Journal of Automotive Technology, 21(2), 451–457. https://doi.org/10.1007/s12239-020-0042-7

Alex LPG. (2022). Direct petrol injection and LPG installation? https://alexlpg.com/en/direct-petrol-injection-and-lpg-installation/

Araclo. (2026). LPG dönüşüm rehberi 2026: Avantajları, dezavantajları ve bilmeniz gereken her şey [LPG conversion guide 2026: Advantages, disadvantages and everything you need to know]. https://www.araclo.com/blog/lpg-donusum-rehberi-2026-avantajlari-dezavantajlari-ve-bilmeniz-gereken-her-sey

Atik, M., Koruvatan, T., Yılmaz, B., & Köse, Y. (2015). Otomobillerdeki yakıt sistem değişikliğinin vergi ve tüketici yatırım politikalarına etkilerinin analizi [Analysis of the effects of fuel system conversion in automobiles on tax and consumer investment policies]. Verimlilik Dergisi, (2), 29–42.

Autogas ALEX. (n.d.). ALEX IDEA system. https://autogas-alex.com/en/alex-idea-system/

Aygaz. (2016). Yatırımcı sunumu [Investor presentation]. https://kurumsal.aygaz.com.tr/uploads/yatirimci-iliskileri/yatirimci-sunumlari/1b8aeccc_df52_4bca_9f79_c4c337d5470d__051619-pm.pdf

ISBN: 978-605-80193-6-2

Aygaz. (n.d.). LPG tankının kullanım ömrü ne kadardır? [What is the service life of an LPG tank?] https://kurumsal.aygaz.com.tr/otogaz/duyurular/lpg-tankinin-kullanim-omru-ne-kadardir

BRC Gas Equipment [BRC]. (n.d.). Handbook on Sequent 24 [Technical manual].

Ceviz, M. A., Kaleli, A., & Güner, E. (2015). Controlling LPG temperature for SI engine applications. Applied Thermal Engineering, 82, 298–305. https://doi.org/10.1016/j.applthermaleng.2015.02.059

Çelik, M. B., Aktaş, A., & Özdalyan, B. (2006). Gerçek yol şartlarında LPG ve benzinle çalışan iki taşıtın emisyon bakımından karşılaştırılması [Comparison of exhaust emissions of two vehicles running on LPG and gasoline under real road conditions]. Uludağ Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi, 11(1), 45–54.

Giechaskiel, B., Lähde, T., Clairotte, M., Suarez-Bertoa, R., Valverde, V., Melas, A. D., Selleri, T., & Bonnel, P. (2022). Emissions of Euro 6 mono- and bi-fuel gas vehicles. Catalysts, 12(6), 651. https://doi.org/10.3390/catal12060651

Grzelak, P., & Taubert, S. (2021). Consumption of gasoline in vehicles equipped with an LPG retrofit system in real driving conditions. Open Engineering, 11(1), 463–469. https://doi.org/10.1515/eng-2021-0044

Karamangil, M. İ. (2007). Development of the auto gas and LPG-powered vehicle sector in Turkey: A statistical case study of the sector for Bursa. Energy Policy, 35(1), 640–649. https://doi.org/10.1016/j.enpol.2006.01.004

Kontses, A., Triantafyllopoulos, G., Ntziachristos, L., & Samaras, Z. (2020). Particle number (PN) emissions from gasoline, diesel, LPG, CNG and hybrid-electric light-duty vehicles under real-world driving conditions. Atmospheric Environment, 222, 117126. https://doi.org/10.1016/j.atmosenv.2019.117126

Langshaw, L., Ainalis, D., Acha, S., Shah, N., & Stettler, M. E. J. (2020). Environmental and economic analysis of liquefied natural gas (LNG) for heavy goods vehicles in the UK: A well-to-wheel and total cost of ownership evaluation. Energy Policy, 137, 111161. https://doi.org/10.1016/j.enpol.2019.111161

Letmathe, P., & Suares, M. (2017). A consumer-oriented total cost of ownership model for different vehicle types in Germany. Transportation Research Part D: Transport and Environment, 57, 314–335.

Makina Mühendisleri Odası [Chamber of Mechanical Engineers, MMO]. (2015). LPG tanklarının kullanım ömrüne dikkat edilmeli [Attention must be paid to the service life of LPG tanks] [Press release]. https://mmo.org.tr/basin-aciklamasi/lpg-tanklarinin-kullanim-omrune-dikkat-edilmeli

Masi, M. (2012). Experimental analysis on a spark ignition petrol engine fuelled with LPG (liquefied petroleum gas). Energy, 41(1), 252–260. https://doi.org/10.1016/j.energy.2011.05.029

Official Gazette. (2004, 21 October). Araçların İmal, Tadil ve Montajı Hakkında Yönetmelik [Regulation on the manufacture, modification and assembly of vehicles] (No. 25620).

Otomerkezi. (2026). LPG hesaplama 2026: LPG'li araç tüketim ve maliyet hesabı [LPG calculation 2026: Consumption and cost calculation for LPG vehicles]. https://www.otomerkezi.net/yakit-tuketimi-hesaplama/lpg-hesaplama

Otomotiv Distribütörleri ve Mobilite Derneği [Automotive Distributors and Mobility Association, ODMD]. (2026). Otomobil ve hafif ticari araç pazarı, 2025 Ocak–Aralık [Passenger car and light commercial vehicle market, January–December 2025] [Press bulletin]. https://www.odmd.org.tr

Oto Tasarruf. (2026). FSI–TSI–GDI–TFSI direkt enjeksiyon LPG montaj fiyatları 2026 [FSI–TSI–GDI–TFSI direct injection LPG installation prices 2026]. https://www.ototasarruf.com/b-82/fsi-tsi-gdi-tfsi-direkt-enjeksiyon-lpg-montaj-fiyatlari-2026

Özdemir, E. D., & Tosun, H. B. (2015, 25–27 May). Total cost of ownership analysis for passenger cars in Turkey [Conference paper]. IAEE International Conference, Antalya.

Petrol Ofisi. (2026). Akaryakıt fiyatları [Fuel prices, September 2026]. https://www.petrolofisi.com.tr

Prins Autogassystemen. (n.d.-a). Service on your Prins DLM system. https://www.prinsautogas.com/en/service-your-prins-dlm-system

Prins Autogassystemen. (n.d.-b). VSI-3 DI – The revolutionary LPG system. https://www.prinsautogas.com/en/systems/vsi-3-di

Raslavičius, L., Keršys, A., Mockus, S., Keršienė, N., & Starevičius, M. (2014). Liquefied petroleum gas (LPG) as a medium-term option in the transition to sustainable fuels and transport. Renewable and Sustainable Energy Reviews, 32, 513–525. https://doi.org/10.1016/j.rser.2014.01.052

Teknobilgi. (2025). LPG dönüşüm fiyatları 2026 [LPG conversion prices 2026]. https://teknobilgi.net/lpg-donusum-fiyatlari/

Turkish Statistical Institute [TurkStat]. (2026). Motorlu kara taşıtları, Aralık 2025 [Motor land vehicles, December 2025] [News bulletin]. https://data.tuik.gov.tr

TÜVTÜRK. (2026). 2026 araç muayene ücretleri [2026 vehicle inspection fees]. https://www.tuvturk.com.tr

TÜVTÜRK. (n.d.). LPG hakkında sorular [Questions about LPG]. https://www.tuvturk.com.tr/musteri-hizmetleri/sikca-sorulan-sorular/lpg-hakkinda-sorular.aspx

 ISBN: 978-605-80193-6-2

Active Packaging Systems for Extending Food Shelf Life and Reducing Food Waste: A Sustainable Food Systems Perspective

Muhammed Yusuf Çağlar[1] Abdullah Çağlar[2]

[1]Ağrı İbrahim Çeçen University, Faculty of Health Sciences, Department of Nutrition and Dietetics, Ağrı, Türkiye.

[2]Kocaeli University, Faculty of Agriculture, Department of Agricultural Economics, Kocaeli, Türkiye

* Corresponding author e-mail: mcaglar@agri.edu.tr

***Abstract*— Food loss and waste are major challenges to sustainable food systems, and premature deterioration caused by microbial growth, oxidation, moisture changes, and physiological processes contributes to loss of food quality and usable shelf life. This narrative review evaluates the potential of active packaging systems to extend food shelf life and contribute to food waste prevention within a sustainable food systems framework. Recent peer reviewed literature on oxygen scavengers, moisture regulators, carbon dioxide based systems, ethylene scavengers, antimicrobial packaging, and antioxidant packaging was evaluated, with attention to preservation mechanisms, applications in different food matrices, sustainability implications, and technological limitations. The reviewed evidence indicates that appropriately designed active packaging can inhibit microbial proliferation, retard oxidative deterioration, regulate moisture and package atmosphere, and delay ripening and senescence. Experimental applications in meat, poultry, seafood, dairy products, and fresh produce demonstrate meaningful improvements in product stability and usable shelf life. However, extending shelf life does not necessarily demonstrate an actual reduction in food waste or an overall environmental benefit. Sustainability also depends on material and energy requirements, food contact safety, migration behavior, recyclability, biodegradability, end of life management, industrial scalability, and environmental impacts avoided through food preservation. Future research should integrate product specific shelf life validation with direct food waste measurements and life cycle assessment. Active packaging can support more sustainable food systems when effective preservation is achieved without disproportionate environmental, safety, or technological burdens.**

***Keywords*— Active packaging, food waste, shelf life, antimicrobial packaging, antioxidant packaging, sustainable food systems.**

I. INTRODUCTION

Food deterioration is a multidimensional process involving microbiological, chemical, enzymatic, physical, and physiological changes that progressively reduce product quality and may also affect safety. Microbial growth, lipid oxidation, pigment degradation, moisture transfer, respiration, ethylene mediated ripening, and oxygen exposure are among the processes that can limit commercial shelf life. Packaging has therefore evolved from passive containment and protection toward systems capable of actively modifying the package environment to preserve food quality [1].

Active packaging deliberately interacts with packaged food or its surrounding environment through absorption, release, or regulation of specific substances. Depending on the application, these systems can remove oxygen or excess moisture, regulate carbon dioxide, adsorb ethylene, or provide antimicrobial and antioxidant functionality [1, 2]. Oxygen scavengers, for example, reduce oxygen exposure that can accelerate oxidative deterioration and support aerobic microbial growth [3]. Antimicrobial and antioxidant systems can further delay deterioration, although their effectiveness depends on the packaging matrix, food composition, environmental conditions, and release behavior of the active agent [4, 5].

The relevance of active packaging extends beyond shelf life. By delaying deterioration, it may increase the time available for distribution, retail, storage, and consumption and thereby create opportunities to prevent avoidable food waste [2]. Nevertheless, shelf life extension should not automatically be interpreted as proof of food waste reduction or environmental sustainability. Additional materials, active substances, processing operations, migration behavior, recyclability, biodegradability, and end of life management must also be considered. This narrative review therefore evaluates major active packaging systems, their preservation mechanisms, applications in different food matrices, and their potential contribution to sustainable food systems.

II. FROM CONVENTIONAL TO ACTIVE FOOD PACKAGING

Conventional food packaging primarily provides containment and passive protection against environmental and physical factors. Passive protection may be insufficient for highly perishable foods in which microbial, oxidative, enzymatic, or physiological deterioration continues during storage. Active packaging addresses some of these limitations by incorporating components that intentionally interact with the food or package environment [1, 6].

Active technologies can broadly operate by removing undesirable compounds or releasing substances that contribute to preservation. Scavenging systems may remove oxygen, moisture, ethylene, or other undesirable components, whereas releasing systems may provide carbon dioxide, antimicrobial agents, antioxidants, or other functional compounds [1, 2]. The selected strategy should correspond to the predominant deterioration mechanism of the target food. In releasing systems, film properties, food characteristics, environmental conditions, and properties of the active compound affect migration and

ISBN: 978-605-80193-6-2

release [4]. Active packaging should therefore be optimized for the specific food package combination rather than treated as a universal preservation technology.

III. MAJOR ACTIVE PACKAGING SYSTEMS

A. Oxygen Scavenging Systems

Oxygen can contribute to lipid oxidation, pigment deterioration, undesirable flavor development, degradation of oxygen sensitive components, and growth of aerobic microorganisms. Oxygen scavengers are designed to remove oxygen from the package headspace or dissolved phase. Iron based systems, platinum group metals, ascorbic acid, tocopherols, unsaturated hydrocarbons, enzymes, and microorganisms have all been investigated for this purpose [3].

B. Moisture and Carbon Dioxide Control

Moisture accumulation can promote condensation, microbial proliferation, and textural deterioration, whereas excessive moisture loss can cause dehydration and sensory quality loss. Absorbent pads, desiccants, and related materials can regulate water accumulation or humidity [1, 6]. Carbon dioxide emitters can help maintain atmospheres that suppress selected spoilage microorganisms, while absorbers can be useful when excessive carbon dioxide accumulation is undesirable [1, 2].

