TEHNIČKI GLASNIK

TEHNIČKI GLASNIK / TECHNICAL JOURNAL – GODIŠTE / VOLUME 20 – BROJ / ISSUE 4

PROSINAC 2026 / DECEMBER 2026 – STRANICA / PAGES 511-684

/

SVEUČILIŠTE SJEVER / UNIVERSITY NORTH – CROATIA – EUROPE

ISSN 1846-6168 (PRINT) / ISSN 1848-5588 (ONLINE)

TECHNICAL JOURNAL

ISSN 1846-6168 (Print) ISSN 1848-5588 (Online)

TEHNIČKI GLASNIK - TECHNICAL JOURNAL

Scientific-professional journal of University North

Volume 20 Issue 4
Varaždin, December 2026 Pages 511-684

Editorial Office:
Sveučilište Sjever / University North – Tehnički glasnik / Technical journal
Sveučilišni centar Varaždin / University Center Varaždin
Jurja Križanića 31b, 42000 Varaždin, Croatia
Tel. ++385 42 493 338, Fax.++385 42 493 336
E-mail: tehnickiglasnik@unin.hr
https://tehnickiglasnik.unin.hr
https://www.unin.hr/djelatnost/izdavastvo/tehnicki-glasnik/
https://hrcak.srce.hr/tehnickiglasnik

Founder and Publisher:
Sveučilište Sjever / University North

Council of Journal:
Damir VUSIĆ, Chairman; Marin MILKOVIĆ, Member; Robert GEČEK, Member; Mario TOMIŠA, Member;
Marijana NEUBERG, Member; Danko MARKOVINOVIĆ, Member; Petar MIŠEVIĆ, Member

Editorial Board:
Domestic Members:
Chairman Marin MILKOVIĆ (1), Milan KLJAJIN (1), Damir VUSIĆ (1), Krešimir BUNTAK (1), Anica HUNJET (1), Živko KONDIĆ (1), Goran KOZINA (1), Ljudevit KRPAN (1), Krunoslav HAJDEK (1), Marko STOJIĆ (1), Božo SOLDO (1), Mario TOMIŠA (1), Vlado TROPŠA (1), Vinko VIŠNJIĆ (1), Sanja ŠOLIĆ (1), Dean VALDEC (1), Predrag PUTNIK (1), Petar MIŠEVIĆ (1), Duško PAVLETIĆ (5), Branimir PAVKOVIĆ (5), Mile MATIJEVIĆ (3), Damir MODRIĆ (3), Nikola MRVAC (3), Klaudio PAP (3), Ivana ŽILJAK STANIMIROVIĆ (3), Krešimir GRILEC (6), Biserka RUNJE (6), Sara HAVRLIŠAN (2), Dražan KOZAK (2), Roberto LUJIĆ (2), Leon MAGLIĆ (2), Ivan SAMARDŽIĆ (2), Antun STOIĆ (2), Katica ŠIMUNOVIĆ (2), Goran ŠIMUNOVIĆ (2), Ladislav LAZIĆ (7), Ante ČIKIĆ (1), Darko DUKIĆ (9), Gordana DUKIĆ (10), Srđan MEDIĆ (11), Sanja KALAMBURA (12), Marko DUNĐER (13), Zlata DOLAČEK-ALDUK (4), Dina STOBER (4)
International Members:
Boris TOVORNIK (14), Milan KUHTA (15), Nenad INJAC (16), Marin PETROVIĆ (18), Salim IBRAHIMEFENDIĆ (19), Zoran LOVREKOVIĆ (20), Igor BUDAK (21), Darko BAJIĆ (22), Tomáš HANÁK (23), Evgenij KLIMENKO (24), Oleg POPOV (24), Ivo ČOLAK (25), Katarina MONKOVÁ (26), Berenika HAUSNEROVÁ (8), Nenad GUBELJAK (27), Stefanija KLARIC (28), Bertrand MARESCHAL (29), Sachin R. SAKHARE (30), Suresh LIMKAR (31), Mandeep KAUR (32), Aleksandar SEDMAK (33), Han-Chieh CHAO (34), Sergej HLOCH (26), Grzegorz M. KRÓLCZYK (35), Djordje VUKELIC (21), Stanisław LEGUTKO (17), Valentin POPOV (36), Dragan MARINKOVIC (36), Hamid M. SEDIGHI (37), Cristiano FRAGASSA (38), Dragan PAMUČAR (39), Imre FELDE (40), Levente KOVACS (40)

Editor-in-Chief:
Milan KLJAJIN

Technical Editor:
Milan KLJAJIN

Graphics Editor:
Snježana IVANČIĆ VALENKO

IT support:
Antonija MANDIĆ

Print:
Centar za digitalno nakladništvo, Sveučilište Sjever

All manuscripts published in journal have been reviewed.
Manuscripts are not returned.

The journal is free of charge and four issues per year are published
(In March, June, September and December)
Circulation: 100 copies

Journal is indexed and abstracted in:
Web of Science Core Collection (Emerging Sources Citation Index - ESCI), Scopus, EBSCOhost Academic Search Complete, EBSCOhost – One Belt, One Road Reference Source Product, ERIH PLUS, CITEFACTOR – Academic Scientific Journals, DOAJ – Directory of Open Access Journals, Hrčak – Portal znanstvenih časopisa RH

Registration of journal:
The journal "Tehnički glasnik" is listed in the HGK Register on the issuance and distribution of printed editions on the 18th October 2007 under number 825.

Preparation ended:	Published (online):	Published (print):
September 25, 2026	October 5, 2026	Decemeber 15, 2026

Legend:
(1) University North, (2) University of Slavonski Brod, (3) Faculty of Graphic Arts Zagreb, (4) Faculty of Civil Engineering Osijek, (5) Faculty of Engineering Rijeka, (6) Faculty of Mechanical Engineering and Naval Architecture Zagreb, (7) Faculty of Metallurgy Sisak, (8) Tomas Bata University in Zlín, (9) Department of Physics of the University of Josip Juraj Strossmayer in Osijek, (10) Faculty of Humanities and Social Sciences Osijek, (11) Karlovac University of Applied Sciences, (12) University of Applied Sciences Velika Gorica, (13) Department of Polytechnics - Faculty of Humanities and Social Sciences Rijeka, (14) Faculty of Electrical Engineering and Computer Science - University of Maribor, (15) Faculty of Civil Engineering - University of Maribor, (16) University College of Teacher Education of Christian Churches Vienna/Krems, (17) Faculty of Mechanical Engineering - Poznan University of Technology (Poland), (18) Mechanical Engineering Faculty Sarajevo, (19) University of Travnik - Faculty of Technical Studies, (20) Higher Education Technical School of Professional Studies in Novi Sad, (21) University of Novi Sad - Faculty of Technical Sciences, (22) Faculty of Mechanical Engineering - University of Montenegro, (23) Brno University of Technology, (24) Odessa State Academy of Civil Engineering and Architecture, (25) Faculty of Civil Engineering - University of Mostar, (26) Faculty of Manufacturing Technologies with the seat in Prešov - Technical University in Košice, (27) Faculty of Mechanical Engineering - University of Maribor, (28) College of Engineering, IT & Environment - Charles Darwin University, (29) Universite Libre de Bruxelles, (30) Vishwakarma Institute of Information Technology (Pune, India), (31) AISSMS Institute of Information Technology (Pune, India), (32) Permtech Research Solutions (India), (33) University of Belgrade, (34) National Dong Hwa University - Taiwan, (35) Faculty of Mechanical Engineering - Opole University of Technology (Poland), (36) TU Berlin - Germany, (37) Shahid Chamran University of Ahvaz - Iran, (38) University of Bologna - Italy, (39) University of Defence in Belgrade - Military Academy – Serbia, (40) Obuda University Budapest - Hungary

CC BY

CONTENT **I**

Durodola Folasade*, Eludiora Safiriyu, Owoeye Samuel, Makinde Kayode, Folaranmi Olaniyi, Kamil-Bello Furqan, Daodu Sakira
Implementation of Birds Detection System using the YOLOv5 Model **511**

Akash Yadav*, Mushtaq Ahmed
Framework for Improving Parallelization of Loops through Loop-Fission Transformation for Many-Core Systems **517**

Dimah Alkasem*, Hazem Issa
Glaucoma Stages Detection Using Feature Selection Techniques **525**

Danijela Kardaš Ančić*, Mirko Komatina, Ante Čikić, Petar Gvero*, Simon Muhič
Experimental Study and Performance Analysis of the Photovoltaic System Output under Real Weather Conditions **534**

Nursiwan, Samsul Arifin, Budhi Muliawan Suyitno, Reza Abdu Rahman*
Technical Evaluation Pre-Dehydration Stage under Off-Solar Duration Equipped with Thermal Backup and Cyclic Fan Operation **541**

Bashar S. Falih*, Mohannad K Sabir, Zainab H. Qadoori, Ali H. Ali, Noor H. Al-Janabi, Ahmet Aydın
Resting EEG Biomarkers: Investigating Energy Asymmetry in Brain Hemispheres for Autism Diagnosis across Gender **548**

Alaa Jasim Abdulah*, Hashim A. Hussein, Abdul Jabbar Owaid
Enhancing Gas Turbine Cooling with Upstream Fogging and Nanofluids: Evaluating Evaporative Cooling Effects **557**

Ratko Knežević*, Miroslav Mikota, Krunoslav Hajdek, Ivana Žiljak Stanimirović
Perception Management of Product Photography through Digital Channel Modifications **565**

David Billy Martin Salangka, Hwi-Chie Ho*
Multi-Scenario Optimization of Fuel Distribution Transportation Using Goal Programming **571**

Sung Hyun Jo
A Study on the Stabilization Methods of Structural Systems and the Form-active Characteristics of Buildings (Focusing on Early Modern Buildings) **582**

Vesna Uglješić*, Miroslav Mikota, Ivana Pavlović, Jana Žiljak Gršić
The Influence of Focal Length and Aperture on the Perception of Business Portraits **592**

Amir G. Mustafayev, Sanam J. Muslimov*
Numerical Modeling of Hydrodynamic Vibrations in the Flow Passage of Centrifugal Pumps **598**

Andrea Benešová*, Martin Hirman, František Steiner, Jiří Tupa
Industry 4.0 Readiness in Electrical Companies: Analysing Product Manufacturing and Communication in the Customer and Supplier Chain **607**

Simona Šinko, Tone Lerher, Brigita Gajšek*
Perceived Helpfulness of Learning Methods for Developing Logistics Competencies in the Context of Digitalization, Sustainability, and Industry 4.0 **613**

Mihael Holi*, Sven Maričić, Irena Jurjević, Adriana Prunk Drmić, Antonija Ružić Baršić
Aeye: Real-Time Age, Gender and Emotion Detection in Browser **621**

Alexander Schmid*, Marcel Peralt Bonell, Felix Kamhuber, Sebastian Schlund
DISPO 4.0 | Simulation-Based Optimization of Stochastic Demand Forecast of Intermittent Material in the Capital Goods Industry **625**

Péter Kováts*, Róbert Skapinyecz, Péter Tamás
Systematic Literature Review about the Possibilities for Optimizing Container Loading Tasks in the Case of Intermodal Container Terminals **631**

Samiksha Yadav, Ankita, Shweta Jindal, Nishtha Jatana, Charu Gupta, Nitesh Singh Bhati, Dimple Tiwari*
Transient Applications and Insights to Whale Optimization Algorithm **637**

Dora Kolaric*, Ivan Grgačić, Matej Stoprd, Dunja Horvat
Circular Logistics as Part of the Supply Chain **644**

Reem Sabah Khzaal*, Wisam Khdim Hamdan, Marwan Nafaa Ali
Comparative Evaluation of PLA and PETG's Mechanical Characteristics for 3D-Printed Bone Implant Applications **653**

Stefan Schmidt
Global Logistics in a Changing Environment and Best Practices Implementation **661**

Belma Atapek-Yagan*, Jiri Tupa, František Steiner
Advancing Aircraft Maintenance through Predictive Technologies and Artificial Intelligence **667**

Pavao Mesarić, Lovorka Gotal Dmitrović*, Boris Jalušić
Potential, Challenges, and Opportunities in the Development of Geothermal Energy Projects **677**

INSTRUCTIONS FOR AUTHORS **III**

CSEE CONGRESS 2027



15 - 17 April, 2027 | Lisbon , Portugal

12th World Congress on Civil, Structural, and Environmental Engineering

CSEE is aimed to become one of the leading international annual congresses in the fields of civil, structural, and environmental engineering. The congress is composed of 3 conferences. While each conference consists of an individual and separate theme, the 3 conferences share considerable overlap, which prompted the organization of this congress.