C. Ethylene Scavenging Systems

Ethylene is involved in ripening and senescence in many fruits and vegetables. Its accumulation may accelerate softening, color change, and other physiological transformations in ethylene responsive commodities. Ethylene scavengers can remove or degrade ethylene and thereby delay postharvest deterioration, although effectiveness remains commodity specific [1, 2].

D. Antimicrobial and Antioxidant Packaging

Antimicrobial packaging incorporates active substances into films, coatings, pads, or other structures to inhibit microorganisms. Investigated agents include bacteriocins, essential oils, organic acids, enzymes, and other functional compounds. Performance depends not only on intrinsic antimicrobial activity but also on film structure, food composition, storage conditions, and release behavior [4]. Antioxidant packaging aims to retard oxidative reactions responsible for lipid oxidation, discoloration, undesirable flavor development, and degradation of oxidation sensitive constituents [5].

IV. MECHANISMS OF SHELF LIFE EXTENSION

Active packaging extends shelf life by targeting processes that limit food stability. Antimicrobial systems can reduce or delay microbial proliferation by delivering active compounds to the food surface or package environment, while oxygen scavengers and antioxidant systems can reduce oxygen availability or inhibit oxidative reactions [3–5]. Moisture and gas regulation can influence microbial and physicochemical stability, and ethylene scavenging can delay ripening and senescence in responsive produce [1, 2].

Controlled release is particularly important for antimicrobial and antioxidant packaging. An active compound released too rapidly may be depleted before the end of storage, whereas insufficient release may fail to achieve an effective concentration. Release behavior depends on the packaging matrix, food composition, environmental conditions, and physicochemical properties of the active substance [4, 5]. Thus, successful active packaging requires matching the active function and its kinetics to the dominant deterioration pathway of the target food.

V. APPLICATIONS IN DIFFERENT FOOD MATRICES

A. Meat and Poultry Products

Fresh meat and poultry are susceptible to microbial deterioration, lipid and protein oxidation, discoloration, and moisture related quality losses. Wongphan et al. developed thermoplastic starch and polybutylene adipate co terephthalate films containing sodium erythorbate and sodium hexametaphosphate for beef. The films inhibited selected bacteria and delayed lipid oxidation and protein deterioration; beef quality was evaluated during nine days of chilled storage, with improved redness and color stability particularly evident during the first six days [7].

Mukkadan et al. developed an arrowroot starch film reinforced with nanocellulose from Borassus flabellifer residues and containing neem oil. The active formulation improved antimicrobial and antioxidant performance and maintained chicken meat freshness for up to 12 days compared with approximately six days for the control film [8]. Sasidharan et al. reported that a chitosan film containing biogenic zinc oxide nanoparticles maintained acceptable poultry meat quality for seven days, compared with five days for the chemogenic zinc oxide formulation and four days for the chitosan control [9]. Caprylic acid incorporated into poly lactic acid films has also been investigated against Salmonella contamination on chicken breast and beef [10].

B. Fish and Seafood Products

Fish and seafood are highly perishable because microbial growth, endogenous enzymatic reactions, water migration, protein degradation, and lipid oxidation can occur simultaneously. Eshaghi et al. developed carboxymethyl cellulose and myrrh gum films containing titanium dioxide nanoparticles and dill essential oil for common carp and reported improved refrigerated preservation [11]. Chen et al. developed chitosan and polyvinyl alcohol films incorporating ginger essential oil loaded bacterial cellulose for sea bass; the optimized system extended refrigerated shelf life by at least six days compared with the control [12].

C. Dairy Products

Cheese is an important target because microbial growth, oxidation, moisture transfer, and surface deterioration can limit stability. Sharma et al. developed poly lactic acid and polybutylene adipate co

ISBN: 978-605-80193-6-2

terephthalate films containing titanium dioxide and cinnamon essential oil and reported enhanced cheese preservation during 12 days at 4 degrees Celsius [13]. Perera et al. developed sodium alginate and gelatin films containing nanoclay and clove essential oil, extending cheddar cheese shelf life to up to 15 days at 4 degrees Celsius [14]. Abdel Rehim et al. also demonstrated the preservation potential of Lepidium sativum gum and polyvinyl alcohol active films for cheddar cheese [15].

D. Fruits and Vegetables

Fresh produce remains metabolically active after harvest. Respiration, transpiration, enzymatic activity, ripening, and senescence continue during storage. Wei et al. developed platinum loaded dendritic mesoporous silica as an ethylene scavenger. In banana application, the system delayed softening and browning, reduced weight loss, and maintained freshness for 14 days [16]. These studies demonstrate that the optimal active function differs among food matrices and should target the mechanism that actually limits usable shelf life. The representative applications discussed above are summarized in Table I.

TABLE I. REPRESENTATIVE APPLICATIONS OF ACTIVE PACKAGING IN DIFFERENT FOOD MATRICES

Food matrix and active system	Main reported effect	Ref.
Beef, TPS/PBAT + sodium erythorbate/hexametaphosphate	Delayed oxidation and quality deterioration; improved color stability during chilled storage	[7]
Chicken meat, arrowroot starch + nanocellulose + neem oil	Maintained freshness up to 12 d vs. about 6 d for control	[8]
Poultry meat,chitosan + biogenic ZnO nanoparticles	Acceptable quality up to 7 d vs. 4 d for chitosan control	[9]
Chicken/beef, PLA + caprylic acid	Antimicrobial activity against Salmonella contamination	[10]
Common carp, CMC/myrrh gum + TiO2 + dill essential oil	Improved refrigerated preservation and quality stability	[11]
Sea bass, chitosan/PVA + ginger essential oil-loaded bacterial cellulose	Extended refrigerated shelf life by at least 6 d	[12]
Cheese, PLA/PBAT + TiO2 + cinnamon essential oil	Enhanced preservation during 12 d at 4 °C	[13]
Cheddar cheese, sodium alginate/gelatin + nanoclay + clove essential oil	Extended shelf life up to 15 d at 4 °C	[14]
Cheddar cheese, Lepidium sativum gum/PVA active film	Demonstrated preservation potential	[15]
Banana, Pt-loaded dendritic mesoporous silica ethylene scavenger	Delayed ripening, softening and browning; maintained freshness for 14 d	[16]

VI. ACTIVE PACKAGING AND FOOD WASTE REDUCTION

Active packaging can potentially contribute to food waste prevention by delaying microbial, chemical, physical, or physiological deterioration and extending the period during which food remains acceptable for distribution, retail, storage, and consumption [2]. Experimental studies demonstrate shelf life and quality preservation benefits across multiple food categories [7–16]. However, shelf life extension and food waste reduction are not equivalent outcomes. Most active packaging studies measure microbial counts, oxidation indices, physicochemical characteristics, sensory quality, or shelf life rather than the amount of food actually discarded.

The relationship can therefore be interpreted as a sequence in which active packaging delays deterioration, extends usable shelf life, increases the opportunity for distribution and consumption, and may consequently reduce avoidable food waste. Whether this potential becomes an actual reduction also depends on retail management, consumer behavior, storage temperature, package size, product demand, and other supply chain factors. Direct waste measurements under realistic retail and household conditions are therefore needed to quantify the food waste benefit.

VII. SUSTAINABILITY CONSIDERATIONS

The sustainability of active packaging cannot be determined solely from the amount or type of packaging material. Environmental assessment should consider direct impacts associated with packaging production and disposal together with the environmental consequences of food losses that packaging may prevent [2]. Gutierrez et al. showed in a packaged cheesecake life cycle assessment that shelf life extension and avoided food losses can materially influence overall environmental performance [17]. Casson et al. similarly demonstrated for beef that conclusions based on packaging alone can change when potential food waste associated with different shelf lives is included in the life cycle assessment [18].

Bio based and biodegradable materials provide opportunities for active packaging, but these descriptors do not automatically demonstrate environmental superiority. Feedstock production, processing energy, functional performance, material quantity, actual biodegradation conditions, and end of life infrastructure must also be considered. Active compounds, coatings, nanoparticles, and multilayer structures can additionally affect recyclability and disposal. Food contact safety is equally important because migration depends on film properties, food characteristics, environmental conditions, and properties of the active substance [4].

Life cycle thinking therefore requires assessment of the complete food package system rather than packaging material alone. The relevant comparison is whether the preservation benefit and avoided food loss compensate for the additional environmental burdens introduced by active functionality [17, 18].

VIII. CHALLENGES AND FUTURE PERSPECTIVES

Transition from laboratory development to widespread commercial application remains challenging. Successful systems require preservation efficacy together with suitable mechanical and barrier properties, food contact safety, manufacturing feasibility, economic viability, and compatibility with existing packaging operations [2, 6]. Laboratory activity may change after incorporation into polymer matrices and contact with real foods because composition, temperature, contact time, and polymer characteristics influence diffusion and release [4]. Product specific

ISBN: 978-605-80193-6-2

validation under realistic storage conditions is therefore essential.

Industrial scalability is another barrier because many experimental films are produced using laboratory casting, coating, encapsulation, or nanostructuring techniques. Commercial production requires reproducibility, high speed processing compatibility, adequate storage stability, reliable incorporation of active agents, and acceptable cost [6]. Future research should also quantify actual food waste prevention rather than relying only on shelf life as a proxy. The most relevant next generation systems will be those that demonstrate effective preservation, food contact safety, industrial scalability, economic feasibility, measurable waste prevention, and favorable life cycle performance.

IX. CONCLUSION

Active packaging provides a versatile approach for controlling several mechanisms responsible for food deterioration. Oxygen scavengers, moisture regulators, carbon dioxide based systems, ethylene scavengers, antimicrobial packaging, and antioxidant packaging can preserve quality and extend shelf life when matched appropriately to the target food [1–6]. Applications in meat, poultry, fish, dairy products, and fresh produce demonstrate meaningful preservation effects [7–16].

From a sustainable food systems perspective, shelf life extension can increase the time available for distribution, retail, storage, and consumption and may contribute to reducing avoidable food waste. However, shelf life extension should not be considered equivalent to demonstrated food waste reduction. Environmental performance should also be assessed at the level of the complete food package system because additional materials and processing may introduce burdens related to production, migration, recycling, biodegradation, and end of life management [17, 18]. Future development should integrate product specific shelf life validation, controlled release, food contact safety, scalable manufacturing, direct food waste measurements, and life cycle assessment.

References

[1] Yildirim, S., Röcker, B., Pettersen, M. K., Nilsen-Nygaard, J., Ayhan, Z., Rutkaite, R., et al. (2018). Active packaging applications for food. Comprehensive Reviews in Food Science and Food Safety, 17(1), 165–199. https://doi.org/10.1111/1541-4337.12322

[2] Alves, J., Gaspar, P. D., Lima, T. M., & Silva, P. D. (2023). What is the role of active packaging in the future of food sustainability? A systematic review. Journal of the Science of Food and Agriculture, 103(3), 1004–1020. https://doi.org/10.1002/jsfa.11880

[3] Dey, A., & Neogi, S. (2019). Oxygen scavengers for food packaging applications: A review. Trends in Food Science & Technology, 90, 26–34. https://doi.org/10.1016/j.tifs.2019.05.013

[4] García-Anaya, M. C., Sepúlveda, D. R., & Acosta-Muñiz, C. H. (2025). Contributing factors to the migration of antimicrobials in active packaging films. Food Research International, 200, 115514. https://doi.org/10.1016/j.foodres.2024.115514

[5] Vasile, C., & Baican, M. (2021). Progresses in food packaging, food quality, and safety—Controlled-release antioxidant and/or antimicrobial packaging. Molecules, 26(5), 1263. https://doi.org/10.3390/molecules26051263

[6] Kadirvel, V., Palanisamy, Y., & Ganesan, N. D. (2025). Active packaging system—An overview of recent advances for enhanced food quality and safety. Packaging Technology and Science, 38(2), 145–162. https://doi.org/10.1002/pts.2863

[7] Wongphan, P., Harnkarnsujarit, N., Nerín, C., & Paiva, R. (2024). Sodium erythorbate and hexametaphosphate loaded TPS/PBAT films: A multifunctional packaging for extending beef shelf-life. Food Packaging and Shelf Life, 46, 101396. https://doi.org/10.1016/j.fpsl.2024.101396

[8] Mukkadan, M. J., Preetha, R., & Sreejit, V. (2024). Development of active packaging film reinforced with nano cellulose extracted from Palmyra palm residues for prolonging the shelf life of chicken meat. International Journal of Food Science & Technology, 59(3), 2033–2041. https://doi.org/10.1111/ijfs.16706

[9] Sasidharan, S., Tey, L.-H., Djearamane, S., Ab Rashid, N. K. M., Rajeshwari, P. A., Rajendran, V., et al. (2024). Innovative use of chitosan/ZnO NPs bio-nanocomposites for sustainable antimicrobial food packaging of poultry meat. Food Packaging and Shelf Life, 43, 101298. https://doi.org/10.1016/j.fpsl.2024.101298

[10] Her, E., Han, S., & Ha, S.-D. (2024). Development of poly(lactic acid)-based natural antimicrobial film incorporated with caprylic acid against Salmonella biofilm contamination in the meat industry. International Journal of Food Microbiology, 425, 110871. https://doi.org/10.1016/j.ijfoodmicro.2024.110871

[11] Eshaghi, R., Mohsenzadeh, M., & Ayala-Zavala, J. F. (2024). Bio-nanocomposite active packaging films based on carboxymethyl cellulose, myrrh gum, TiO2 nanoparticles and dill essential oil for preserving fresh-fish meat quality. International Journal of Biological Macromolecules, 263, 129991. https://doi.org/10.1016/j.ijbiomac.2024.129991

[12] Chen, X., Lan, W., & Xie, J. (2024). Characterization of active films based on chitosan/polyvinyl alcohol integrated with ginger essential oil-loaded bacterial cellulose and application in sea bass packaging. Food Chemistry, 441, 138343. https://doi.org/10.1016/j.foodchem.2023.138343

[13] Sharma, S., Byrne, M., Perera, K. Y., Duffy, B., Jaiswal, A. K., & Jaiswal, S. (2023). Active film packaging based on bio-nanocomposite TiO2 and cinnamon essential oil for enhanced preservation of cheese quality. Food Chemistry, 405, 134798. https://doi.org/10.1016/j.foodchem.2022.134798

 ISBN: 978-605-80193-6-2

[14] Perera, K. Y., Jaiswal, A. K., & Jaiswal, S. (2023). Extending cheese shelf-life using eco-friendly sodium alginate-gelatin films reinforced with nanoclay. Food Bioscience, 56, 103304. https://doi.org/10.1016/j.fbio.2023.103304

[15] Abdel Rehim, M., Zahran, H. A., & Al-Moghazy, M. (2023). Synthesis of active packaging films from Lepidium sativum gum/polyvinyl alcohol composites and their application in preserving cheddar cheese. Scientific Reports, 13, 1647. https://doi.org/10.1038/s41598-023-28173-3

[16] Wei, H., Li, L., Zhang, T., Seidi, F., & Xiao, H. (2023). Platinum-loaded dendritic mesoporous silica as novel ethylene scavenger to extend shelf life of banana. Food Chemistry, 424, 136415. https://doi.org/10.1016/j.foodchem.2023.136415

[17] Gutierrez, M. M., Meleddu, M., & Piga, A. (2017). Food losses, shelf life extension and environmental impact of a packaged cheesecake: A life cycle assessment. Food Research International, 91, 124–132. https://doi.org/10.1016/j.foodres.2016.11.031

[18] Casson, A., Giovenzana, V., Frigerio, V., Zambelli, M., Beghi, R., Pampuri, A., et al. (2022). Beyond the eco-design of case-ready beef packaging: The relationship between food waste and shelf-life as a key element in life cycle assessment. Food Packaging and Shelf Life, 34, 100943. https://doi.org/10.1016/j.fpsl.2022.100943

ISBN: 978-605-80193-6-2

The Role of "Best Before" and "Use By" Dates in Food Waste: Consumer Perception, Food Safety, and Sustainability

Muhammed Yusuf Çağlar[1] Abdullah Çağlar[2]

[1]Ağrı İbrahim Çeçen University, Faculty of Health Sciences, Department of Nutrition and Dietetics, Ağrı, Türkiye.

[2]Kocaeli University, Faculty of Agriculture, Department of Agricultural Economics, Kocaeli, Türkiye

* Corresponding author e-mail: mcaglar@agri.edu.tr

***Abstract*— Food date marking can influence whether consumers retain or discard foods, yet printed dates are not interpreted uniformly. This study integrates evidence on "Best Before" and "Use By" labels, consumer decision making, microbiological shelf-life assessment, food waste, communication interventions, and intelligent packaging. Consumer studies distinguish a quality-oriented interpretation of "Best Before" from the stronger safety concerns elicited by "Use By", while also documenting substantial confusion between these meanings. Such misunderstanding may contribute to the premature disposal of foods that remain suitable for consumption or, conversely, to inappropriate consumption decisions when safety related dates are exceeded. Microbiological evidence shows that shelf life is product and condition specific and that sensory acceptability cannot be equated with microbial status. Studies of foods at and beyond marked dates further demonstrate that results obtained for one matrix, packaging system, storage condition, or temperature regime should not be generalized to other products. Communication studies indicate that consumers may focus more strongly on the printed date than on accompanying terminology, while Look, Smell, Taste cues have shown limited effects on actual overdate product use. Time temperature indicators and freshness monitoring systems may complement static dates by providing information related to storage history or product condition, but these technologies require product specific calibration, validation, and appropriate consumer interpretation. Overall, reducing date-related food waste requires an integrated approach combining clear and consistent date label communication, scientifically justified shelf-life determination, appropriate household food management practices, and validated technological interventions. Such strategies should aim to reduce avoidable food disposal while maintaining a clear distinction between quality-related guidance and safety related date marking, thereby ensuring that food waste reduction initiatives do not weaken consumer food safety protection.**

***Keywords*— Food date marking, best before, use by, food waste, shelf life, consumer behaviour, food safety, intelligent packaging.**

I. INTRODUCTION

Food loss and waste represent a substantial inefficiency in contemporary food systems because discarded food also embodies losses of agricultural land, water, energy, labor, processing inputs, packaging, transport, and household expenditure [1, 2]. At household level, disposal decisions are influenced by multiple factors, including shopping and storage practices, sensory expectations, perceived food-safety risk, product perishability, and interpretation of date labels. Date marking is therefore not a single cause of food waste, but it can become an important decision cue when consumers use the printed date as the principal indicator of whether a food should be retained or discarded.

The central difficulty is that different date labels communicate different types of information and are not interpreted uniformly by consumers. In consumer research, "Best Before" is predominantly interpreted as a quality or minimum-durability cue, whereas "Use By" elicits stronger safety-related concerns. European studies nevertheless document persistent confusion between these meanings, and such confusion can contribute to unnecessary disposal or inappropriate food-use decisions [3, 4].

This narrative review examines food date marking through four connected perspectives: label meaning and consumer interpretation, microbiological shelf-life determination, food-waste prevention, and emerging communication and intelligent-packaging approaches. The evidence base is intentionally interdisciplinary and draws on peer reviewed consumer behaviour experiments and surveys, product specific microbiological shelf-life studies, and reviews of intelligent packaging and food waste interventions. Priority was given to recent studies directly addressing date-label interpretation, disposal behaviour, shelf-life evidence, or condition-responsive packaging, while older studies were retained where they provided relevant foundational evidence. Because the review is narrative rather than systematic, the literature is synthesized to explain mechanisms, contrasts, and product-specific findings rather than to provide an exhaustive search record or pooled effect estimate. Particular attention is given to the distinction between sensory quality and microbiological status and to the limits of generalizing findings across food matrices, packaging systems, and storage regimes.

II. MEANING and INTERPRETATION of FOOD DATE LABELS

"Best Before" and "Use By" labels convey different signals to consumers. "Best Before" is generally interpreted as a quality or minimum-durability cue, whereas "Use By" produces stronger safety and quality concerns and can reduce willingness to consume a product [5]. European consumer research likewise

ISBN: 978-605-80193-6-2

shows that many consumers recognize the distinction, but a substantial minority still misunderstand or conflate the two labels [3, 4]. These findings are important for food-waste prevention because disposal decisions depend not only on the printed date but also on the meaning consumers assign to it.

The wording of a date label can itself alter perceived risk. In a controlled experiment, Gong et al. [5] compared several commonly used date-label expressions and found that "Use By" generated greater safety and quality concerns than "Best if Used By". These concerns mediated willingness to consume the tested food, while information indicating normal color and odor or appropriate storage increased willingness to consume. Date-label terminology should therefore be evaluated not only as technical information but also as a behavioural signal that shapes consumer interpretation.

III. CONSUMER UNDERSTANDING, RISK PERCEPTION, AND BEHAVIOURAL MECHANISMS

Consumers do not always translate date-label terminology into accurate disposal decisions. In a six-country European survey involving 1507 consumers, 75.2% understood the distinction between "Use By" and "Best Before," while 86.7% and 82.3% correctly understood the individual meanings of "Best Before" and "Use By," respectively [4]. Greater understanding was associated with lower reported household food waste, but knowledge alone did not explain behaviour. Earlier Polish evidence similarly showed persistent confusion, with 42% of respondents considering "Use By" and "Best Before" to have the same meaning [3].

This knowledge behaviour gap is important because date labels interact with product type and individual decision preferences rather than producing a uniform response. In an experimental study using deli meat and spaghetti sauce, Wilson and Miao [6] found that the "Use by" treatment generated more premeditated waste than the "Best by" treatment for both products. Greater loss aversion was also associated with higher premeditated waste, while among highly loss-averse participants the difference between the two label treatments was not statistically significant. These findings indicate that label simplification alone is unlikely to produce the same waste-reduction effect across foods and consumer groups.

Date information can also influence expected and experienced product evaluation. In a consumer sensory study, Sielicka-Różyńska and Samotyja [7] showed that "Best Before" information affected expected food liking and examined how date-related expectations interacted with actual liking. This demonstrates that a printed date can shape product evaluation before consumption and should therefore be considered a behavioural cue as well as a technical item of package information.

IV. MICROBIOLOGICAL BASIS of DATE MARKING and SHELF-LIFE DETERMINATION

Scientifically defensible shelf-life assessment requires more than assigning a fixed number of days to a broad food category. Microbial spoilage and safety are influenced by the characteristics of the food matrix, the initial microbial population, processing and packaging conditions, and the environment encountered during storage. These factors interact, so shelf-life evidence should be interpreted for the specific food and storage system rather than generalized across unrelated products [8, 9].

Intrinsic factors relevant to microbial stability include water activity (a_w), pH, and nutrient availability, whereas important extrinsic factors include storage temperature, oxygen availability, humidity, and package atmosphere. Karanth et al. [8] emphasized that microbial spoilage results from the combined influence of these conditions and the surrounding microflora. Product-specific studies illustrate this interaction: in modified-atmosphere-packed chicken, bacterial populations differed according to storage time and temperature, with higher counts observed under the warmer storage condition [9].

Because the same environmental condition can affect microorganisms differently across food matrices, no single physicochemical parameter should be treated as a universal predictor of shelf life. The relevant organism, initial contamination level, food composition, storage temperature, packaging environment, and microbial interactions all contribute to the observed spoilage trajectory. Quantitative microbial spoilage risk assessment has therefore been proposed as a probabilistic framework for integrating microbial behavior, variability, and uncertainty when estimating spoilage risk along the food chain [8].

This distinction is also important for date marking. A printed date is a static information cue, whereas the microbiological condition of a food develops according to its formulation and actual storage history. Consequently, evidence used to justify or evaluate a shelf life should be tied to the tested product and conditions. The product-specific chicken data of Söderqvist et al. [9], together with the cross-national variation in date labels and storage instructions documented for chilled ready-to-eat foods by Smigic et al. [10], illustrate why date-marking decisions should not be reduced to a single rule across unrelated foods and storage systems.

V. FOOD QUALITY and MICROBIOLOGICAL CHANGES at and BEYOND the MARKED DATE

Studies that examine foods at and beyond a printed date are particularly informative when microbiological and sensory outcomes are measured in parallel. Their interpretation must nevertheless remain tied to the food matrix, packaging system, storage temperature, sampling design, and microorganisms investigated. Evidence that sensory characteristics remain acceptable beyond a printed date is not, by itself, evidence that a safety-oriented shelf life can be extended.

Söderqvist et al. [9] investigated modified-atmosphere-packed fresh chicken breast fillets at the labelled use-by date, two days after the date, and four days after the date under storage at 4 °C and 8 °C. Changes in total aerobic count (TAC), Enterobacteriaceae, and lactic acid bacteria were

ISBN: 978-605-80193-6-2

monitored, and cooked samples were evaluated using sensory tests. Across the selected time points, TAC populations were approximately 6.5 $\log_{10}$ CFU/g at the use-by date and approximately 8.0 $\log_{10}$ CFU/g four days after the use-by date. Despite this increase in bacterial counts, storage for two or four additional days did not significantly affect the measured sensory parameters of cooked chicken compared with chicken consumed at the use-by date [9].