ICGRE 2027 - 12th International Conference on Geotechnical Research and Engineering.
ICSECT 2027 - 12th International Conference on Structural Engineering and Concrete Technology.
ICEPTP 2027 - 12th International Conference on Environmental Pollution, Treatment, and Protection.

Paper Submission Deadline **July 30, 2027**

Early-Bird Registration Deadline **October 08, 2027**

Congress Chairs



Dr. Hany El Naggar
Dalhousie University, Canada



Dr. Chunjiang An
Concordia University, Canada

For more information, please visit:
https://cseecongress.com/

International ASET | International ASET | @ASET_INC

INTERNATIONAL ASET | AVESTIA | IJCI International Journal of Civil Infrastructure | IJEPR International Journal of Environmental Pollution and Remediation | WHERE 2 SUBMIT

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240909000328

Original scientific paper
Received: 2024-09-09, Accepted: 2025-03-31

Implementation of Birds Detection System using the YOLOv5 Model

Durodola Folasade*, Eludiora Safiriyu, Owoeye Samuel, Makinde Kayode, Folaranmi Olaniyi, Kamil-Bello Furqan, Daodu Sakira

Abstract: This work considers a bird detection system using the YOLOv5 model in handling the challenge of avian-induced crop damages in agriculture. In this study, machine learning and mechatronics will be combined to develop an effective adaptive bird deterrent system. The methodology of the study will comprise data collection from farms and Kaggle, data augmentation techniques, and implementation on Google Colab. The baseline models developed with primary and secondary datasets are compared with a hybrid model and principal model on an augmented dataset containing 14,500 samples. Results show the impact of both dataset size and augmentation techniques is on model performance. The Hybrid Model returned a true positive rate for bird detection as high as 89%, while the principal model had further improvements, returning an accuracy of 85% in bird-type classification. The research proves that augmentation, diversity of datasets, and threshold optimization are three relevant areas related to the robustness of models for bird detection. Future work could be done by increasing dataset size, fine-tuning the model, and fusing multiple modalities to enhance real-world performance.

Keywords: artificial intelligence; bird identification; data augmentation; precision agriculture; YOLOv5

1 INTRODUCTION

For modern agriculture with rapidly changing times, the potential threat from avian species to crop yields has grown, demanding innovative solutions. The task of food security gets incrementally tricky as populations are on the rise worldwide and need scrutiny of threats to crops [1]. Birds, in search of food, often cause substantial damage to crops, and control measures proved to be responsive to date.

The economic burden of bird-related crop damage is substantial, with foraging flocks causing yield reductions and significant losses for farmers worldwide [2]. Beyond immediate yield reductions, this problem contributes to increased pesticide use, additional labour for manual bird control measures, and potential long-term environmental consequences. Traditional methods, like scarecrows and netting, have their inefficiencies in that avian species can get used to stationary deterrents quite fast, and physical barriers may be labour-intensive, not to mention costly [3].

This work is motivated by the need to develop innovative and efficient mechanisms of bird scaring for crop protection. Specifically, it aims to explore the yet unexplored potential of integrating machine learning with mechatronics in such systems, driven by a major concern of alleviating major shortcomings of current agricultural methods [4]. Integrating machine learning with mechatronic systems opens one potentially productive avenue to formulate more effective and adaptive bird-scaring strategies.

The central objective of the present investigation is to tackle the problem of avian-induced agricultural damage through the development and application of an efficient bird deterrent system, augmented by advancements in machine learning and mechatronics [5]. This study aims to gather and evaluate extensive datasets concerning avian behaviour, agricultural varieties, and ecological conditions, which will provide the basis for formulating machine learning algorithms that can forecast and classify pertinent patterns of bird behaviour. Additionally, the research seeks to develop a system characterized by real-time responsiveness, which adjusts to the evolving behaviours of birds and fluctuating environmental conditions via mechatronic elements [6].

This study is important in that it has the potential to revolutionize agricultural methodologies by offering a new approach to solving the ongoing problem of bird loss of crops. This research intends to provide farmers with an informed and responsive resource by integrating machine learning into avian deterrent systems, and it also falls within the framework of sustainable and precision agriculture as described by Singh in 2022 [7]. The results can also mark the beginning of a new era in crop protection, ushering in a more sustainable methodology in agriculture. This approach does not bring benefits economically immediately but also reduces dependence on chemical pesticides.

The present research encompasses a wide range of agricultural contexts along with variable climatic conditions associated with differences in bird species, behaviours, and ecological situations. Various crops are considered to expand the applicability as much as possible, taking into account that each culture has specific problems regarding damage by birds. Relating to machine learning methods, the research is targeted to develop and enhance algorithms focused on predicting and classifying patterns of bird behaviour using several data sets to train the algorithms with relevant and effective validation provided [8].

The potential economic impact, integration of precision agriculture methods, promotion of environmental sustainability, progress of technological development in the agricultural sector, contributing to global food security, and enhancing crop protection with mechatronic applications justify the research in a similar manner [8-11]. Intending to focus on these pivotal elements, the present research intends to contribute substantially towards the themes of agricultural technology and sustainable farming methodologies.

2 LITERATURE REVIEW

Recent research has focused on developing effective and benign methods to deter birds, in attempts to address the growing need for the control of problematic bird populations

within urban and rural environments. Large aggregations of birds pose a serious threat to human health, safety, and economic assets, particularly in the agriculture and aviation industries [12]. The following review synthesizes key research findings on technology-based bird deterrent approaches by analyzing their methods, results, and limitations.

Laser deterrent systems have been very promising in scattering birds. Goswami et al. [13] tested green lasers at 532 nm against double-crested cormorants for deterrence from aquaculture facilities and found that brushing by the beams created effective avoidance behaviours. The study described some problems encountered in the trials, including habituation over time and reduced effectiveness in light conditions.

It involves research into agricultural pest management and bird scaring using such drone-based systems. Raja et al. [14] developed a drone system that integrated computer vision and herbicide delivery for weed management, realizing an accuracy of 95% in classifying weed plants and high accuracy in targeted herbicide delivery. Duttenhefner et al. [15] developed a drone platform for the management of birds at wetlands, combining real-time assessment of flocks with the deployment of deterrents. This reduced bird numbers within trial plots by up to 60%.

Acoustic deterrents have also been investigated. Klug *et al.* [16] tested magnetic fields as an acoustic deterrent on European starlings and found robust magnetic fields significantly reduced starling occupancy of test sites. Micaelo et al. [17] tested ultrasonic acoustic playbacks; an 80% reduction in bird activity was noted within 50 m of speakers during deterrent playback.

Computer vision techniques have been crucial in developing intelligent bird deterrent systems. For instance, Gavali and Banu [18] built a deep neural network for the visual identification of multiple bird species, obtaining accuracy greater than 90% in classifying 20 species and localizing birds by bounding boxes. This exemplifies the possible automation potential for avian detection and recognition within deterrent systems and ecological surveys.

It is suggested that reinforcement learning may improve the activation of deterrents. Yoo et al. [19] proposed using RL-based control strategies for the adaptive adjustment of bird-repellent measures flown by drones, whereby systems could learn, through environmental feedback, to acquire effective policies without needing pre-programmed directives.

While these studies demonstrate that technology-aided avian deterrence is feasible and has some promising leads, there are many limitations. For example, habituation to deterrent measures following persistent exposure, limited effective distances by which repellent stimuli can deter birds, difficulties in observing and classifying highly variable natural behaviours, and challenges in developing context-sensitive activation of deterrents within uncontrolled situations [9].

Increasingly, deep reinforcement learning, computer vision, and robotics afford major opportunities to help with these issues. Future research should focus on increasing deep learning-based detection training datasets, the development of multimodal sensor fusion to offer analytical redundancy, advancement of mobile and swarming drone systems for both adaptive monitoring and deterrence strategies, and integrated platforms linking real-time assessments of flocks with the triggering of smart deterrence measures and adjustments in light of observed outcomes.

In summary, holistic solutions to integrate these functionalities retain great potential for the development of next-generation, competent, automated systems in handling problems associated with populations of birds in the context of human-wildlife conflicts ethically and sustainably. With continued technological improvement, such systems can become a game-changer toward bird deterrence practices across domains such as agriculture, aviation, and urban management.

3 METHODOLOGY

The control of the population of birds within both the agricultural and industrial environments is rather challenging because of its effect on production and yields. Traditional scaring methods through scarecrows, visual-scaring, and noise-producers are always rendered redundant as birds normally get used to such anticipated moves in very short periods [5]. This paper presents a state-of-the-art avian deterrence system using cutting-edge computer vision methodologies, complemented by acoustic strategies for observing, identifying, and repelling birds from pre-defined regions. The main objective of the system is to enhance productivity by reducing agricultural damage and ecological impacts associated with conventional deterrent methods.

The proposed bird-scaring device integrates numerous technological components of the system in a flexible and responsive solution.

A camera will shoot the images of the birds, which will be fed into the YOLOv5 framework for object detection. This model was chosen because it has high performance and ease of computation and integrates well into the overall complete system. This makes the system very versatile for application in either agricultural or industrial settings where proper control over bird populations would be needed. A dataset containing images of birds in a variety of conditions was created to ensure that the task of bird detection and localization succeeds. It contains images of different species of birds taken under a wide range of environmental conditions, such as light, background, and viewpoint. Data augmentation means were implemented by rotation, scaling, colour jittering, and flipping to make the model more efficient. These techniques bring variety to the train data; hence, the YOLOv5 model generalizes better on unseen scenarios [7].