The microbiological results require careful interpretation. Söderqvist et al. [9] reported that total aerobic counts were close to or above levels associated with spoilage and concluded that further work was required before any extension of shelf life could be considered. Total aerobic count represents a broad bacterial population and is not equivalent to direct pathogen enumeration. The study therefore demonstrates a divergence between sensory acceptability and microbial load under the tested conditions; it does not establish pathogen absence or justify a general recommendation to consume chicken beyond its use-by date.

Evidence on foods beyond their marked dates should also be distinguished from evidence describing how dates are assigned in the marketplace. Smigic et al. [10] analyzed 2,665 chilled ready-to-eat food labels collected across six European Union and four non-EU countries and documented substantial variation in the use of "Use by" and "Best before" dates and in recommended storage temperatures. Because this was a transnational label analysis rather than a microbiological challenge or shelf-life study, it cannot establish whether any individual marked date was microbiologically conservative or permissive. Its value here is to show that real-world date marking and storage instructions are heterogeneous, reinforcing the need for product specific validation rather than universal post-date rules.

VI. STRATEGIES to IMPROVE DATE-LABEL COMMUNICATION and REDUCE FOOD WASTE

Improving date-label communication requires more than replacing one phrase with another. Consumers differ in their understanding of date terminology and in the extent to which they rely on labels when deciding whether to retain or discard food [4, 6]. Intervention effectiveness should therefore be assessed using behavioral outcomes rather than inferred from label clarity alone.

Eye-tracking evidence indicates that consumers may attend more strongly to the printed date itself than to the accompanying date phrase: in the milk-discard experiment of Badiger et al. [11], more than half of decisions involved no visual fixation on the phrase. Experimental work on Look-Smell-Taste cues likewise found that the cues did not significantly extend overdate acceptance or actual product use [12]. These findings caution against assuming that a visually clearer or simplified label will necessarily translate into less food waste.

A second strategy is to encourage consumers to use sensory assessment when dealing with foods carrying quality-oriented dates. Wallnoefer et al. [12] tested look-smell-taste (LST) cues using an online study complemented by a food-laboratory study. Although participants preferred LST-style cues over text-only alternatives, the intervention did not significantly extend acceptance of overdate products or actual product use. Sensory assessment itself was positively related to product use, suggesting that visual prompts may need to be accompanied by stronger consumer education or behavioral support rather than being treated as a stand-alone solution.

VII. INTELLIGENT PACKAGING AND DYNAMIC SHELF-LIFE INFORMATION

Conventional date labels provide static calendar information, whereas intelligent packaging can provide information about conditions experienced by a product or physicochemical changes occurring during storage. Reviews of food-packaging technologies describe time–temperature indicators (TTIs), freshness indicators, gas and pH indicators, sensors, and radio-frequency identification (RFID) systems as distinct tools for monitoring storage history, quality-related changes, or supply-chain information [13, 14]. These technologies should not be treated as interchangeable measures of microbiological safety.

TTIs integrate temperature exposure over time and can reveal cold-chain deviations that a fixed printed date cannot capture. Their response is based on time- and temperature-dependent physical, chemical, enzymatic, or biological processes. For shelf-life applications, the response kinetics of a TTI must be appropriately matched to the deterioration kinetics of the intended food. A TTI therefore provides information about time–temperature history and associated deterioration potential; it does not directly demonstrate the presence or absence of a foodborne pathogen [13].

Freshness indicators operate on a different principle. They respond to chemical or biochemical changes associated with deterioration, including carbon dioxide, volatile nitrogen compounds, sulfides, amines, and pH-related metabolites. Such responses can provide qualitative or semi-quantitative information about freshness-related changes without opening the package [14]. However, metabolite profiles and spoilage pathways differ among foods. Indicator response must therefore be calibrated and validated for the intended food matrix, packaging environment, and storage conditions; a generic color change should not be interpreted as a universal microbiological safety threshold.

Optical freshness-monitoring systems have also been investigated as tools for reducing avoidable food waste. Poyatos-Racionero et al. [15] reviewed 45 developments in optical freshness monitoring across fruits, vegetables, fish, and meat. Their review demonstrates the breadth of potential applications while also identifying research challenges that must be addressed before intelligent indicators can be used broadly as decision-support tools. The evidence supports their role as complementary information systems rather than as universal replacements for conventional shelf-life validation.

ISBN: 978-605-80193-6-2

Data-carrier technologies such as RFID and QR-based systems serve a related but technically distinct function. RFID can support product identification, traceability, inventory management, and transmission or retrieval of supply-chain information [13]. A barcode, QR code, or RFID tag by itself does not measure microbial growth, pathogen presence, or freshness. Any inference about product condition requires linkage to validated sensor data, recorded time–temperature history, or other product-specific evidence.

Intelligent packaging may therefore complement calendar-based date labels by adding condition-responsive information, but the evidentiary requirements remain product specific. TTIs require kinetic matching to the target deterioration process; freshness indicators require validation against relevant chemical, sensory, or microbiological changes; and data carriers require reliable underlying measurements. These technologies can improve the information available for shelf-life decisions, but current evidence does not support treating an unvalidated sensor signal, color response, RFID tag, or QR code as a universal indicator of food safety [13–15].

VIII. FOOD WASTE and SUSTAINABILITY IMPLICATIONS

Date marking is relevant to sustainability because it can influence whether consumers retain or discard foods that are approaching or have passed a printed date. Experimental and survey studies link date-label interpretation with willingness to consume and anticipated or reported food-waste behavior, but the response is not uniform across label wording, products, and consumers [4–6]. Date-label communication should therefore be regarded as one component of household food-waste prevention rather than as a stand-alone sustainability intervention.

Recent European evidence illustrates the complexity of this relationship. In a six-country survey of 1507 consumers, Cliceri et al. [4] reported high correct-understanding rates for the individual "Best Before" and "Use By" labels and found that 36.4% of respondents consumed recently expired foods as a strategy against food waste. Their analysis identified perceived behavioral use of foods beyond expiration dates as a predictor of lower reported household food waste. Because the study was observational, these associations should not be interpreted as proof that label knowledge alone causes reductions in household food waste.

Attention to the printed date itself may be as important as the wording that accompanies it. Using eye tracking during milk-discard decisions, Badiger et al. [11] found that participants fixated significantly more on the printed date than on the accompanying phrase, and more than half of discard decisions occurred without visual fixation on the phrase. This result cautions against assuming that terminology changes alone will substantially reduce date-related waste; the process used to determine the displayed date and the way consumers attend to it also matter.

From a sustainability perspective, the objective is to reduce avoidable disposal without treating waste reduction as evidence that a safety-oriented shelf life can be extended. Gong et al. [5] showed that date-label effects on willingness to consume overdate yogurt were mediated by consumers' quality and safety concerns, and that normal sensory information or appropriate storage information increased willingness to consume. These findings support clearer communication and informed food assessment, but they do not substitute for product-specific microbiological evidence when safety-oriented shelf lives are established or reconsidered.

Food-waste prevention also requires interventions beyond labeling. Household food waste arises from interacting psychological, socio-cultural, economic, and contextual drivers, and effective interventions may therefore need to combine information with practical food-management opportunities and capabilities [16]. Within this broader system, date labeling can contribute to sustainability when it communicates quality and safety information clearly, is supported by scientifically justified shelf-life determination, and is integrated with storage guidance and consumer food-management practices.

IX. CONCLUSION

"Best Before" and "Use By" should not be treated as interchangeable cues. Across the reviewed consumer studies, date terminology is associated with differences in understanding, perceived quality and safety, willingness to consume, and anticipated or reported waste behavior, with substantial variation across products and consumers [3–6]. Improving date-label understanding is therefore relevant to food-waste prevention, but the evidence does not support treating knowledge alone as sufficient to eliminate avoidable disposal.

Microbiological evidence also argues against interpreting a printed date as a universal biological threshold. Shelf life depends on the food matrix, initial microbiota, processing, packaging, temperature, and storage history. Product-specific evidence shows that sensory acceptability and microbial status may diverge, while cross-national label data show substantial variation in the date and storage information applied to chilled ready-to-eat foods [9, 10]. Consequently, proposals to modify safety-oriented shelf lives should be supported by product-specific microbiological evidence rather than by waste-reduction objectives alone.

Communication interventions show non-uniform effects. Consumers may focus more strongly on the printed date than on the accompanying phrase, while Look-Smell-Taste cues have not consistently increased actual use of overdate foods [11, 12]. Intelligent packaging may complement static date labels by providing information on time-temperature exposure or freshness-related physicochemical changes, but such systems require product-specific calibration and validation and should not be interpreted as generic measures of pathogen safety [13–15].

Taken together, the reviewed evidence supports an integrated approach combining scientifically justified shelf-life determination, clear differentiation of quality- and safety-related information, consumer food-management support, and product-validated condition-monitoring technologies. The evidence does not justify

ISBN: 978-605-80193-6-2

indiscriminate extension of marked dates. Rather, it supports improving the information used in consumer and shelf-life decisions while maintaining product-specific microbiological validation where safety is at issue.

References

[1] Ishangulyyev, R., Kim, S., & Lee, S. H. (2019). *Understanding food loss and waste Why are we losing and wasting food? Foods*, *8*(8), 297. https://doi.org/10.3390/foods8080297

[2] Tutar, H., Streimikiene, D., Tahiri Mutlu, H., Kloudova, J., & Bilan, Y. (2025). *Global food waste as an anti-sustainability trend:* Analysis of economic and environmental impacts across countries. *British Food Journal*, *127*(13), 674–692. https://doi.org/10.1108/BFJ-05-2025-0650

[3] Zielińska, D., Bilska, B., Marciniak-Łukasiak, K., Łepecka, A., Trząskowska, M., Neffe-Skocińska, K., et al. (2020). *Consumer understanding of the date of minimum durability of food in association with quality evaluation of food products after expiration*. *International Journal of Environmental Research and Public Health*, *17*(5), 1632. https://doi.org/10.3390/ijerph17051632

[4] Cliceri, D., Pedrotti, M., Gasperi, F., & Endrizzi, I. (2025). *European consumers' involvement with date labels and implications for household food waste. Food Quality and Preference*, *129*, 105522. https://doi.org/10.1016/j.foodqual.2025.105522

[5] Gong, Z., Su, L. Y.-F., Zhang, J. S., Chen, T., & Wang, Y.-C. (2022). *Understanding the association between date labels and consumer-level food waste*. *Food Quality and Preference*, *96*, 104373. https://doi.org/10.1016/j.foodqual.2021.104373

[6] Wilson, N. L. W., & Miao, R. (2025). *Food waste, date labels, and risk preferences:* An experimental exploration. *Applied Economic Perspectives and Policy*, *47*(3), 1029–1057. https://doi.org/10.1002/aepp.13507

[7] Sielicka-Różyńska, M., & Samotyja, U. (2022). *Influence of "best before" dates on expected and actual food liking*. *British Food Journal*, *125*(4), 1317–1334. https://doi.org/10.1108/BFJ-08-2021-0935

[8] Karanth, S., Feng, S., Patra, D., & Pradhan, A. K. (2023). *Linking microbial contamination to food spoilage and food waste:* The role of smart packaging, spoilage risk assessments, and date labeling. *Frontiers in Microbiology*, *14*, 1198124. https://doi.org/10.3389/fmicb.2023.1198124

[9] Söderqvist, K., Peterson, M., Johansson, M., Olsson, V., & Boqvist, S. (2024). *A microbiological and sensory evaluation of modified atmosphere-packed (MAP) chicken at use-by date and beyond*. *Foods*, *13*(13), 2140. https://doi.org/10.3390/foods13132140

[10] Smigic, N., Ozilgen, S., Gómez-López, V. M., Osés, S. M., Miloradovic, Z., Aleksic, B., et al. (2026). *Food labelling in focus*: A transnational analysis of chilled ready-to-eat products in the EU, the Western Balkans and Turkey. *International Journal of Food Sciences and Nutrition*, *77*(3), 225–237. https://doi.org/10.1080/09637486.2026.2655369

[11] Badiger, A., Katz, T., Simons, C. T., & Roe, B. E. (2023). *When considering whether to waste food, consumers focus attention on food label dates rather than phrases*. *Waste Management*, *168*, 230–234. https://doi.org/10.1016/j.wasman.2023.06.006

[12] Wallnoefer, L. M., Meixner, O., & Riefler, P. (2024). *Look-smell-taste labels on food date marking*: Assessing their effectiveness for reducing food waste at a consumer level as part of the European Green Deal. *Food Quality and Preference*, *120*, 105253. https://doi.org/10.1016/j.foodqual.2024.105253

[13] Kalpana, S., Priyadarshini, S. R., Maria Leena, M., Moses, J. A., & Anandharamakrishnan, C. (2019). *Intelligent packaging*: Trends and applications in food systems. *Trends in Food Science & Technology*, *93*, 145–157. https://doi.org/10.1016/j.tifs.2019.09.008

[14] Shao, P., Liu, L., Yu, J., Lin, Y., Gao, H., Chen, H., et al. (2021). *An overview of intelligent freshness indicator packaging for food quality and safety monitoring. Trends in Food Science & Technology*, *118*, 285–296. https://doi.org/10.1016/j.tifs.2021.10.012

[15] Poyatos-Racionero, E., Ros-Lis, J. V., Vivancos, J.-L., & Martínez-Máñez, R. (2018). *Recent advances on intelligent packaging as tools to reduce food waste. Journal of Cleaner Production, 172*, 3398–3409. https://doi.org/10.1016/j.jclepro.2017.11.075

[16] Vittuari, M., Garcia Herrero, L., Masotti, M., Iori, E., Caldeira, C., Qian, Z., et al. (2023). *How to reduce consumer food waste at household level*: A literature review on drivers and levers for behavioural change. *Sustainable Production and Consumption*, *38*, 104–114. https://doi.org/10.1016/j.spc.2023.03.023

ISBN: 978-605-80193-6-2

Nanotechnology Applications in Foods: Nanoencapsulation of Phenolic Compounds

Serdar YEŞİL[1*]

[1]Karabük University, Eflani Vocational School, Food Processing Department, Karabük, Türkiye

* Corresponding author e-mail: serdaryesil@karabuk.edu.tr

***Abstract*—Nanotechnology is an interdisciplinary field that offers new opportunities in food processing, preservation, packaging, quality and safety control, and the delivery of bioactive ingredients. Phenolic compounds found in plant sources such as fruits, vegetables, cereals, tea, coffee and cocoa are among the most important constituents of functional foods because of their antioxidant, anti-inflammatory and antimicrobial effects. However, their effective use in food systems is limited by their low water solubility, their bitter and astringent taste, their sensitivity to heat, light, oxygen and pH changes, their rapid degradation in the gastrointestinal tract and the resulting low bioavailability. Nanoencapsulation is an approach that entraps phenolic compounds within nanoscale carriers, thereby improving their stability, masking their undesirable sensory attributes, enabling controlled release and enhancing their bioavailability. In this study, the main applications of nanotechnology in foods are summarized. The systems used for the nanoencapsulation of phenolic compounds are then comparatively evaluated, namely lipid-based carriers (nanoemulsions, liposomes, solid lipid nanoparticles), biopolymer-based nanoparticles, equipment-based techniques (electrospinning, electrospraying, nano spray drying) and natural nanocarriers (cyclodextrins, casein micelles). Release mechanisms from nanocarriers, the characterization of nanoencapsulated phenolics and their use in food fortification and active packaging are also discussed. Selected studies from the literature show that nanoencapsulation can markedly improve the stability and bioavailability of phenolics such as catechins, curcumin, resveratrol and epigallocatechin gallate. Nevertheless, the toxicological safety of nanomaterials, scale-up costs, regulation and consumer acceptance stand out as priority research needs in the field.**

***Keywords*—Nanotechnology, nanoencapsulation, phenolic compounds, bioavailability, functional food**

I. INTRODUCTION

Nanotechnology deals with the design, production and use of structures that have at least one dimension in the range of approximately 1–100 nm, and it gives materials a high surface area-to-volume ratio together with physical, chemical and biological properties that differ from those observed at the macroscale [1, 2]. However, there is no agreed definition of the term “nano” for foods, and structures of up to several hundred nanometers are also frequently referred to as nanoparticles in the scientific literature [3]. In food science, nanotechnology has found applications in many areas, including processing, preservation, packaging, the detection of contaminants and pathogens, and the targeted delivery of nutrients and bioactive compounds [4, 5]. Recent comprehensive reviews highlight the potential of nanotechnology to extend the shelf life of foods, improve their nutritional and sensory quality and enhance the bioavailability of functional ingredients, while noting that safety and regulatory aspects also need to be addressed [6, 7, 8].

Phenolic compounds are a large group of secondary metabolites synthesized by plants, with more than 8000 different structures identified [9]. Phenolic acids, flavonoids (flavonols, flavanols, anthocyanins, etc.), stilbenes and lignans are the main classes of this group [9]. Chlorogenic acid in coffee, catechins in tea and cocoa, quercetin in onion and apple, anthocyanins in red fruits, resveratrol in grapes and curcumin in turmeric are widely studied examples of phenolic compounds in foods [9, 10, 11]. The antioxidant, anti-inflammatory, anticarcinogenic and antimicrobial effects of these compounds have been associated with the prevention of chronic diseases such as cardiovascular disease, neurodegenerative disorders and diabetes [9, 11]. However, the physicochemical instability and low bioavailability of phenolics limit the transfer of these health effects to food products. For example, the oral bioavailability of tea catechins has been reported to be only 2–5% [12].

Encapsulation is the process of protecting a sensitive ingredient (the core material) from environmental factors by entrapping it within a wall material or matrix [13]. Nanoencapsulation is the application of this process at the nanoscale, and compared with microencapsulation it offers advantages such as a larger surface area, better solubility and dispersibility, and controlled release [14, 15, 16]. The aim of this study is to evaluate, in light of the current literature, the applications of nanotechnology in foods, the carrier systems used for the nanoencapsulation of phenolic compounds, the effects of these systems on stability and bioavailability, their food applications, and the related safety and regulatory aspects.

II. MAIN APPLICATIONS OF NANOTECHNOLOGY IN FOODS

Applications of food nanotechnology can generally be grouped under four main headings [17]:

1) **Food processing and formulation:** Incorporation of bioactive compounds, vitamins, flavors and colorants into foods by means of nanoemulsions, nanocapsules and nanofibers, and improvement of texture, flavor and color stability.

2) **Packaging:** Enhancement of gas and moisture barrier properties with polymer nanocomposites containing clay, nanocellulose or metal/metal oxide nanoparticles, antimicrobial active

ISBN: 978-605-80193-6-2

packaging based on silver, zinc oxide and chitosan, and intelligent packaging that monitors freshness and temperature history. For example, the antimicrobial efficiency of silver–polyamide nanocomposites has been reported to last for 28 days, while silver migration from silver/poly(vinyl chloride) films into chicken meat was found to remain below the legal migration limits set by the European Union (EU) for plastic films [2].

3) **Food safety and analysis:** Sensitive nanosensors for the detection of pathogens, pesticides and toxins, and nanobarcodes for product tracking and traceability [1, 8].

4) **Delivery of bioactive compounds:** Nanoencapsulation of sensitive ingredients such as phenolic compounds, carotenoids, omega-3 fatty acids and vitamins [17].

III. FACTORS LIMITING THE USE OF PHENOLIC COMPOUNDS IN FOODS

The main factors that limit the efficacy of phenolic compounds in food systems and in the human body are as follows [13, 18, 19]:

- The very low water solubility of many phenolics, such as curcumin, quercetin and resveratrol.
- Their rapid degradation through oxidation and epimerization when exposed to high temperature, light, oxygen, metal ions and neutral to alkaline pH.
- Their breakdown in the gastrointestinal tract by pH changes, digestive enzymes and the microbiota, their limited permeation across the intestinal epithelium and their extensive phase II metabolism.
- Their adverse effect on the sensory acceptance of products, as they impart a bitter and astringent taste and cause color changes.

Although the daily dietary intake of polyphenols can reach approximately 1 g [10], the low concentrations attained in plasma are the clearest indication of the bioavailability problem. Nanoencapsulation offers a strategy that addresses each of these limitations: the carrier matrix protects the phenolic compound from adverse environmental conditions, the nanoscale size increases the surface-to-volume ratio and cellular uptake, and the wall material can mask unpleasant tastes and enable controlled or targeted release [11, 13, 15].

IV. SYSTEMS USED FOR THE NANOENCAPSULATION OF PHENOLIC COMPOUNDS

Nanocarriers developed for phenolics can be classified into four groups according to their production principles and carrier materials: lipid-based systems, biopolymer-based nanostructures, nanocarriers produced by equipment-based techniques and natural nanocarriers [18]. The comparative characteristics of the main systems are given in Table I.

A. Lipid-Based Nanocarriers

Nanoemulsions are kinetically stable systems that are prepared by high- or low-energy emulsification methods and generally have droplet sizes below 200 nm. Owing to their small droplet size, they can be used in optically transparent products (e.g., beverages), and they improve the bioavailability of encapsulated lipophilic bioactive compounds [20]. Liposomes are vesicles composed of phospholipid bilayers, and they have the advantage of carrying hydrophilic phenolics in their inner aqueous phase and lipophilic phenolics within the lipid bilayer at the same time. Owing to their lipid matrices, which are solid at room temperature, solid lipid nanoparticles (SLNs) and nanostructured lipid carriers (NLCs) protect labile compounds against degradation and provide controlled release. In NLCs, the less ordered lipid matrix created by the blend of solid and liquid lipids increases the loading capacity [3, 19].

B. Biopolymer-Based Nanostructures

Proteins (whey proteins, β-lactoglobulin, casein, zein, gelatin, soy protein) and polysaccharides (chitosan, alginate, pectin, starch, gum arabic) are widely used for the nanoencapsulation of phenolics because they are food-grade, biocompatible and biodegradable materials, most of which are generally recognized as safe (GRAS) [21]. Biopolymer nanoparticles can be produced by relatively mild methods such as ionic gelation, nanoprecipitation and complex coacervation. Examples of the protection of catechin and epigallocatechin gallate (EGCG) with chitosan- and β-lactoglobulin-based carriers are given in Table II. Poly(lactic-co-glycolic acid) (PLGA), a synthetic but biodegradable polymer, is frequently used in bioavailability studies of phenolics because it is approved for pharmaceutical use [11].

C. Equipment-Based Techniques and Natural Nanocarriers

Electrospinning and electrospraying produce nanofibers or nanoparticles under high voltage and, because they do not require heat, are suitable for heat-sensitive phenolics [22]. Nano spray drying, which was developed to extend spray drying to the submicron level, yields much smaller particles than conventional spray drying [21, 23]. Supercritical fluid techniques are generally applied with supercritical CO_2. Because CO_2 has a low critical temperature and the process requires very little organic solvent, these techniques are an environmentally friendly option for the nanoencapsulation of heat-sensitive bioactives [3, 23]. Among natural nanocarriers, cyclodextrins form inclusion complexes with phenolics through their hydrophobic inner cavity and thus increase their solubility and stability, whereas casein micelles are nanoscale structures that occur naturally in dairy products and can bind hydrophobic bioactives [18, 21].

ISBN: 978-605-80193-6-2

TABLE I. Comparison of the main carrier systems used for the nanoencapsulation of phenolic compounds [3, 18, 19, 21, 24].

Carrier system	Typical materials (production method)	Advantages	Limitations
Nanoemulsion	Vegetable oil, lecithin, Tween, proteins (high-pressure homogenization, ultrasonication)	Transparent products, high bioaccessibility, easy scale-up	Need for synthetic emulsifiers, Ostwald ripening
Liposome	Phospholipids, cholesterol (thin-film hydration, ethanol injection)	Co-delivery of hydrophilic and lipophilic compounds	Phospholipid oxidation, leakage, cost
SLN / NLC	Solid fats, waxes, liquid oil (hot/cold homogenization)	Protection against oxidation, sustained release	Polymorphic transition and expulsion of the loaded compound during storage
Biopolymer nanoparticle	Chitosan, zein, casein, β-lactoglobulin, alginate (ionic gelation, nanoprecipitation, complex coacervation)	Food-grade materials, pH-responsive controlled release	Particle size distribution may be broad
Nanofiber / nanoparticle (electrohydrodynamic)	Zein, gelatin, poly(vinyl alcohol), pullulan (electrospinning, electrospraying)	No heat required, suitable for active packaging	Low production rate
Natural nanocarrier	Cyclodextrins, casein micelles (inclusion complexation, spontaneous binding)	Simple preparation, increased solubility	Limited loading capacity

V. RELEASE MECHANISMS FROM NANOCARRIERS

The success of nanoencapsulation depends not only on protecting the phenolic compound but also on releasing it at the right time and at the right site. In controlled release systems, the aim is to release the bioactive compound with a predetermined profile over a predetermined period [23]. Diffusion is the dominant mechanism in this process. Phenolic molecules migrate from the interior of the carrier to the surrounding medium along a concentration gradient, and the diffusion rate depends on the solubility of the compound and on the permeability and structure of the carrier matrix [23]. In erosion-controlled release, the compound is released as the carrier dissolves or degrades. Degradation confined to the outer layer is termed surface erosion, whereas degradation throughout the structure is termed bulk erosion, and the pores formed by erosion can accelerate diffusion [23]. In swelling- and osmosis-based mechanisms, release is triggered by expansion of the carrier as it absorbs solvent or by its rupture under osmotic pressure. This behavior can be tuned by selecting an appropriate polymer matrix and through external conditions such as temperature and pH [23]. This is particularly important for phenolics. For example, chitosan–tripolyphosphate nanoparticles have enabled the controlled release of tea catechins [12], and casein nanoparticles have been shown to act as gastro-resistant carriers that release their load only under simulated intestinal conditions [23].