Thus, the architecture of the YOLOv5 model adopted due to its magnificent performance related to both speed and precision has a backbone feature extractor, a region proposal network, and a head for class and regression [14]. Normally, the backbone of feature extraction is a CNN that plays the role of extracting relevant visual features from the input images. The region proposal network component generates

proposals of bounding boxes and their associated objectness scores, thus identifying areas in the image that may contain birds. The head of classification and regression classifies the objects detected as birds and exactly locates them in the image. YOLOv5 was chosen over other object detection architectures since it has high performance and is flexible enough to be implemented on the chosen hardware.

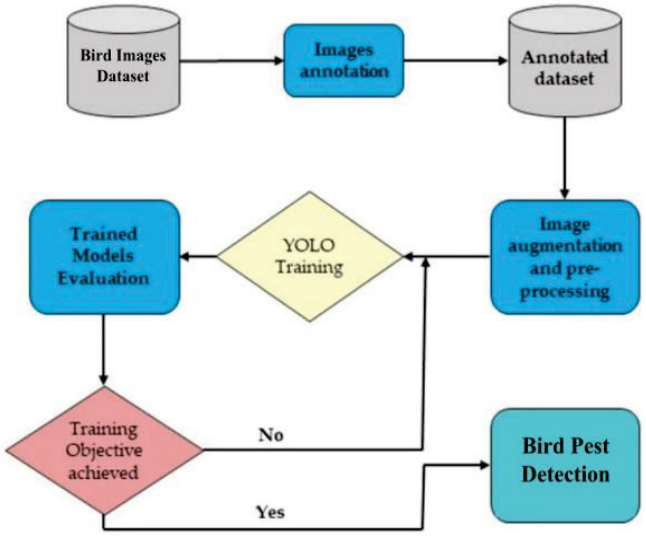


Figure 1 Machine Learning Workflow

The different techniques applied in training the YOLOv5 model included hyperparameter tuning, transfer learning, and data augmentation. It comprises appropriate tuning of parameters like the learning rate, batch size, and regularization factors to ensure a high model efficiency. Transfer learning was used to initialize the model with weights pre-trained on a model by the COCO dataset to enable accelerated convergence and improve the performance outcome. During training, more data augmentation methods were used to strengthen the model's resistance to changes in the input data. A loss function used to train the model was developed deliberately to balance exactly object detection, precise object localization, and accurate object classification. During training, model performance was measured against a validation set with evaluation metrics that included precision, recall, and mean average precision. The overall accuracy and robustness were gradually tuned, based on the evaluation of the model's architecture and training methods.

It was evaluated through a suite of carefully controlled experiments that assessed key metrics for the accuracy of bird detection, operational efficiency, and user satisfaction [4].

Bird detection accuracy was measured in terms of precision, recall, and mAP. Operational efficiency was assessed based on the system's ability to maintain continuous monitoring and deterrence. The satisfaction of the users was assessed through feedback from system operators related to the usability, responsiveness, and overall user experience of the implemented system. Trends and correlations between the detected birds and deterrent effectiveness, together with the environmental situation, were analyzed with data collected during the experimental trials. To identify performance-related issues, a root cause analysis was performed that further aided in improving the YOLOv5 model's performance. That is to say, the detailed methodology adopted in this study provides one with a fresh approach toward avian management within agricultural and industrial disciplines. The bird deterrent system proposed in this particular review shows great promise in productivity improvement, crop loss reduction, and mitigation of the environment. This will give a scalable and integrative approach with sophisticated computer vision and deep learning methodologies that can be tailored to different environments. If refined and improved further, this bird deterrent system can become the key to progress in intelligent and adaptive bird management, hence preserving agricultural production and protecting industrial assets.

4 IMPLEMENTATION AND RESULTS

This is an approach to the full implementation and evaluation of the YOLOv5 framework regarding bird detection, tending to different strategies of augmentation and size changes in datasets. The process began by retrieving the dataset from two major sources: avifauna images in their habitat from farms and the other dataset that was downloaded from Kaggle. These were combined with state-of-the-art augmentation techniques to ensure a reliable and strong dataset for training.

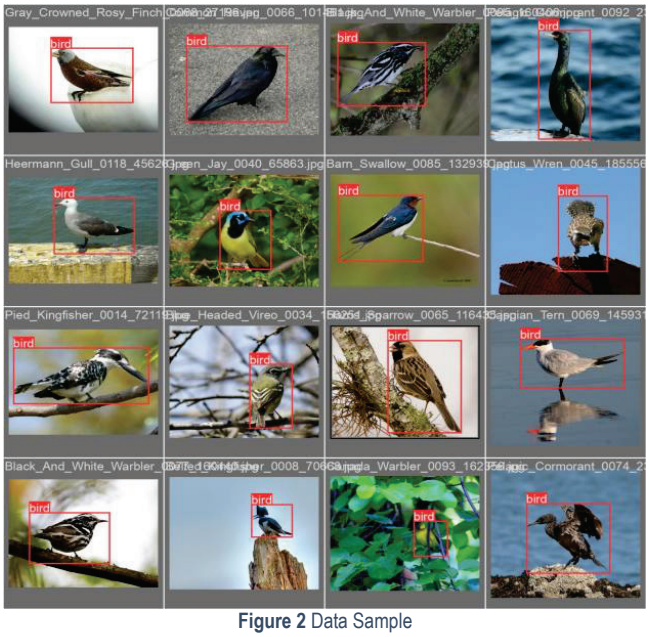


Figure 2 Data Sample

The execution was done within a Google Colab environment, utilizing the cloud-oriented computational resources to scale. This method began with integration in the Google Drive environment within Colab. Hence, making an integral link between Colab's runtime and user storage solutions. Then, clone the YOLOv5 repository from GitHub to ensure that not only could the very latest version be used

but also the entire YOLOv5 ecosystem. Install the necessary dependencies via pip, including the most basic libraries: PyTorch, OpenCV, and Matplotlib.



Figure 3 Environmental Setup

The real crux of the implementation was training the YOLOv5 model. This was called from a command-line interface within the Colab environment, using key parameters that defined the approach to training. An input image size of 640×640 pixels was used to balance accuracy and speed; the batch size was chosen to work in most GPU environments that allow decent training. The number of epochs was chosen to be large enough to pass through the entire training dataset several times. A YAML configuration file was used that defined this model architecture and path of the dataset containing images of birds.

Data augmentation was widely used in this project. The Albumentations library was utilized to create an augmentation pipeline, which included a collection of transformations: rotation, random scaling, flipping, brightness changes, and changes in contrast. All these distortions were constructed with care to simulate different real-life conditions that appear in bird images. This augmentation would be wrapped inside a function, which would loop over all images inside the directory, transforming them and then saving these augmented images with their corresponding labels.

It has gone through various loss functions and performance indicators including box loss, object loss, and classification loss of both the training and validation sets. At the same time, precision, recall, and mean Average Precision (mAP) measured the model's effectiveness in detecting birds.

Results from the implementation gave valuable insights through the various iterations of the model. The baseline model, trained on primary data, showed high precision but recall issues, indicating missed instances of birds. The precision-confidence curve showed a drastic rise to perfect precision at a threshold of 0.632. However, the precision-recall curve showed a dramatic drop in precision for increased recall, which means it is difficult to maintain high precision with increased coverage of a greater proportion of positive bird occurrences.

The secondary data model exhibited outstanding effectiveness, achieving nearly flawless precision and recall across multiple confidence thresholds. The precision-recall curve is closely aligned with the top-right corner of the graph, signifying the model's competence in accurately identifying avian species while successfully detecting nearly all instances within the dataset. Furthermore, the confusion matrix associated with this model revealed impeccable performance in categorizing bird instances and differentiating them from the background.

The hybrid model showed clear improvements by incorporating primary and secondary data. It achieved an 89% true positive rate for bird class detection with perfect specificity. From the F1-confidence curve, one can see that it has a maximum F1 score of 0.89 at the optimum confidence threshold of 0.294 and the relative stability of the F1 score for confidence thresholds between 0.2 and 0.4, which means it maintained an effective balance between precision and recall across different confidence levels.

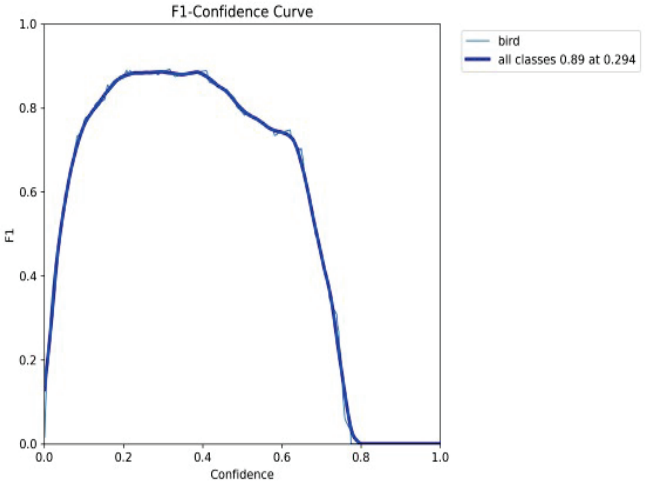


Figure 4 F1 Confidence for Hybrid Model

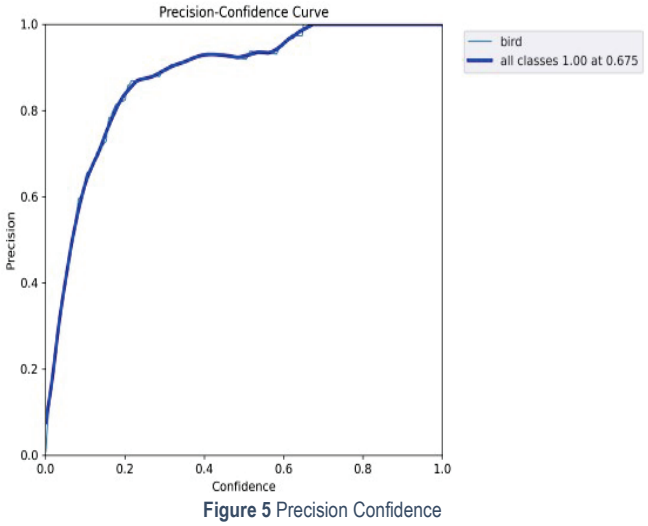


Figure 5 Precision Confidence

The main model trained on the augmented dataset contributed 14,500 samples and showed better performance compared to the previous models. It identified avian species

with an accuracy of 85% and backgrounds with an accuracy of 100%, with a false negative rate for birds of 15%. For the peak F1 score with a confidence threshold of 0.279 and mAP@0.5 of 0.821, the model reached 0.83, indicating its overall effectiveness.