VI. CHARACTERIZATION OF NANOENCAPSULATED PHENOLICS

Whether a nanoencapsulation process actually improves the properties of the phenolic compound should be confirmed by comparative tests with the nonencapsulated compound [3]. The parameter most commonly used in characterization is encapsulation efficiency (EE), which is calculated using (1):

$$EE\ (\%) = (W_E / W_T) \times 100 \quad (1)$$

Here, W_E is the amount of phenolic compound retained in the carrier and W_T is the total amount of phenolic compound added to the formulation. Loading capacity is the ratio of the amount of retained phenolic compound to the total particle mass. Particle size and size distribution are determined by dynamic light scattering, surface charge by zeta potential measurement, and morphology by scanning and transmission electron microscopy (SEM/TEM). Fourier transform infrared spectroscopy (FTIR), X-ray diffraction (XRD) and differential scanning calorimetry (DSC) reveal phenolic–wall material interactions, changes in crystallinity and thermal behavior, respectively [23]. Finally, in vitro digestion models based on simulated gastric and intestinal fluids are used to assess the release profile, bioaccessibility and retention of antioxidant activity. These results need to be supported by cell culture and in vivo studies [21, 25].

VII. STABILITY, BIOAVAILABILITY AND FOOD APPLICATIONS

The most striking effects of nanoencapsulation on the biological efficacy of phenolics are seen in bioavailability studies (Table II). For example, the oral bioavailability of curcumin loaded into PLGA nanoparticles was at least 9-fold higher than that of curcumin administered with piperine [26]. These increases are attributed to enhanced solubility, protection of the phenolic compound in the gastrointestinal environment, and increased absorption and cellular uptake [11, 19].

TABLE II. Selected studies on the nanoencapsulation of phenolic compounds and their findings.

Phenolic / carrier	Main finding	Reference
Tea catechins / chitosan–tripolyphosphate nanoparticles	Encapsulation efficiency of 24–53%, control of size and charge by fabrication conditions, controlled release	Hu et al. [12]
Catechins / chitosan nanoparticles	Improved stability against degradation in alkaline solution	Dube et al. [27]
EGCG / β-lactoglobulin nanocomplexes	Protection against oxidative degradation, suitability for clear beverages	Shpigelman et al. [28]
Resveratrol / nanoemulsion	Retention of chemical stability and antioxidant activity during in vitro digestion	Sessa et al. [25]
Curcumin / PLGA nanoparticles	At least 9-fold higher oral bioavailability than curcumin with piperine	Shaikh et al. [26]
EGCG / PLGA nanoparticles	About 30-fold dose advantage over free EGCG in preventing DNA damage	Rahim et al. [11]
Curcumin / nanoliposomes (supercritical CO_2-assisted)	128–228 nm, encapsulation efficiency of up to 95% for curcumin, antioxidant activity retained	Singh et al. [23]

ISBN: 978-605-80193-6-2

Phenolic / carrier	Main finding	Reference
Catechin / starch nanoparticles (ultrasonication)	322.7–615.6 nm, encapsulation efficiency of 48.3–59.1%, better retention of bioactivity during in vitro digestion	Singh et al. [23]

In terms of food applications, nanoencapsulated phenolic compounds and extracts have been applied to meat products and vegetable oils, and coconut milk has been enriched with nanoencapsulated catechin [16]. Encapsulation can also alleviate the unpleasant taste that limits the use of phenolics [13]. Another important application is the incorporation of plant-derived antioxidant and antimicrobial agents, such as essential oils, plant extracts and epicatechin gallate nanocapsules, into edible nanocoatings and active packaging materials, for example in the form of nanoemulsions or electrospun nanofibers, to extend shelf life [2, 17]. Nevertheless, further studies are needed to improve existing encapsulation systems and to explore their application in food and gastrointestinal systems before laboratory-scale successes can be translated into industrial products [24].

VIII. CHALLENGES IN COMMERCIALIZATION AND SCALE-UP

Despite the favorable results obtained at the laboratory scale, the conversion of nanoencapsulated phenolics into commercial products remains limited. One of the reasons is that there is no universal method applicable to all compounds, because each core material and intended use requires a specific design [23]. Ease of scale-up, energy and raw material costs, and product yield are decisive in the choice of method. For example, although spray drying is a rapid, economical and continuous process, its high temperature can cause losses of heat-sensitive phenolics such as anthocyanins, and adhesion to the dryer wall can reduce the yield. Freeze drying is suitable for heat-sensitive compounds but requires long processing times and high energy input. Nanoemulsions offer easy production, high loading capacity and optical transparency, but control over release is limited [23]. For PLGA nanoparticles, high production costs and difficulties in scale-up also restrict their clinical use [11]. A general comparison of micro- and nanoencapsulation is summarized in Table III.

TABLE III. Comparison of micro- and nanoencapsulation with respect to phenolic compounds [3, 11, 16, 23, 29, 30].

Property	Microencapsulation	Nanoencapsulation
Particle size	Generally 1–1000 μm	Generally below 1000 nm (1–100 nm under the strict definition)
Surface area / volume	Low	High, with higher solubility and dissolution rate
Effect on bioavailability	Limited	Marked increase through enhanced cellular uptake
Product appearance	Turbidity and sedimentation may occur	Can be incorporated into clear beverages
Production and scale-up	Established industrial methods (e.g., spray drying)	Most methods at laboratory scale, high cost
Safety and regulation	Conventional food ingredient assessment	Nanomaterial-specific risk assessment and "(nano)" labeling

The situation is similar with regard to regulation. Gutiérrez et al. [3] stated that, despite the available technologies, there were no approved nanofoods in the United Kingdom and the EU and that nanoparticles in foods or food supplements were regarded as novel foods. However, the existence of commercial examples, such as a nanoparticulate lycopene formulation with GRAS status [3], shows that products developed with food-grade carriers and with proven safety can reach the market.

IX. SAFETY, REGULATION AND CONSUMER ACCEPTANCE

The high reactivity and cellular uptake associated with nanoscale structures also bring potential toxicological risks. The accumulation of nanoparticles in the digestive tract, their passage across the intestinal barrier into the systemic circulation and their long-term effects have not yet been sufficiently elucidated [4]. It has been noted that, because nanotoxicology is a relatively new discipline, standardized and validated toxicity tests for nanoparticles are not yet available [3]. For this reason, the recommended assessment steps are detailed characterization of the nanostructure to be tested and of its state after each processing step, determination of its penetration into the biological system and evaluation of the selected toxicity endpoint [3]. In addition to in vitro genotoxicity tests, in vivo models such as zebrafish embryos are also used [3]. In the EU, foods consisting of engineered nanomaterials are subject to pre-market authorization as novel foods [31], and such ingredients must be indicated with the word "nano" in the list of ingredients [29]. The European Food Safety Authority (EFSA) has published comprehensive guidance on the risk assessment of nanomaterials in the food and feed chain, which sets out requirements for particle characterization, dissolution and degradation behavior during digestion, and toxicological testing [30]. Consumer perception of "nano" is also an important factor affecting commercialization, and regulatory frameworks that address consumer concerns are needed to increase acceptance [3, 23].

X. CONCLUSION AND FUTURE PERSPECTIVES

Nanoencapsulation stands out as an effective strategy for overcoming the main limitations of phenolic compounds, such as low solubility, chemical instability, unfavorable sensory properties and low bioavailability. This approach also lays the groundwork for the development of "smart foods" that deliver bioactive compounds to the target site in a controlled manner [23]. Lipid-based systems and nanocarriers made from food-grade biopolymers appear to be the most suitable options for food applications in terms of both safety and regulatory compliance. Examples from the literature show that, with a suitable carrier, the stability and bioavailability of phenolics can be increased manyfold. Future research should focus on (i) developing low-cost, solvent-free "clean label" production methods applicable in real food matrices and at industrial scale, (ii) confirming bioavailability and safety through in vivo and clinical studies in addition to standardized in vitro

ISBN: 978-605-80193-6-2

digestion models, (iii) improving sustainability by valorizing phenolics and wall materials obtained from agri-food by-products, and (iv) strengthening consumer awareness of and trust in nanotechnology.

References

[1] Chellaram, C., Murugaboopathi, G., John, A. A., Sivakumar, R., Ganesan, S., Krithika, S., & Priya, G. (2014). Significance of nanotechnology in food industry. *APCBEE Procedia*, *8*, 109–113. https://doi.org/10.1016/j.apcbee.2014.03.010

[2] Nile, S. H., Baskar, V., Selvaraj, D., Nile, A., Xiao, J., & Kai, G. (2020). Nanotechnologies in food science: Applications, recent trends, and future perspectives. *Nano-Micro Letters*, *12*(1), 45. https://doi.org/10.1007/s40820-020-0383-9

[3] Gutiérrez, F. J., Albillos, S. M., Casas-Sanz, E., Cruz, Z., García-Estrada, C., García-Guerra, A., García-Reverter, J., García-Suárez, M., Gatón, P., González-Ferrero, C., Olabarrieta, I., Olasagasti, M., Rainieri, S., Rivera-Patiño, D., Rojo, R., Romo-Hualde, A., Sáiz-Abajo, M.-J., & Mussons, M.-L. (2013). Methods for the nanoencapsulation of β-carotene in the food sector. *Trends in Food Science & Technology*, *32*(2), 73–83. https://doi.org/10.1016/j.tifs.2013.05.007

[4] He, X., & Hwang, H.-M. (2016). Nanotechnology in food science: Functionality, applicability, and safety assessment. *Journal of Food and Drug Analysis*, *24*(4), 671–681. https://doi.org/10.1016/j.jfda.2016.06.001

[5] Singh, T., Shukla, S., Kumar, P., Wahla, V., Bajpai, V. K., & Rather, I. A. (2017). Application of nanotechnology in food science: Perception and overview. *Frontiers in Microbiology*, *8*, 1501. https://doi.org/10.3389/fmicb.2017.01501

[6] Sahoo, M., Vishwakarma, S., Panigrahi, C., & Kumar, J. (2021). Nanotechnology: Current applications and future scope in food. *Food Frontiers*, *2*(1), 3–22. https://doi.org/10.1002/fft2.58

[7] Mohammad, Z. H., Ahmad, F., Ibrahim, S. A., & Zaidi, S. (2022). Application of nanotechnology in different aspects of the food industry. *Discover Food*, *2*(1), 12. https://doi.org/10.1007/s44187-022-00013-9

[8] Singh, R., Dutt, S., Sharma, P., Sundramoorthy, A. K., Dubey, A., Singh, A., & Arya, S. (2023). Future of nanotechnology in food industry: Challenges in processing, packaging, and food safety. *Global Challenges*, *7*(4), 2200209. https://doi.org/10.1002/gch2.202200209

[9] Pandey, K. B., & Rizvi, S. I. (2009). Plant polyphenols as dietary antioxidants in human health and disease. *Oxidative Medicine and Cellular Longevity*, *2*(5), 270–278. https://doi.org/10.4161/oxim.2.5.9498

[10] Manach, C., Scalbert, A., Morand, C., Rémésy, C., & Jiménez, L. (2004). Polyphenols: Food sources and bioavailability. *The American Journal of Clinical Nutrition*, *79*(5), 727–747. https://doi.org/10.1093/ajcn/79.5.727

[11] Rahim, R. A., Jayusman, P. A., Muhammad, N., Ahmad, F., Mokhtar, N., Mohamed, I. N., Mohamed, N., & Shuid, A. N. (2019). Recent advances in nanoencapsulation systems using PLGA of bioactive phenolics for protection against chronic diseases. *International Journal of Environmental Research and Public Health*, *16*(24), 4962. https://doi.org/10.3390/ijerph16244962

[12] Hu, B., Pan, C., Sun, Y., Hou, Z., Ye, H., Hu, B., & Zeng, X. (2008). Optimization of fabrication parameters to produce chitosan–tripolyphosphate nanoparticles for delivery of tea catechins. *Journal of Agricultural and Food Chemistry*, *56*(16), 7451–7458. https://doi.org/10.1021/jf801111c

[13] Fang, Z., & Bhandari, B. (2010). Encapsulation of polyphenols – a review. *Trends in Food Science & Technology*, *21*(10), 510–523. https://doi.org/10.1016/j.tifs.2010.08.003