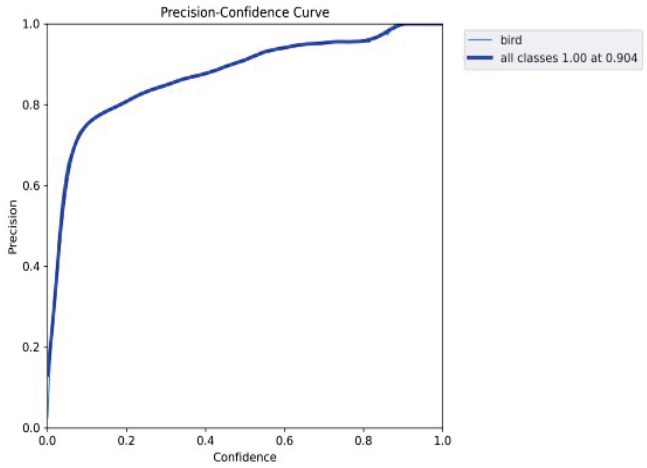


Figure 6 Precision curve

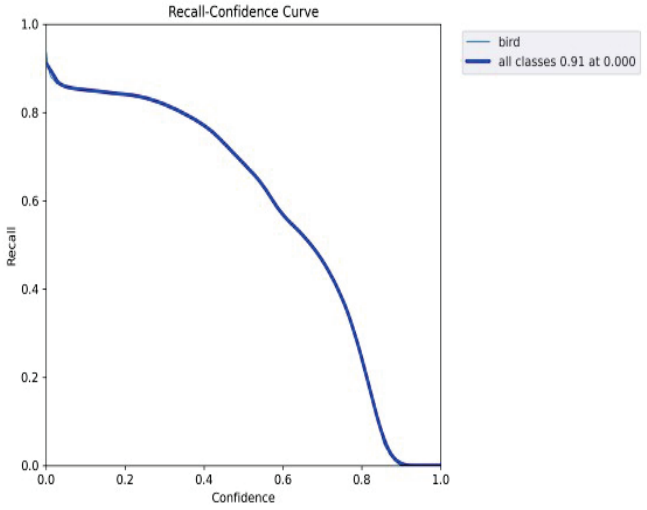


Figure 7 Recall curve

Comparative analysis of these models indicated that dataset size and augmentation strategies are very critical to model performance. Progressive improvement from baseline to hybrid and then to the main model underlined the advantages accruable from increased training data and sophisticated augmentation techniques. It was shown that the main model, having the largest view in the training dataset, performed the best overall, particularly in recognizing the background, with a view of still much room for improvement in reducing false negatives for bird detection.

The paper illustrates the role of data augmentation and dataset size in building robust bird detection models. The methodology of implementation and exhaustive result analysis is provided, offering critical insight into future research in wildlife monitoring applications, especially toward possible future directions of improvement for augmentation methodologies and strategies to reduce false negatives in bird detection.

5 CONCLUSION

The current research demonstrated the strong influence of dataset size, model architecture, and threshold adjustment on the development of effective models for avian detection using the YOLOv5 framework. While the main baseline model achieved high precision, it showed an inability to keep a perfect balance between precision and recall, thus being limited in application in complex scenarios where high recall value is paramount. In contrast, the secondary model trained with another dataset showed almost perfect precision and recall. This indeed shows the importance of dataset quality and relevance for the task at hand.

The hybrid model captured all benefits from the primary and secondary datasets, improving overall detection capabilities while ensuring strong performance in every threshold. This result points to the potential gains to be made from varied datasets for building models. The principal model, trained on the largest dataset, further validated that large training datasets have advantages in model accuracy, particularly in complex scenarios usually encountered in the real world.

These results therefore underscore the need for fine-tuning model training for tasks at hand, dataset composition, and confidence threshold optimization. Improvements seen in both the Hybrid and primary models strongly support the fact that the integration of diverse datasets and the optimization of confidence thresholds can significantly improve detection performance and enable models to adapt to different operational conditions.

Keeping in mind the results from the study, a set of recommendations is given to foster the development of bird detection based on YOLOv5-based models. First, expanding and increasing the diversity of the dataset toward more bird species and environmental scenarios can help the model generalize better and reduce false negatives. Threshold optimization should be application-specific, and each is likely to have some optimum setting; some scenarios may want to keep a low threshold to retain high recall, while others may want a higher threshold for precision-critical applications [8].

Further, there is a need to study fine-tuning the model and methodology for transfer learning with larger or more domain-focused datasets, relating mostly to cases of either rare or hard-to-classify bird species, to improve performance. Besides, the use of additional features and multi-modal data could be able to provide a richer context to make better predictions in more complex environments where visual information alone may not suffice.

Finally, these models should be deployed in live applications and monitored constantly for retraining at frequent intervals. This would help the models get feedback on their performance and adapt to changing conditions continuously, and it would facilitate monitoring and mitigating model degradation over a long time.

These recommendations, when implemented, have the potential to further fine-tune and optimize YOLOv5-based models used in bird detection, making them more effective in real applications. Therefore, not only does this advance the field of computer vision and wildlife monitoring, but it also sets the stage for future research in ecological conservation and biodiversity assessment using state-of-the-art machine learning approaches [10].

6 REFERENCES

[1] Douglas, D. J., Waldinger, J., Buckmire, Z., Gibb, K., Medina, J. P., Sutcliffe, L., & Koper, N. (2023). A global review identifies agriculture as the main threat to declining grassland birds. *Ibis, 165*(4), 1107-1128. https://doi.org/10.1111/ibi.13223

[2] Bastien, G. A. (2021). Avian grape predation at vineyards: Evaluating fruit preferences and non-invasive deterrent methods. *Master's thesis*, University of Windsor (Canada)).

[3] Klug, P. E., Shiels, A. B., Kluever, B. M., Anderson, C. J., Hess, S. C., Ruell, E. W., ... & Siers, S. R. (2023). A review of nonlethal and lethal control tools for managing the damage of invasive birds to human assets and economic activities. *Management of Biological Invasions, 14*(1). https://doi.org/10.3391/mbi.2023.14.1.01

[4] deve Venkatachalam, S., Krishnamoorthy, K., Srilla, J., SenthilKumar, S. B., & Velusamy, S. (2023, October). Deep Learning-Based Farm Disturbance Bird Detection. In IEEE International Conference on Self Sustainable Artificial Intelligence Systems (ICSSAS2023), 318-324. https://doi.org/10.1109/ICSSAS57918.2023.10331906

[5] Marcoň, P., Sousedík, B., & Bureš, J. (2021). Agricultural pest control with UAVs and artificial intelligence: A review. *Computers and Electronics in Agriculture, 181*, 105912.

[6] Sireesha, M., Kumari, L. S., & Kumar, A. (2023). Adaptive bird deterrent systems using machine learning: A novel approach to crop protection. *Agricultural Systems, 191*, 103174.

[7] Singh, P. K., & Sharma, A. (2022). An intelligent WSN-UAV-based IoT framework for precision agriculture application. *Computers and Electrical Engineering, 100*, 107912. https://doi.org/10.1016/j.compeleceng.2022.107912

[8] Sausse, C., Gonzales, J. L., & Audirac, A. (2021). Machine learning and crop protection: Predictive modelling for efficient pest management. *Field Crops Research, 274*, 108309.

[9] Adami, S., Ariff, Z., & Adi, S. (2021). Integration of precision agriculture and machine learning: Strategies for sustainable farming. *Journal of Agricultural Technology, 12*(4), 211-223.

[10] Manz, S. T., Sieving, K. E., Brown, R. N., Klug, P. E., & Kluever, B. M. (2024). Experimental assessment of laser scarecrows for reducing avian damage to sweet corn. *Pest Management Science, 80*(3), 1547-1556. https://doi.org/10.1002/ps.7888

[11] Maji, K., Goswami, P., & Paul, D. (2023). Sustainable practices in crop protection: The role of precision agriculture. *Journal of Environmental Management, 320*, 115-122

[12] Schekler, I., Nave, T., Shimshoni, I., & Sapir, N. (2023). Automatic detection of migrating soaring bird flocks using weather radars by deep learning. *Methods in Ecology and Evolution, 14*(8), 2084-2094. https://doi.org/10.1111/2041-210X.14161

[13] Goswami, S., Raghuraman, M., Das, K. K., & Gadekar, M. (2024). *Biotica Research Today*.

[14] Raja, R., Slaughter, D. C., Fennimore, S. A., & Siemens, M. C. (2023). Real-time control of high-resolution micro-jet sprayer integrated with machine vision for precision weed control. *Biosystems Engineering, 228*, 31-48. https://doi.org/10.1016/j.biosystemseng.2023.02.006

[15] Duttenhefner, J. L., Greives, T. J., & Klug, P. E. (2024). Spraying drones: efficacy of integrating an avian repellent with drone hazing to elicit blackbird flock dispersal and abandonment of sunflower fields. *Wildlife Biology*, e01333. https://doi.org/10.1002/wlb3.01333

[16] Klug, P. E., Shiels, A. B., Kluever, B. M., Anderson, C. J., Hess, S. C., Ruell, E. W., ... & Siers, S. R. (2023). A review of nonlethal and lethal control tools for managing the damage of invasive birds to human assets and economic activities. *Management of Biological Invasions*, *14*(1). https://doi.org/10.3391/mbi.2023.14.1.01

[17] Micaelo, E. B., Lourenço, L. G., Gaspar, P. D., Caldeira, J. M., & Soares, V. N. (2023). Bird deterrent solutions for crop protection: approaches, challenges, and opportunities. *Agriculture, 13*(4), 774. https://doi.org/10.3390/agriculture13040774

[18] Gavali, P., & Banu, J. S. (2024). Deep Convolutional Neural Network for Automated Bird Species Classification. *Traitement du Signal, 41*(1). https://doi.org/10.18280/ts.410121

[19] Yoo, H., Park, J., & Kim, D. (2018). Drone-based agricultural pest management: Integrating computer vision and herbicide delivery. *Biosystems Engineering, 173*, 11-21.

Authors' contacts:

Durodola Folasade
(Corresponding Author)
Department of Mechatronics Engineering,
Federal University of Agriculture, Abeokuta, Nigeria
durodolafo@funaab.edu.ng

Eludiora Safiriyu
Department of Computer Science and Engineering,
Obafemi Awolowo University, Ile Ife, Nigeria

Owoeye Samuel
Department of Mechatronics Engineering,
Federal University of Agriculture, Abeokuta, Nigeria

Makinde Kayode
Department of Mechatronics Engineering,
Federal University of Agriculture, Abeokuta, Nigeria

Folarami Olaniyi
Department of Mechatronics Engineering,
Federal University of Agriculture, Abeokuta, Nigeria

Kamil- Bello Furqan
Department of Mechatronics Engineering,
Federal University of Agriculture, Abeokuta, Nigeria

Daodu Sakira
Department of Mechatronics Engineering,
Federal University of Agriculture, Abeokuta, Nigeria

CC BY

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20241206111011

Original scientific paper
Received: 2024-12-06, Accepted: 2025-09-30

Framework for Improving Parallelization of Loops through Loop-Fission Transformation for Many-Core Systems

Akash Yadav*, Mushtaq Ahmed

Abstract: Many-core systems rely on parallel processing to achieve high-performance computing, but writing efficient parallel code remains challenging, especially for researchers with limited programming expertise. Automatic parallelization tools such as autoPar, Cetus, and Par4All help transform sequential programs into parallel ones, yet they often fail to handle certain programming constructs efficiently. This paper proposes a loop-fission transformation-based framework to improve the parallelization of dependency-free loops with control flow statements. The framework restructures loops to maximize parallel execution, overcoming limitations in existing tools. Experiments include computational statements like array assignments, addition, and multiplication to evaluate performance. Results show that the proposed approach reduces execution time by 4% to 15% on many-core systems compared to existing tools. These findings highlight the inefficiencies of current automatic parallelization techniques and demonstrate the effectiveness of the proposed method in enhancing execution performance.

Keywords: Loop Parallelization; Loop-Fission; OpenMP; Parallelizing Compilers; Superword-Level Parallelization

1 INTRODUCTION

Traditionally, CPU processing power was increased by increasing the CPU's clock speed. The CPU consumes power and generates more heat to operate at a higher clock speed. The power that can be supplied to the CPUs is limited. This limitation of power supply is known as the Power Wall problem [1]. This problem downgrades the overall system performance. However, there are more and more data-intensive and computationally intensive applications that demand high-performance processor design, e.g., big data, machine learning, and deep learning techniques and applications [2-5]. To overcome this problem, manufacturers increase the number of cores. Today, multi and many-core architectures are developing on a larger scale as parallel computing and High-Performance Computing (HPC) are now the necessary key technologies for the latest industry standards [6].

Multi-core and many-core processors are distinct types of computer processors based on the number of cores they contain. Multi-core processors feature a limited number of independent processing units within a single package, which is suitable for parallel processing in general-purpose computing tasks. On the other hand, many-core processors have a significantly larger number of cores, exceeding dozens or even hundreds, and are optimized for highly parallel workloads like scientific simulations and data analytics [7]. While multi-core processors are commonly found in desktops and laptops, many-core processors are often used in specialized applications, HPC environments, and data centers.

Both multi-core and many-core systems support parallel computing, with multi-core processors enabling low to moderate levels of parallelism for general-purpose tasks, while many-core processors excel in handling highly parallel workloads in specialized applications [8]. The many-core processors maximize parallelism and performance in specialized applications. These many-core machines can run parallel applications to give better performance. This led to the rise of parallel or concurrent programming to give optimum performance [9, 10]. Also, programmers and developers need to create applications or programs that can fully utilize the processing power of these parallel processors. However, creating a parallel program is not a simple task. It requires in-depth knowledge of parallel programming, its underlying concepts, and parallel paradigms [11]. Depending on the machine architecture, parallelism can be achieved in many ways. A parallel application can use any available method (such as message passing, shared memory, etc.) to run in parallel [12].

Most programmers find it a tedious task to write parallel codes, which leads to the requirement of parallelization tools [13, 14]. A parallelization tool can automatically convert a sequentially written program into a parallelized one. Today, many parallelization tools are available to perform automatic parallelization [15-17]. These tools insert OpenMP directives into the sequential code to achieve parallelism [18, 19]. However, this automatic conversion is still not very efficient, particularly when a program contains loops with conditional statements inside their body, such as if-else constructs [20, 21].

Many researchers and developers rely on automatic parallelization tools to simplify this process, but these tools often fail to achieve optimal performance in such cases. This limitation motivates the need for improved techniques that can effectively handle control flow within loops. This paper explores a loop-fission transformation-based approach to enhance parallelism, making automatic parallelization more effective for complex loop structures.

The rest of the paper is structured as follows. Section 2 provides information about automatic parallelization tools, loop transformation techniques, and background information on OpenMP API. The proposed framework is discussed in Section 3. Section 4 presents the experimental setup and test cases used in the experiments and results evaluation of the

proposed framework on these test cases, followed by a conclusion.

2 EXISTING TECHNIQUES FOR AUTOMATIC PARALLELIZATION

In parallel programming, many types of parallelization exist, such as Instruction-Level Parallelization (ILP), Loop-Level Parallelization (LLP), Superword-Level Parallelization (SLP) and many others [22, 23]. In SLP, similar types of independent isomorphic statements within a basic block are grouped into a single Single Instruction Multi Data (SIMD) operation. Instead of executing those isomorphic statements one at a time, this SIMD operation is executed, allowing them to run simultaneously [23]. Fig. 1 shows an example of SLP.

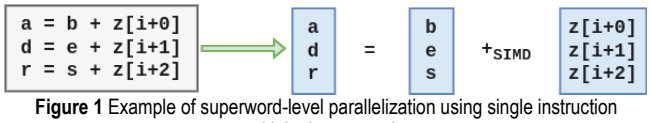


Figure 1 Example of superword-level parallelization using single instruction multiple data operation

Most of the programs spend almost 90% of their execution times in the loops [24]. Hence, LLP is a very common technique to parallelize a sequential code to speed-up the execution time [25]. Many loop transformation methods (such as loop unrolling, loop splitting, loop fission, etc.) are available that can be used to parallelize the loops. Loop unrolling generally leads to superword-level parallelism, as shown in Fig. 2.

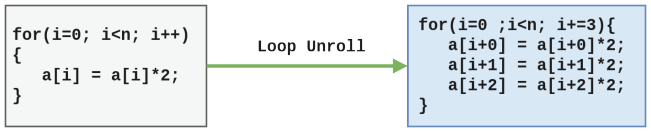


Figure 2 Example of loop unrolling transformation

SLP cannot be done when a loop contains the conditional statement in its body, as shown in Listing 1.

Listing 1 Example of a loop with conditional statements in its body

```
for(i=0;i<256;i++)
{
  if(a[i]%2==0)
  {
    a[i]=a[i]*2;
  }
}
```

In this case, loop unrolling does not create independent isomorphic statements that can be efficiently used for SIMD execution. In addition, automatic parallelization tools apply the OpenMP directive to the outer loop, which does not fully utilize parallelism. This type of loop is most common in multimedia-related areas such as image processing and video coding, such as MPEG.

2.1 Shared Memory Model and OpenMP

The shared memory model is a common programming model for parallel computing in which multiple processes or threads share access to a common memory pool, as shown in Fig. 3. This allows them to share data and communicate with each other more efficiently than other methods of interprocess communication, such as message passing.

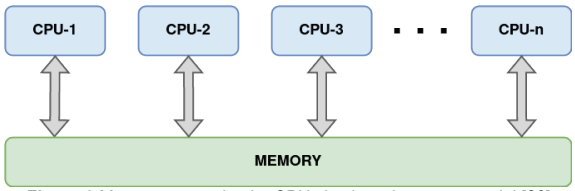


Figure 3 Memory accessing by CPUs in shared memory model [26]

One of the primary advantages of the shared memory model is its speed and efficiency. The same memory location can be accessed directly by all processes or threads, eliminating the requirement for message passing, which copies data between them [27]. This property can significantly decrease overhead and enhance performance, particularly for highly computational tasks that require a significant amount of communication between different processes. Shared memory systems can be utilised to enhance the performance of parallel computing applications, especially those involving large data sets [28, 29].

However, there are also some challenges that need to be considered when using the shared memory model. One of the most important issues is the requirement for meticulous synchronisation in order to prevent race situations and other concurrency problems. Furthermore, the approach may also be less scalable than alternative parallel programming models, such as message-passing models, for specific high-performance computing applications, especially when the number of processes or threads increases [30].

OpenMP is one of the prominent parallel programming standards for FORTRAN, C, and C++, enabling programmers to create parallel code using environment variables, runtime library functions, and compiler directives. The main objective of OpenMP is to simplify the writing of high-performance parallel programs that can utilise multiple CPU cores in a single system. OpenMP does have some restrictions, though. Its inability to enable distributed memory parallelism is one of its main drawbacks, which may limit its applicability to some complex HPC applications.

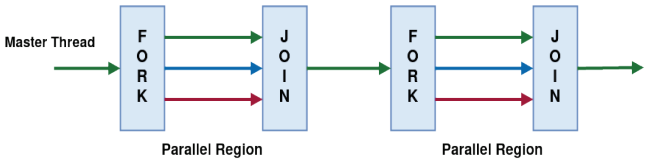


Figure 4 Parallelization by OpenMP using fork-join [31]

OpenMP uses the fork-join model to provide multi-threading-based parallelism. The fork-join model diagram is shown in Fig. 4. When the master thread encounters the parallel region, it executes the fork operation and generates

slave threads and then both master and slave threads start executing code in parallel. At the end of the parallel region, the master thread waits to complete the task assigned to the slave threads, and after completion, it executes the join operation.

Fig. 5 shows an example of parallelizing a loop using four threads. The master thread (in green colour) executes the program sequentially. When a parallel region occurs (*#pragma omp parallel for*), master threads create three slave threads and distribute the loop iterations between them (including itself) to run the loop in parallel. The loop iterations are divided equally between the master thread and slave threads. Each thread executes four iterations of loops. After completion of the task, master threads use the join operation to end the parallel execution and continue the execution of the sequential region.

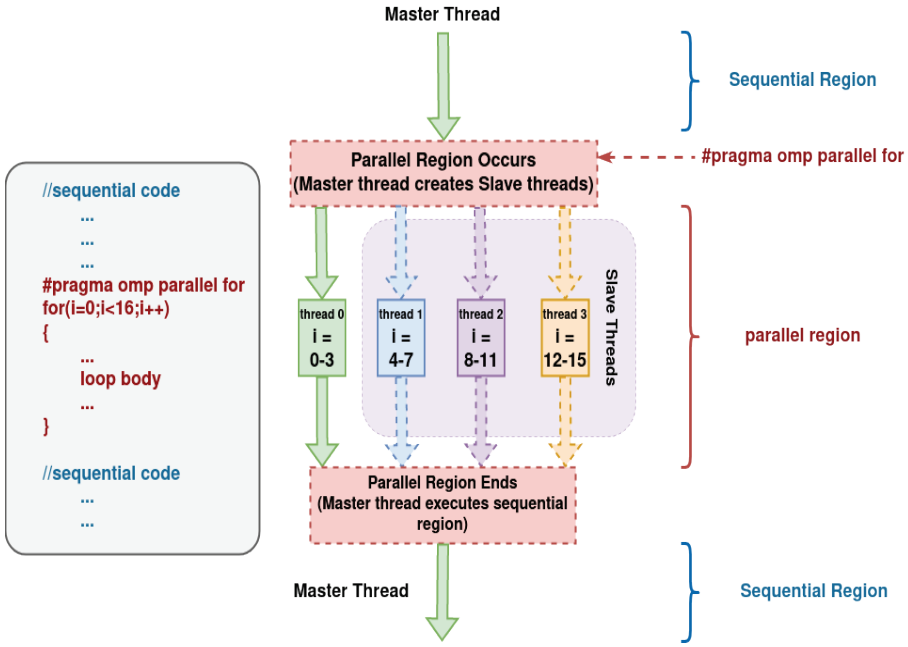


Figure 5 Parallelizing a for loop with 16 iterations using OpenMP with four threads

2.2 Automatic Parallelization Tools

Several automatic parallelization compilers can convert sequential code into its parallelized equivalent by inserting OpenMP directives [32]. In our experiments, we used AutoPar and Cetus, as they are free, up-to-date, and well-suited for parallelizing scientific codes. Both are source-to-source compilers that utilize the OpenMP API for parallelization.

```
for (i = 0; i < n ; i ++)
{
    a[i] = a[i] + b[i];
}
```

(a) Simple loop

```
for (i = 0; i < n ; i ++)
{
    if (a[i] != 0)
    {
        a[i] = a[i] + b[i];
    }
}
```

(b) Loop with control flow statement

Figure 6 Example of two different loops

The following subsections provide brief background information about autoPar and Cetus. These subsections also contain examples of how autoPar and Cetus parallelize the two loops with different body statements, as shown in Fig. 6. Listings 2–5 are code snippets that highlight the OpenMP directives inserted by autoPar and Cetus. The full programs are not shown here. In the complete codes, loop variables such as int i are declared. The tools also generate the long type casting and fixed values like 100000. These are shown exactly as produced.