[14] Ezhilarasi, P. N., Karthik, P., Chhanwal, N., & Anandharamakrishnan, C. (2013). Nanoencapsulation techniques for food bioactive components: A review. *Food and Bioprocess Technology*, *6*(3), 628–647. https://doi.org/10.1007/s11947-012-0944-0

[15] McClements, D. J. (2015). Nanoscale nutrient delivery systems for food applications: Improving bioactive dispersibility, stability, and bioavailability. *Journal of Food Science*, *80*(7), N1602–N1611. https://doi.org/10.1111/1750-3841.12919

[16] Gündoğan, R., Tomar, G. Ş., & Can Karaça, A. (2023). Gıda bileşenlerinin nanoenkapsülasyonuna yönelik güncel çalışmalar. *ITU Journal of Food Science and Technology*, *1*(1), 1–14. https://dergipark.org.tr/en/pub/itujfst/article/1258208

[17] Biswas, R., Alam, M., Sarkar, A., Haque, M. I., Hasan, M. M., & Hoque, M. (2022). Application of nanotechnology in food: Processing, preservation, packaging and safety assessment. *Heliyon*, *8*(11), e11795. https://doi.org/10.1016/j.heliyon.2022.e11795

[18] Assadpour, E., Jafari, S. M., & Faridi Esfanjani, A. (2017). Protection of phenolic compounds within nanocarriers. *CAB Reviews*, *12*(057), 1–8. https://doi.org/10.1079/PAVSNNR201712057

[19] Faridi Esfanjani, A., Assadpour, E., & Jafari, S. M. (2018). Improving the bioavailability of phenolic compounds by loading them within lipid-based nanocarriers. *Trends in Food Science & Technology*, *76*, 56–66. https://doi.org/10.1016/j.tifs.2018.04.002

[20] Thiruvengadam, M., Rajakumar, G., & Chung, I. M. (2018). Nanotechnology: Current uses and future applications in the food industry. *3 Biotech*, *8*, 74. https://doi.org/10.1007/s13205-018-1104-7

[21] Faridi Esfanjani, A., & Jafari, S. M. (2016). Biopolymer nano-particles and natural nano-carriers for nano-encapsulation of phenolic compounds. *Colloids and Surfaces B: Biointerfaces*, *146*, 532–543. https://doi.org/10.1016/j.colsurfb.2016.06.053

ISBN: 978-605-80193-6-2

[22] Anandharamakrishnan, C., & Parthasarathi, S. (Eds.). (2019). *Food nanotechnology: Principles and applications*. CRC Press. https://doi.org/10.1201/9781315153872

[23] Singh, A. K., Pal, P., Pandey, B., Goksen, G., Sahoo, U. K., Lorenzo, J. M., & Sarangi, P. K. (2023). Development of "smart foods" for health by nanoencapsulation: Novel technologies and challenges. *Food Chemistry: X*, *20*, 100910. https://doi.org/10.1016/j.fochx.2023.100910

[24] Shishir, M. R. I., Xie, L., Sun, C., Zheng, X., & Chen, W. (2018). Advances in micro and nano-encapsulation of bioactive compounds using biopolymer and lipid-based transporters. *Trends in Food Science & Technology*, *78*, 34–60. https://doi.org/10.1016/j.tifs.2018.05.018

[25] Sessa, M., Tsao, R., Liu, R., Ferrari, G., & Donsì, F. (2011). Evaluation of the stability and antioxidant activity of nanoencapsulated resveratrol during in vitro digestion. *Journal of Agricultural and Food Chemistry*, *59*(23), 12352–12360. https://doi.org/10.1021/jf2031346

[26] Shaikh, J., Ankola, D. D., Beniwal, V., Singh, D., & Kumar, M. N. V. R. (2009). Nanoparticle encapsulation improves oral bioavailability of curcumin by at least 9-fold when compared to curcumin administered with piperine as absorption enhancer. *European Journal of Pharmaceutical Sciences*, *37*(3–4), 223–230. https://doi.org/10.1016/j.ejps.2009.02.019

[27] Dube, A., Ng, K., Nicolazzo, J. A., & Larson, I. (2010). Effective use of reducing agents and nanoparticle encapsulation in stabilizing catechins in alkaline solution. *Food Chemistry*, *122*(3), 662–667. https://doi.org/10.1016/j.foodchem.2010.03.027

[28] Shpigelman, A., Israeli, G., & Livney, Y. D. (2010). Thermally-induced protein–polyphenol co-assemblies: Beta lactoglobulin-based nanocomplexes as protective nanovehicles for EGCG. *Food Hydrocolloids*, *24*(8), 735–743. https://doi.org/10.1016/j.foodhyd.2010.03.015

[29] European Parliament and Council. (2011). Regulation (EU) No 1169/2011 on the provision of food information to consumers. *Official Journal of the European Union*, L 304, 18–63.

[30] EFSA Scientific Committee. (2021). Guidance on risk assessment of nanomaterials to be applied in the food and feed chain: Human and animal health. *EFSA Journal*, *19*(8), 6768. https://doi.org/10.2903/j.efsa.2021.6768

[31] European Parliament and Council. (2015). Regulation (EU) 2015/2283 on novel foods. *Official Journal of the European Union*, L 327, 1–22.

 ISBN: 978-605-80193-6-2

Food Non-Compliances in Türkiye and Their Distribution by Public Health Risk: An Analysis of 559 Official Public Disclosure Records

Serdar YEŞİL[1*]

[1]Karabük University, Eflani Vocational School, Food Processing Department, Karabük, Türkiye

* Corresponding author e-mail: serdaryesil@karabuk.edu.tr

***Abstract*—Food counterfeiting and adulteration mislead consumers and, depending on the substance or practice involved, can also pose a risk to public health. This study evaluated the 559 product records in the 27th public announcement published by the Republic of Türkiye Ministry of Agriculture and Forestry on 1 March 2022 by food group, type of non-compliance and public health relevance. The non-compliances were grouped into 11 main categories, and the records were further assigned to three classes: adulteration and counterfeiting, potential chemical risk, and direct health-risk indicator. Milk and dairy products (27.9%), meat and meat products (23.8%) and vegetable oils (20.4%) together accounted for 72.1% of all records. The most frequent non-compliances were fat substitution (35.2%), meat species or offal substitution (23.4%) and pharmaceutical active substances (14.7%). Of the records, 396 (70.8%) were classified as adulteration and counterfeiting, 66 (11.8%) as potential chemical risk and 97 (17.4%) as direct health-risk indicators. Records in the direct health-risk indicator or potential chemical risk classes made up 29.2% of the total. The distribution of the three classes differed significantly between food groups ($p<0.001$). The findings show that evaluating official food control data by product group and nature of the non-compliance together, rather than by notification counts alone, yields more informative results.**

Keywords—Food adulteration, food fraud, public health, official control, Türkiye

I. INTRODUCTION

Food fraud covers practices such as the intentional substitution, dilution, misdeclaration or concealment of the nature of a food or food ingredient for economic gain [1], [2]. Although these practices are primarily economically motivated, they can also have serious public health consequences when undeclared allergens, pharmaceutical active substances, non-compliant additives or unexpected animal species enter the food chain [1]–[3]. The assessment of food fraud is therefore important not only in terms of economic losses and consumer deception, but also for food safety and the protection of public health.

The forms of counterfeiting and adulteration differ markedly between food groups. In meat and meat products, species substitution and the addition of other tissues are prominent [4], whereas in honey, adulteration with sugar syrups and verification of the product's true composition are among the main issues [5]. Blending of olive oil with lower-value oils is monitored through multiple analytical indicators, particularly the fatty acid and sterol profiles [6]. The presence of undeclared pharmaceutical active substances such as sildenafil, tadalafil and sibutramine in food supplements is directly related to consumer safety [7]–[9]. Similarly, colorants that are not permitted or are used under inappropriate conditions constitute a separate monitoring area for food safety [10].

In Türkiye, official food controls are carried out under Law No. 5996, and products in which laboratory analyses reveal counterfeiting, adulteration or a non-compliance that may endanger health can be disclosed to the public. The 27th public announcement, published by the Ministry of Agriculture and Forestry on 1 March 2022, listed 559 product records belonging to 371 companies [11]. Public announcements for the 2016–2020 period have been evaluated previously, and vegetable oils, meat products and dairy products were reported to be the most prominent groups [12]. However, examining total record counts alone does not adequately explain the nature of the non-compliances in different food groups or their priority for public health.

This study aimed to determine the distribution of the 559 product records in the 2022 public announcement by food group, to classify the reported non-compliances under common headings, to classify the records from a public health perspective, and to describe the distribution of these classes across food groups.

II. MATERIALS AND METHODS

A. Data Source and Record Preparation

A retrospective analysis of official records was performed. The primary data source was the 54-page official public announcement list published by the Ministry of Agriculture and Forestry on 1 March 2022 [11]. Each product lot in the list was treated as one record and unit of analysis. Food group, product name, company and brand, reported non-compliance, and lot or serial number were recorded. The dataset comprised 13 food groups: alcoholic beverages, non-alcoholic beverages, bee products, spices, herbal, tea and coffee products, vegetable oils, chocolate and cocoa products, energy drinks, meat and meat products, bakery products and desserts, milk and dairy products, confectionery, and food supplements.

B. Classification of Non-Compliances and Public Health Assessment

Before statistical evaluation, the non-compliance statements in the Ministry records were grouped under common headings. Eleven main non-compliance groups were defined: fat substitution, meat species or offal substitution, pharmaceutical active substances, preservative non-compliance, colorant non-compliance, honey composition non-compliance, undeclared fillers

ISBN: 978-605-80193-6-2

or substitutes, milk species substitution, tert-butyl alcohol, alcohol composition non-compliance, and other compositional non-compliances. The non-compliance groups are not mutually exclusive; records with more than one finding were counted in each relevant group. In the public health classification, by contrast, each record was assigned to only one class. Records in which pharmaceutical active substances or tert-butyl alcohol were detected were classified as "direct health-risk indicator", records with preservative or colorant non-compliance as "potential chemical risk", and all other records as "adulteration and counterfeiting". Records with more than one reported non-compliance were classified according to the highest-risk non-compliance they contained. This classification is based solely on the nature of the reported non-compliance; it is not a toxicological risk assessment involving concentration and exposure data.

C. Statistical Analysis

Categorical data are presented as counts and percentages. The association between food group and the three-class public health classification was evaluated with the chi-square test. Because some cells had expected frequencies below 5, the p value was calculated by Monte Carlo simulation (2000 replicates), and the strength of the association was expressed with Cramér's V. Statistical significance was set at $p<0.05$. In addition, the combined share of the three food groups with the most records in 2022 was compared with published data for the 2016–2020 period [12].

III. RESULTS AND DISCUSSION

A. Distribution by Food Group

Of the 559 product records, 156 (27.9%) belonged to milk and dairy products, 133 (23.8%) to meat and meat products and 114 (20.4%) to vegetable oils. Together, these three groups accounted for 72.1% of all records. Food supplements represented 7.3%, bee products 4.5%, alcoholic beverages 4.1%, herbal, tea and coffee products 3.9%, and chocolate and cocoa products 3.8%. The remaining five food groups (non-alcoholic beverages, spices, energy drinks, confectionery, and bakery products and desserts) accounted for 4.3% (n=24) of all records (Fig. 1). The concentration of records in dairy products, meat products and vegetable oils is consistent with the literature, in which these products are reported to be among the food groups most vulnerable to counterfeiting and adulteration because of their high economic value and complex composition [2], [4], [6].

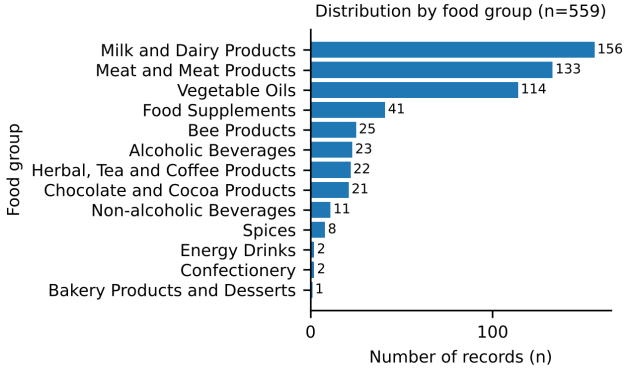


Fig. 1. Distribution of the 559 official public disclosure records by food group.

B. Non-Compliance Types and Public Health Classification

A detailed examination of the non-compliances revealed a more informative distribution than total record counts. The most frequent non-compliance was fat substitution (n=197, 35.2%), which includes records such as the blending of seed oils into olive oil and the presence of non-milk fat in dairy products. Meat species or offal substitution ranked second with 131 records (23.4%). Pharmaceutical active substances were detected in 82 records (14.7%), most of which were concentrated in food supplements, chocolate and cocoa products, and some beverages. Sildenafil was the most frequently detected active substance, found in 80 of the 82 records with pharmaceutical active substances (97.6%). Tadalafil was reported in 6 records, sibutramine in 4 and paracetamol in 3, while unapproved sildenafil analogues (dimethylsildenafil, thiodimethylsildenafil) were reported in 2 records. More than one active substance was found together in 11 of these 82 records (13.4%). The undeclared presence of these substances in foods or food supplements is of public health concern because of dose uncertainty and drug interactions [7]–[9].

TABLE I. CLASSIFICATION OF RECORDS FROM A PUBLIC HEALTH PERSPECTIVE

Class	n	%
Adulteration and counterfeiting	396	70.8
Potential chemical risk	66	11.8
Direct health-risk indicator	97	17.4

In the public health classification, 396 records (70.8%) were classified as adulteration and counterfeiting, 66 (11.8%) as potential chemical risk and 97 (17.4%) as direct health-risk indicators. Taken together, the direct and potential risk classes comprised 163 records, or 29.2% of the dataset. Of the 66 records in the potential chemical risk class, 41 contained natamycin and 25 contained colorant findings. Of the 97 records classified as direct health-risk indicators, 82 contained pharmaceutical active substances and 15 contained tert-butyl alcohol. Records in the adulteration and counterfeiting class were not considered a direct toxicological risk. However, two records in this class (peanut detected in a burma kadayıf sold as containing pistachio, and soy detected in cooked döner) are undeclared allergen findings that may pose a health risk to allergic individuals. In addition, undeclared species or ingredients are important for allergen management, traceability, nutritional quality and consumer choice [2], [4].

C. Public Health Distribution by Food Group

The public health classification of the records differed markedly between food groups (Fig. 2). All records for vegetable oils, meat and meat products, and bee products fell into the adulteration and counterfeiting class. In food supplements and in chocolate and cocoa products, by contrast, all records were classified as direct health-risk indicators. All non-alcoholic beverage, energy drink and confectionery records, and 65.2% of alcoholic beverage records, were also in this class. In

ISBN: 978-605-80193-6-2

herbal, tea and coffee products, 77.3% of records, and in spices all records, were in the potential chemical risk class. In milk and dairy products, 26.3% of records were classified as potential chemical risk because of preservative non-compliance.

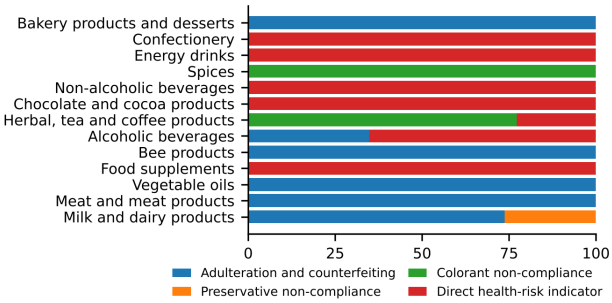


Fig. 2. Public health classification by food group (%). Preservative and colorant non-compliances together form the potential chemical risk class.

The distribution of the three classes differed significantly between food groups (χ^2=726.9; df=24; Monte Carlo p<0.001; Cramér's V=0.81). Direct health-risk indicators predominated in food supplements and in chocolate and cocoa products, potential chemical risks in herbal, tea and coffee products and in spices, and adulteration and counterfeiting records in meat and meat products and in vegetable oils. These results show that the non-compliance profile differs markedly by food group and support the importance of food-group-specific analytical panels in official controls. However, part of this difference may also arise from the different analyses applied to each food group.

D. Public Health Assessment and Analytical Findings

The most striking finding from a public health perspective is the concentration of pharmacologically active substances in food supplements and in some chocolates and beverages. The literature reports that phosphodiesterase-5 inhibitors such as sildenafil and tadalafil can be added to food supplements without declaration, that their amounts can vary widely, and that this can increase the risk of overdose and drug interactions [7]–[9]. Similarly, it has been emphasized that colorants that are not permitted or are used inappropriately may in some cases have genotoxic or carcinogenic properties, and that regular monitoring programs are therefore necessary [10]. Tert-butyl alcohol detected in alcoholic beverages is a substance used to denature ethyl alcohol, so its presence indicates the use of alcohol unfit for human consumption; the 15 related records were therefore classified as direct health-risk indicators. Preservative records, on the other hand, should not be interpreted at the same toxicological level. For example, natamycin is a preservative permitted only for surface treatment of certain cheeses and dried sausages [13]. The natamycin records in the announcement, however, mainly concern products outside this scope of use, such as yoghurt, ayran, processed cheese and fresh cheese. Because the announcement contains no information on substance levels, exposure-related health risk cannot be calculated from presence data alone [13].

The official announcement also contains analytical indicators specific to product groups. For vegetable oils, specific UV absorbance, sterol composition, fatty acid profile, trans fatty acids, ECN-42 difference and fatty acid ethyl esters were used. These indicators are particularly important for detecting the addition of seed oils to olive oil or blending with lower-quality oils [6]. For honey records, parameters such as C4 sugar content, δ13C indicators, proline, diastase activity, fructose/glucose ratio and maltose were used. The literature also recommends evaluating isotope ratio and compositional indicators together to verify the true composition of honey [5]. In meat products, species and tissue detections were prominent, while in dairy products non-milk fat, starch, gelatin and preservative findings stood out. This diversity shows that food-group-specific analytical panels are needed in risk-based official controls.

E. Comparison with Earlier Data

Yalım Kaya and Yayla [12] evaluated 2,112 products disclosed to the public in 2016–2020 and reported 625 records for vegetable oils, 584 for meat products and 559 for milk and dairy products. While these three groups made up 83.7% of the earlier dataset, their share in the 2022 data was 72.1%. This difference should not be interpreted as a decrease or increase in the true national prevalence of food fraud, because public disclosure data are affected by inspection priorities, sampling intensity, analytical capacity and disclosure policies.

F. Limitations of the Study

The products disclosed to the public are not a random sample of the Turkish food market, so the findings cannot be used to estimate national prevalence. Several product records may belong to the same company or brand, and inspection intensity is not equal across product groups. In addition, because the official list does not include the concentration of non-compliant substances, consumption amounts or individual exposure, no quantitative toxicological risk calculation could be made. The three-class classification used in this study was developed to compare records from a public health perspective; it does not replace the official regulatory classification or a toxicological risk assessment.

Despite these limitations, record-level classification went beyond total notification counts and showed which non-compliances are concentrated in which food groups and in which products records associated with health risks are more frequent. This approach can help plan laboratory capacity and official control resources on the basis of product-group-specific findings.

IV. CONCLUSION

The evaluation of the 559 official public disclosure records from 2022 showed that milk and dairy products, meat and meat products, and vegetable oils accounted for 72.1% of all records. The most common non-compliances were fat substitution, meat species or offal substitution, and pharmaceutical active substances. Of the records, 70.8% were classified as adulteration and counterfeiting, 11.8% as potential chemical risk and 17.4% as direct health-risk indicators, and this

ISBN: 978-605-80193-6-2

distribution varied markedly by food group. Direct health-risk indicators predominated in food supplements, chocolate and cocoa products and some beverages, whereas adulteration and counterfeiting records dominated in vegetable oils and meat products. The findings show that analyzing official food control data by product group and type of non-compliance together is useful for inspection and analytical monitoring strategies.

ACKNOWLEDGMENT

The author thanks the Republic of Türkiye Ministry of Agriculture and Forestry for making official food control and public disclosure data publicly available.

References

[1] J. Spink and D. C. Moyer, "Defining the public health threat of food fraud," Journal of Food Science, vol. 76, no. 9, pp. R157–R163, 2011. https://doi.org/10.1111/j.1750-3841.2011.02417.x

[2] K. Everstine, J. Spink, and S. Kennedy, "Economically motivated adulteration (EMA) of food: Common characteristics of EMA incidents," Journal of Food Protection, vol. 76, no. 4, pp. 723–735, 2013. https://doi.org/10.4315/0362-028X.JFP-12-399

[3] L. Manning and J. M. Soon, "Food safety, food fraud, and food defense: A fast evolving literature," Journal of Food Science, vol. 81, no. 4, pp. R823–R834, 2016. https://doi.org/10.1111/1750-3841.13256

[4] N. Z. Ballin, "Authentication of meat and meat products," Meat Science, vol. 86, no. 3, pp. 577–587, 2010. https://doi.org/10.1016/j.meatsci.2010.06.001

[5] S. Soares, J. S. Amaral, M. B. P. P. Oliveira, and I. Mafra, "A comprehensive review on the main honey authentication issues: Production and origin," Comprehensive Reviews in Food Science and Food Safety, vol. 16, no. 5, pp. 1072–1100, 2017. https://doi.org/10.1111/1541-4337.12278

[6] J. Lozano-Castellón et al., "Extra virgin olive oil: A comprehensive review of efforts to ensure its authenticity, traceability, and safety," Comprehensive Reviews in Food Science and Food Safety, vol. 21, no. 3, pp. 2639–2664, 2022. https://doi.org/10.1111/1541-4337.12949

[7] T. Rocha, J. S. Amaral, and M. B. P. P. Oliveira, "Adulteration of dietary supplements by the illegal addition of synthetic drugs: A review," Comprehensive Reviews in Food Science and Food Safety, vol. 15, no. 1, pp. 43–62, 2016. https://doi.org/10.1111/1541-4337.12173

[8] B. J. Venhuis and D. de Kaste, "Towards a decade of detecting new analogues of sildenafil, tadalafil and vardenafil in food supplements: A history, analytical aspects and health risks," Journal of Pharmaceutical and Biomedical Analysis, vol. 69, pp. 196–208, 2012. https://doi.org/10.1016/j.jpba.2012.02.014

[9] E. Petkova-Gueorguieva et al., "Survey on sildenafil, tadalafil, and vardenafil concentrations in food supplements for erectile dysfunction," International Journal of Analytical Chemistry, vol. 2022, Article 3950190, 2022. https://doi.org/10.1155/2022/3950190

[10] M. Oplatowska-Stachowiak and C. T. Elliott, "Food colors: Existing and emerging food safety concerns," Critical Reviews in Food Science and Nutrition, vol. 57, no. 3, pp. 524–548, 2017. https://doi.org/10.1080/10408398.2014.889652

[11] Republic of Türkiye Ministry of Agriculture and Forestry, "Kamuoyu duyurusu: 1 Mart 2022 tarihli 27. kamuoyu duyuru listesi [Public announcement: 27th public disclosure list of 1 March 2022]," 2022 (in Turkish).

[12] S. Yalım Kaya and F. Yayla, "Resmi verilere göre Türkiye'de taklit ve tağşiş yapılan yiyecek ve içecekler [Counterfeit and adulterated foods and beverages in Türkiye according to official data]," Journal of Tourism and Gastronomy Studies, vol. 8, no. 4, pp. 3108–3128, 2020 (in Turkish). https://doi.org/10.21325/jotags.2020.754

[13] EFSA Panel on Food Additives and Nutrient Sources added to Food (ANS), "Scientific Opinion on the use of natamycin (E 235) as a food additive," EFSA Journal, vol. 7, no. 12, Article 1412, 2009. https://doi.org/10.2903/j.efsa.2009.1412

 ISBN: 978-605-80193-6-2

Sponsor

DURU BULGUR

Karaman, Türkiye

Duru Gıda, headquartered in Karaman, is one of Türkiye's long-established and leading food companies, and the only company that aims both to bring bulgur back to Turkish tables and to introduce it to kitchens around the world.

Founded in 1935 by Ziya Duru, the company has been committed since its establishment to delivering its products to consumers with consistent quality. Today, with 400 experienced employees, three bulgur production facilities and a daily production capacity of 300 tonnes, it supplies consumers with natural and healthy products. With 38 dealers and 6 regional directorates in Türkiye and a wide distribution network abroad, Duru Gıda exports products under its own brand to 58 countries, including all European Union member states, the United States, the United Kingdom, Australia, Canada and Japan.

In addition to bulgur, Duru Bulgur also processes and packages natural and healthy pulses and rice products. Its commitment to quality is demonstrated by its ISO 9001, ISO 22000, BRC Food, IFS Food, ISO 14001, ISO 45001, ISO 50001 and ISO 27001 certifications. Duru products also hold Halal and Kosher certificates.

www.durubulgur.com

ISBN: 978-605-80193-6-2

INTERNATIONAL CONFERENCE ON ENGINEERING AND REGIONAL DEVELOPMENT

October 22–24, 2026 · Karaman, Türkiye

Conference Website

www.kmu.edu.tr/intcerd

KARAMANOĞLU MEHMETBEY UNIVERSITY

Karaman, Türkiye