2.2.1 autoPar

The autoPar is an auto parallelization tool in ROSE. ROSE is a compiler-based infrastructure for building program transformation and analysis tools [33]. It is developed at Lawrence Livermore National Laboratory. It is an open-source source-to-source compiler. ROSE tools support a wide range of programming languages, including C/C++, Python, PHP, and FORTRAN. ROSE includes an open-source automatic parallelization compiler called autoPar, which can currently parallelize C and C++ programmes by inserting OpenMP pragma in the sequential code. Listings 2 and 3 show how autoPar parallelizes the loops from Fig. 6 with an iteration count of 100000. Instead of the full code, the listings highlight only the changes made by the automatic parallelization tool, specifically the addition of OpenMP directives to parallelize the loops. All variables used in the loops are already declared in the original code, and the tool itself applies the typecasting of constants, such as long.

Listing 2 Parallelization of the simple loop by autoPar

```
#pragma omp parallel for private(i)
for (i=0; i<=((long )100000)-1;i+=1){
  a[i] = a[i] + b[i];
}
```

Listing 3 Parallelization of the loop with conditional statements in its body by autoPar

```
#pragma omp parallel for private(i)
for (i=0; i<=((long )100000)-1;i+=1){
  if (a[i] != 0) {
    a[i] = a[i] + b[i];
  }
}
```

2.2.2 Cetus

Cetus is a source-to-source C compiler written in Java and was developed and maintained by the ParaMount research group at Purdue University [34]. It is considered the most powerful and easy-to-use tool. Currently, it parallelizes only C programs.

Listing 4 Parallelization of the simple loop by Cetus

```
#pragma cetus private(i)
#pragma loop name main#0
#pragma cetus parallel
#pragma omp parallel for if((10000<(1L+3L*n)))) private(i)
for(i = 0; i < n; i++)
{
a[i] = a[i] + b[i];
}
```

Cetus's analysis passes can perform two types of data dependency analysis: Banerjee-Wolfe Test and Range Test. Its analysis pass also includes array privatization, symbolic expression manipulators, reduction recognition, etc. Transformation passes of Cetus include loop parallelizer, induction variable substitution pass, and loop dependent analysis. Automatic parallelization capability is a built-in capability of the Cetus compiler.

Listing 5 Parallelization of the loop with the conditional statement in its body by Cetus

```
#pragma cetus private(i)
#pragma loop name main#0
#pragma cetus parallel
#pragma omp parallel for if((10000<(1L+(3L*n)))) private(i)
for(i=0;i<n;i++){
  if(a[i]!=0){
    a[i]=a[i]+b[i];
  }
}
```

The listings illustrate how Cetus applies OpenMP to the loops in Fig. 6 for parallelization. Listings 4 and 5 show that Cetus parallelizes both loops in the same manner.

3 PROPOSED FRAMEWORK

As discussed in Section 2.2, parallelization tools apply the same parallelization technique to transform the loops with or without the control flow statements. Most of the time, compilers apply if-conversion to optimize the loop, but it is only applicable when the condition and data associated with it remain the same in every iteration of the loops.

These types of loops can be easily parallelized. But when there is an array-based if condition where data associated with the condition changes in every iteration (e.g., if($a[i]$=0)), it becomes complex to optimally parallelize them. Fig. 7 shows the basic working of automatic parallelization compilers to parallelize loops.

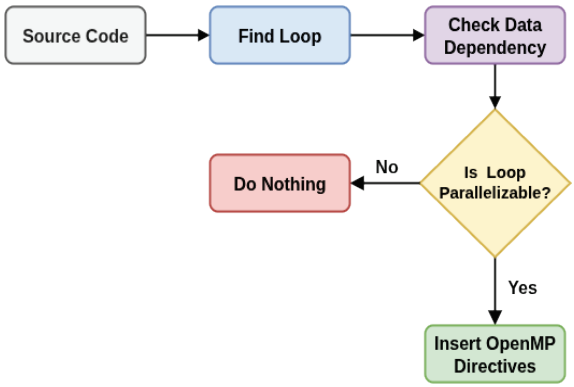


Figure 7 Basic approach of automatic parallelization compilers for parallelizing loops

The proposed framework (named as loopFiss) uses loop-fission transformation to parallelize data-dependency-free loops, even with control statements, for both multi-core and many-core systems. However, it is particularly suited for many-core systems, as HPC workloads typically run on such architectures. In loop-fission, a loop is split into separate loops, as shown in Fig. 8. Independent computations are parallelized, while statements with dependencies, such as data dependencies or control flow constraints, remain sequential, ensuring efficient execution and correctness.

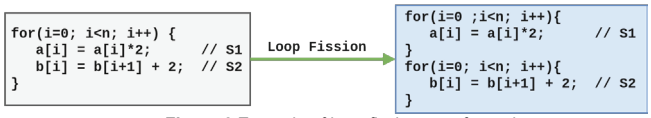


Figure 8 Example of loop-fission transformation

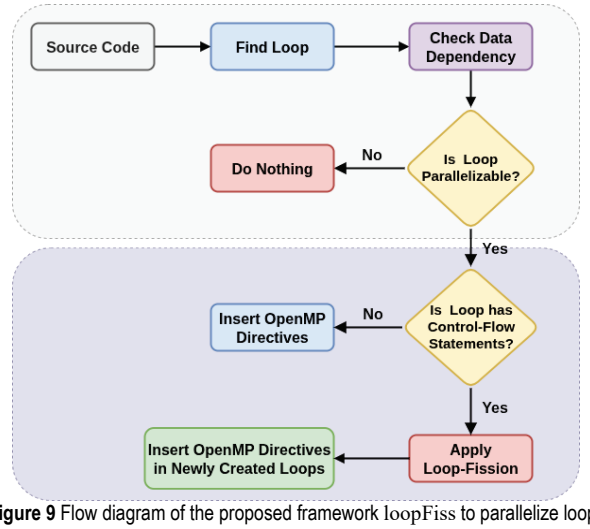


Figure 9 Flow diagram of the proposed framework loopFiss to parallelize loops

The proposed approach checks for dependency-free loops that can be parallelized. Fig. 9 shows the flow diagram of the proposed approach. If the loop has no control flow statements, then OpenMP pragma directives are inserted to parallelize the loop. If a control flow statement is present

inside the loop body, then loop-fission transformation is applied.

For example, to parallelize the loop given in Listing 6, first, it searches for the dependency-free loop with control flow statements.

Listing 6 Syntax of parallelizable loop

```
for(i=0;for_cond;i++){
  if(if_cond){
    body statements containing arrays;
  }
}
```

Then, the loop-fission is applied to the loop to break it into two loops. The first loop stores every array index for which condition is true in a temporary array *trueArr, and* then the second loop uses the index store in *trueArr* to run the statements in parallel. With loop-unrolling transformation, the second loop can use superword-level parallelism to speed-up the processing time.

Algorithm 1 Steps of Proposed Framework loopFiss

```
1:  Create an array trueArr of size N;
2:  Initialize m←0;
3:  while for_cond is true:
4:      if if_cond is true:
5:          trueArr[m]←i;
6:          m←m+1;
7:      end if
8:  end while
9:  Set k ← 0:
10: Apply OpenMP-API to parallelize the loop in Step 11
11: while k< m :
12:     i ← trueArr[k];
13:     Use the value of i as an index for the arrays in body statements;
14: end while
```

Algorithm 1 shows the overall steps of parallelizing the loop in the proposed framework. Listing 7 shows an example of a parallelized loop using this framework. The proposed loop-fission method divides the original loop into two loops. The first loop finds the elements that satisfy the condition. The second loop performs the computations on these elements in parallel. This separation makes it possible to parallelize loops that contain conditional statements.

Listing 7 Code snippet of the output program after applying the loopFiss approach

```
int trueArr[n];
int i,m=0;
#pragma omp parallel for simd private(i)
for(i=0;i<n;i++){
  if(a[i]!=0){
    trueArr[m++]=i;
  }
}

#pragma omp parallel for simd private(z)
for(int z=0;z<m;z++){
  i=trueArr[z];
  a[i]=a[i]+b[i];
}
```

4 EXPERIMENTAL EVALUATION

In this section, the experimental setup and test cases used in the evaluation are described, followed by a discussion of the experimental results. The experiment is conducted on a 72-core Intel® Xeon® processor with 64GB of RAM. For the C-compiler, GCC 9.4.0 is used to compile all test cases as it supports OpenMP 4.5 specifications. The detailed specifications of the hardware and software used for the experiments are shown in Tab. 1.

Table 1 Details of hardware/software specifications used for experiments

Tools	Specification
Ubuntu OS	Ubuntu 20.04.2 (64-bit)
Processor	Intel® Xeon® CPU E5-2699 v3
Clock speed	2.30GHz
Total Cores	72
RAM	64 GB
Cetus	Version 1.4.4
autoPar	ROSE-0.11.46.0
Compiler	GCC 9.4.0

A Python-based generator program is used to generate test cases that contain array-based statements in a loop body. The code snippet shown in Listing 6 is used as the template for test cases in the experiments.

Table 2 Test cases with their description and example of statements used as the loop body

Test Case	Example	Description
AASSIG	$a[i] = a[i] + int\_value$	Assign a value to array elements
AADD	$c[i] = a[i] + b[i]$	Addition of two arrays
AMUL	$c[i] = a[i] * b[i]$	Vector Multiplication of two arrays
AOPs:	$b[i] = a[i] + int\_value$ $c[i] = a[i] * b[i]$	Multiple Computation with different arrays

Different statements such as array value assignment, array addition, array multiplication and their combinations are used as the body of the control statement (if-statement). Tab. 2 shows the different test cases with descriptions.

The experiments are performed to compare the execution time efficiency of the proposed approach with existing approaches used by autoPar and Cetus using different sizes of arrays for every test case. The arrays of varying sizes from 10K to 200k are used in experiments, and the number of loop iterations in each test case is equal to the size of the array. The test cases are run ten times, and their average value is used as the result value to minimize the error. In addition, during the execution of the test case, no other tasks are assigned to the system to keep the system resources free.

4.1 Result Analysis

Tab. 3 shows the execution times of experiments for different array sizes for test cases *AASSIG* and *AADD*. In *AASSIG*, the assignment operation ($a[i] = a[i] + int\_value$) involves arithmetic along with array access. This makes it more computationally intensive than a simple array assignment. In *AADD*, the operation represents element-wise addition of two arrays.

Table 3 Execution time comparison between Cetus, autoPar and loopFiss approach for test cases *AASSIGN* and *AADD*

No. of Iterations	Execution Time (in ms.)					
	AASIGN			AADD		
	Cetus	autoPar	loopFiss	Cetus	autoPar	loopFiss
10000	1.195	1.186	1.201	1.341	1.331	1.346
30000	1.389	1.383	1.394	1.571	1.563	1.571
50000	1.487	1.484	1.468	1.677	1.671	1.650
80000	1.697	1.692	1.663	1.938	1.934	1.892
100000	1.814	1.810	1.769	2.281	2.274	2.214
130000	2.090	2.087	2.032	2.655	2.650	2.565
150000	2.396	2.392	2.321	3.103	3.090	2.982
180000	2.782	2.775	2.679	3.481	3.473	3.331
200000	3.023	3.014	2.901	3.830	3.818	3.639

Results show that for the small number of inputs, the proposed approach takes more time compared to the existing methods. However, as the input size increases, it performs better. From the results, it can be observed that the operation type also affects the execution time. Compared to simple operations like array assignment and addition, a complex operation like multiplication or multiple operations in a single statement take more time.

The proposed approach shows significant improvement in execution time for complex operations, as shown in Fig. 10 and Fig. 11, which show the execution time for test cases *AMUL* and *AOPs*, respectively.

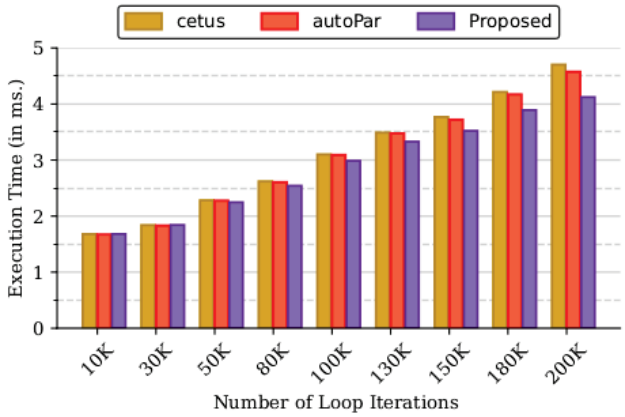


Figure 10 Execution time comparison between Cetus, autoPar and loopFiss for test cases AMUL (1K = 1000)

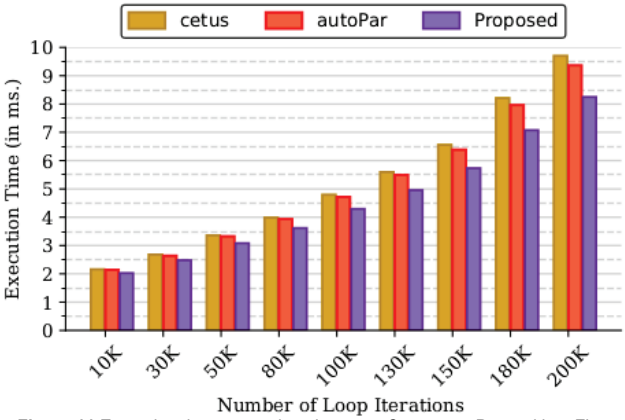


Figure 11 Execution time comparison between Cetus, autoPar and loopFiss Approach for test cases AOPs (1K = 1000)

The speed-up metric is used to determine the performance of methods. The speed-up improvement percentage between the proposed approach and the existing ones is calculated using Eq. (1).

$$Speed\ Up\ (\%) = \left(1 - \frac{ET_{\text{proposed}}}{ET_{\text{existing}}}\right) \cdot 100 \tag{1}$$

Where: ET_{proposed} is the execution time of the proposed approach, and ET_{existing} is the execution time of the existing approach.

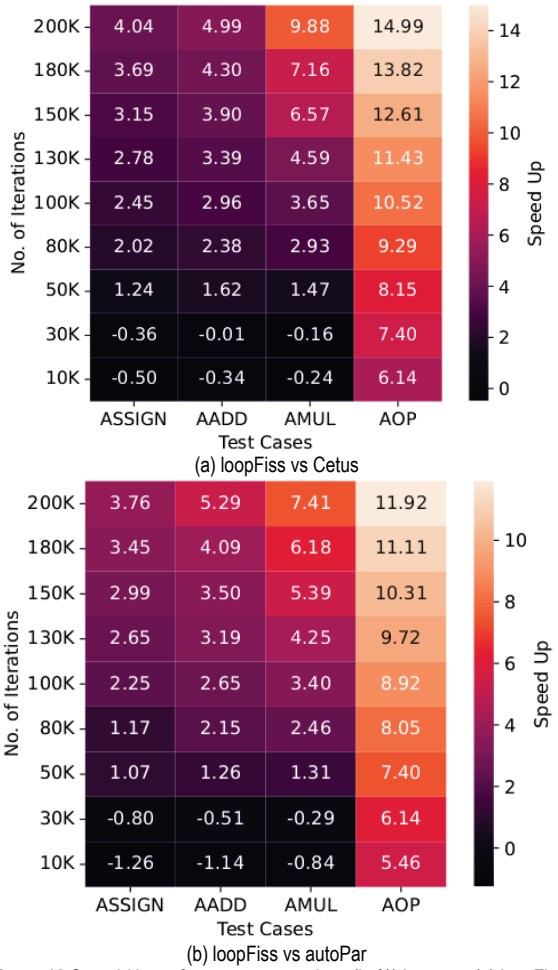


Figure 12 Speed-Up performance comparison (in %) between (a) loopFiss approach vs Cetus, and (b) loopFiss approach vs autoPar

Fig. 12 shows the speed-up differences between the proposed approach and existing approaches for all the test cases. From the results, it can be observed that the proposed approach shows low performance (−ve values) when the input size is small compared to large input sizes, but for large-size inputs, it outperforms the existing methods and shows a significant improvement in execution time.

When the size of the array is small, the proposed approach takes more time compared to existing methods because it breaks the loops and then parallelizes them by creating threads, whereas the existing methods parallelize

only one loop. When performing parallelization, the overhead associated with dividing the work among threads and synchronizing their operations has a significant impact on the overall execution time. For small input sizes, this overhead may outweigh the benefits of parallel execution, resulting in longer execution times when parallelizing two loops compared to parallelization of a single loop.

Fig. 12(a) shows that compared to Cetus, the proposed approach gives a speed-up of 4% to 15% for simple to complex operations. Similarly, compared to autoPar, it shows a speed-up improvement of approximately 3.7% to 12% for the same, as shown in Fig. 12(b). The results analysis indicate that in performing parallelization using OpenMP, simple operations such as array addition exhibit shorter execution times compared to multiplication or more complex computations. This outcome can be attributed to the inherent computational complexity associated with the operation. Additionally, the overhead of parallelization, encompassing thread creation, synchronization, and data communication, has a greater impact on complex operations compared to the simple ones.

5 CONCLUSION AND FUTURE WORK

This work is motivated by the fact that writing a parallel program is a difficult task, specifically for researchers who don't have enough programming knowledge. Therefore, automatic parallelization tools are developed to convert sequential programs into parallel versions to perform computations faster. However, there are still many programming constructs that cannot be efficiently parallelized by these tools. This paper introduced a framework to effectively parallelize data-dependency-free loops with control flow statements using the loop-fission method.

In experiments, various statements like array assignment, addition, multiplication, etc., are used as the body of control flow statements to compare the approaches using execution time as the performance metric. Experimental results show that our approach improves the execution time by 4% to 15% on many-core systems compared to existing tools. These improvements are modest, reflecting the specific test cases and array sizes used in our experiments. To better assess performance over a wider range, future studies will consider larger array sizes that grow more rapidly. Additionally, future work will focus on modifying the approach so that it can be applied to loops containing more than one conditional statement, thereby extending its applicability to more complex scenarios.

6 REFERENCES

[1] Bhaskar, A. V. (2022). A new method of power analysis of network-on-chip using analytical modelling. In *IEEE Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC2022)*, 222–227. https://doi.org/10.1109/PDGC56933.2022.10053136

[2] Lai, Z., Li, S., Tang, X., Ge, K., Liu, W., Duan, Y., Qiao, L., & Li, D. (2023). Merak: An efficient distributed DNN training framework with automated 3D parallelism for giant foundation models. *IEEE Transactions on Parallel and Distributed Systems, 34*(5), 1466–1478. https://doi.org/10.1109/TPDS.2023.3247001

[3] Dozier, J. (2022). Revisiting topographic horizons in the era of big data and parallel computing. *IEEE Geoscience and Remote Sensing Letters, 19*, 1–5. https://doi.org/10.1109/LGRS.2021.3125278

[4] Mohseni, M., Othman, B. A., Raturi, P., Mishra, A. B., Priya, S., & Saravanan, V. (2022). The role of parallel computing towards implementation of enhanced and effective industrial internet of things (IoT) through MANOVA approach. In *IEEE 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE2022)*, 160–164. https://doi.org/10.1109/ICACITE53722.2022.9823698

[5] Tripathi, A. K., Sharma, K., Bala, M., Kumar, A., Menon, V. G., & Bashir, A. K. (2021). A parallel military-dog-based algorithm for clustering big data in cognitive industrial internet of things. *IEEE Transactions on Industrial Informatics, 17*(3), 2134–2142. https://doi.org/10.1109/TII.2020.2995680

[6] Hoe, M., & Dargham, J. (2020). High performance computing (HPC) applications in industry 4.0 (I4.0) for the betterment of humanity. In *IEEE 8th R10 Humanitarian Technology Conference (R10-HTC2020)*, 1–6. https://doi.org/10.1109/R10-HTC49770.2020.9356990

[7] Groen, M., & Gramoli, V. (2016). Multicore vs manycore: The energy cost of concurrency. In *Euro-Par 2016: Parallel Processing* (Dutot, P.-F., & Trystram, D., Eds.), 545–557. Springer International Publishing. https://doi.org/10.1007/978-3-319-43659-3_40

[8] Geer, D. (2005). Chip makers turn to multicore processors. *Computer, 38*(5), 11–13. https://doi.org/10.1109/MC.2005.160

[9] Rogowski, M., Aseeri, S., Keyes, D., & Dalcin, L. (2023). mpi4py.futures: MPI-based asynchronous task execution for Python. *IEEE Transactions on Parallel and Distributed Systems, 34*(2), 611–622. https://doi.org/10.1109/TPDS.2022.3225481

[10] Wang, X., Yang, R., & Liu, X. (2022). Research on general system of remote sensing satellite data preprocessing based on MPI + CUDA. In *IEEE 3rd International Conference on Geology, Mapping and Remote Sensing (ICGMRS2022)*, 607–612. https://doi.org/10.1109/ICGMRS55602.2022.9849313

[11] Watkinson, N., Shivam, A., Nicolau, A., & Veidenbaum, A. (2019). Teaching parallel computing and dependence analysis with Python. In *IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW2019)*, 320–325. https://doi.org/10.1109/IPDPSW.2019.00061

[12] Mitrovski, J., Djinevski, L., Gusev, M., & Arsenovski, S. (2021). TPC-H benchmark Q3, Q6 and Q12 sequential, OpenMP parallel, and CUDA parallel implementation. In *IEEE 44th International Convention on Information, Communication and Electronic Technology (MIPRO2021)*, 938–943. https://doi.org/10.23919/MIPRO52101.2021.9597197

[13] Arabnejad, H., Bispo, J., Barbosa, J. G., & Cardoso, J. M. P. (2018). An OpenMP-based parallelization compiler for C applications. In *2018 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Ubiquitous Computing & Communications, Big Data & Cloud Computing, Social Computing & Networking, Sustainable Computing & Communications (ISPA/IUCC/BDCloud/SocialCom/Sustain Com)*, 915–923. https://doi.org/10.1109/BDCloud.2018.00135

[14] Sah, S., & Vaidya, V. G. (2012). A review of parallelization tools and introduction to EasyPar. *International Journal of Computer Applications, 56*(12). https://doi.org/10.5120/8944-3108

[15] Palkowski, M., & Bielecki, W. (2017). TRACO: Source-to-Source Parallelizing Compiler. *Computing and Informatics, 35*(6), 1277–1306.

[16] Dever, M. (2015). *AutoPar: Automating the parallelization of functional programs* (PhD thesis). Dublin City University.

[17] Dave, C., Bae, H., Min, S.-J., Lee, S., Eigenmann, R., & Midkiff, S. (2009). Cetus: A source-to-source compiler infrastructure for multicores. *Computer, 42*(12), 36–42. https://doi.org/10.1109/MC.2009.385

[18] Arabnejad, H., Bispo, J. A., Cardoso, J. A. M. P., & Barbosa, J. G. (2020). Source-to-source compilation targeting OpenMP-based automatic parallelization of C applications. *Journal of Supercomputing, 76*, 6753–6785. https://doi.org/10.1007/s11227-019-03109-9

[19] Dagum, L., & Menon, R. (1998). OpenMP: An industry standard API for shared-memory programming. *IEEE Computational Science and Engineering, 5*(1), 46–55. https://doi.org/10.1109/99.660313

[20] Prema, S., Jehadeesan, R., & Panigrahi, B. (2017). Identifying pitfalls in automatic parallelization of NAS parallel benchmarks. In *IEEE National Conference on Parallel Computing Technologies (PARCOMPTECH2017)*, 1–6. https://doi.org/10.1109/PARCOMPTECH.2017.8068329

[21] Liang, X., Humos, A. A., & Pei, T. (2017). Vectorization and parallelization of loops in C/C++ code. In *Proceedings of the International Conference on Frontiers in Education: Computer Science and Computer Engineering (FECS)* (pp. 203–206). The Steering Committee of the World Congress in Computer Science.

[22] Singh, D., & Garg, R. (2023). Performance analysis of a multicore approach proposed for efficient community detection and recommendation system. *IETE Technical Review*, 0(0), 1–15.

[23] Zhou, H., & Xue, J. (2016). A compiler approach for exploiting partial SIMD parallelism. *ACM Transactions on Architecture and Code Optimization (TACO), 13*(1), 1–26. https://doi.org/10.1145/2886101

[24] Lee, J. et al. (2014). Improving performance of loops on DIAM-based VLIW architectures. SIGPLAN Not. 49, 5 (May 2014), 135–144. https://doi.org/10.1145/2666357.2597825

[25] Devan, P. S., & Kamat, R. (2014). A Review-LOOP Dependence Analysis for Parallelizing Compiler. *International Journal of Computer Science and Information Technologies, 5*(3), 4038–4046.

[26] Barney, B. (n.d.). LLNL HPC Tutorials: OpenMP Programming Model. Retrieved from https://hpc-tutorials.llnl.gov/openmp/programming_model/

[27] Naik, D., Ramesh, D., Gandomi, A. H., & Gorojanam, N. B. (2022). Parallel and distributed paradigms for community detection in social networks: A methodological review. *Expert Systems with Applications, 187*, 115956. https://doi.org/10.1016/j.eswa.2021.115956

[28] Rauber, T., & Rünger, G. (2023). Parallel Programming Models (pp. 109–168). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-031-28924-8

[29] Eichstädt, J., Vymazal, M., Moxey, D., & Peiró, J. (2020). A comparison of the shared-memory parallel programming models OpenMP, OpenACC, and Kokkos in the context of implicit solvers for high-order FEM. *Computer Physics Communications, 255*, 107245. https://doi.org/10.1016/j.cpc.2020.107245

[30] He, G., Vialle, S., & Baboulin, M. (2022). Parallel and accurate K-means algorithm on CPU-GPU architectures for spectral clustering. *Concurrency and Computation: Practice and Experience, 34*(14), e6621. https://doi.org/10.1002/cpe.6621

[31] Liu, X., Lv, H., & Liu, S. (2021). Design and implementation of multi-core parallel compiler based on OpenMP. *Journal of Physics: Conference Series, 1802*, 032137. https://doi.org/10.1088/1742-6596/1802/3/032137

[32] Harel, R., Mosseri, I., Levin, H., Alon, L.-o., Rusanovsky, M., & Oren, G. (2020). Source-to-source parallelization compilers for scientific shared-memory multi-core and accelerated multiprocessing: Analysis, pitfalls, enhancement, and potential. *International Journal of Parallel Programming, 48*(1), 1–31. https://doi.org/10.1007/s10766-019-00640-3

[33] Quinlan, D. (2000). Rose: Compiler support for object-oriented frameworks. *Parallel Processing Letters, 10*(02n03), 215–226. https://doi.org/10.1142/S0129626400000214

[34] Lee, S.-I., Johnson, T. A., & Eigenmann, R. (2003). Cetus: An extensible compiler infrastructure for source-to-source transformation. In *International Workshop on Languages and Compilers for Parallel Computing* (pp. 539–553). Springer. https://doi.org/10.1007/978-3-540-24644-2_35

Authors' contacts:

Akash Yadav, Dr.
(Corresponding author)
GLA University,
Mathura, UP- 281406, India
Malaviya National Institue of Technology,
Jaipur, Rajasthan-302017, India
akash.yadav@gla.ac.in, 2019rcp9150@mnit.ac.in

Mushtaq Ahmed, Dr.
Malaviya National Institue of Technology,
Jaipur, Rajasthan-302017, India
mahmed.cse@mnit.ac.in

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250116094958

Original scientific paper
Received: 2025-01-16, Accepted: 2025-08-26

Glaucoma Stages Detection Using Feature Selection Techniques

Dimah Alkasem*, Hazem Issa

Abstract: Glaucoma, also known as the "silent thief of sight", is a chronic eye disease caused by increased intraocular pressure, leading to irreversible vision loss without the affected individual ever being aware of it. Over time, it results in damage to the optic nerve, which carries visual images from the eye to the brain. It slowly destroys the eyes of the patients unsuspectingly. There are no early symptoms of glaucoma. Therefore, the challenge lies in detecting the early stages of the disease. This paper puts forth a methodology of feature extraction and relevant feature selection. Feature selection is a much important Stage because it cuts the time required for training time and the features extracted by selecting only worthy and important features while neglecting all unnecessary or non-important features that helps in classifier boosting accuracy. ANN were the methods used in classifying. The result obtained from the proposed method was evaluated with accuracy, specificity and sensitivity. The experimental results of the proposed system achieved an accuracy of 96.12 % with the use of the Harvard dataset and 91.67 % using the Rim-One1 dataset. The experimental proved that the proposed model outperform advanced methods in diagnosing early stages of glaucoma.

Keywords: ANN (Artificial Neural Network); CatBoost; feature selection; glaucoma; HOG (Histogram of Oriented Gradients)

1 INTRODUCTION

Glaucoma is a progressive neurodegenerative disease that causes permanent vision impairment and ranks as the second leading cause of blindness with an estimated prevalence of 80 million individuals in 2020 and is expected to increase to 111 million in the year 2040 [1]. Glaucoma can also go unnoticed until its later stages because of its silent nature; thus, its timely diagnosis becomes very difficult and essential to hinder irreversible vision loss [2]. Conventional modes of diagnosis like the measurement of intraocular pressure (IOP), cup- to disc ratio (CDR), retinal nerve fibre layer (RNFL) and the detection of visual field (VF) abnormalities by optical coherence tomography (OCT) are commonly used [3].Nevertheless, these tests have such drawbacks as the inter-observer variability, extreme high cost, as well as limited accessibility in low-resource environments, which substantiates the necessity to develop automated, reliable, and scalable solutions of diagnostic methods .

Lately, AI and, in particular, ML demonstrated its enormous possibilities in analysing medical images and enables defining the presence of ocular diseases more precisely and within a small amount of time. Research articles have tried out the method utilising support vector machines (SVMs), random forests, and convolutional neural networks (CNNs). SVMs and random forests have shown good performance in the handcrafted features cases but also tend to perform poorly in high-dimensional data and generic transfer across datasets. Conversely, the CNN based deep learning methods have demonstrated high performance with automatic discrimination of features, but they usually demand huge annotated datasets and enormous computing resources which may not always be practical in practice in medical context. These shortcomings demonstrate the necessity of finding an alternative, which would strike a balance between the interpretability and efficiency of hand engineered features, and the robustness of contemporary machine learning classifiers--a solution that this research will attempt to find.

This paper seeks to build and test an automatic glaucoma detection system that combines Histogram of Oriented Gradients (HOG) to extract features and CatBoost to select features and an Artificial Neural Network (ANN) to classify. The scientific value of the present work resides in: (i) shows how to efficiently and effectively integrate handcrafted features with modern ML-based feature selection to diminish dimensionalities with preservation of diagnosis power, (ii) verifying the suggested model on two publicly available datasets (Harvard and RIM-ONE) to guarantee generalizability, (iii) showing superior results in terms of accuracy at various stages of glaucoma as compared to numerous approaches, previously described in the literature.

This research will help in guiding clinicians and researchers in developing credible, accessible and scalable methods of early-stage diagnosis or screening of potential vision loss subjects, which lessen the global burden of vision loss [4].

The article was organized as follows: Section 2 will describe the papers related to the work; Section 3 will explain the methodology used for glaucoma detection; Section 4 will describe the databases used and discuss the evaluation criteria used; Section 5 is for discussion; and Section 6 is the conclusion.

2 RELATED WORKS

In recent years, a large number of studies have been devoted to the application of machine learning methods to automatic glaucoma diagnosis. The researchers develop a group of machine learning algorithms that are intended for the development of an automated glaucoma detection system. Today, machine learning has supplanted other tools when it comes to the prediction and treatment of this dangerous illness. Hatanaka et al. [5] proposed a method for the calculation of the CDR. Through the assessment of the second derivative of the vertical profile curve with respect to

the area under the curve, it was evaluated if the patient was affected by glaucoma. The researchers Parashar and Agrawal have presented a novel technique for the automatic classification of glaucoma stages on a 2D wavelet transform (2D-T-EWT) [6]. More than half of the approaches related to the automated recognition of glaucoma problems seized on these existing elements (CDR and RDR) [7, 8] when they worked individually. Classification of glaucoma, which depends on the optical boundaries, can sometimes be misleading based on the fact that it might generate falsely incorrect glaucoma classifications. To defeat these shortcomings, scientists started using pipelines in multiple stages, including the preprocessing, feature extraction, and classification. To give an example, Sumaya et al. [9] presented a multi-stage and automatic model of diagnosis of glaucoma, and Alice et al. [10] accepted image feature descriptors as a contribution to the mixture of machine learning classifiers, resulting in an accuracy of 80.43%. The same case applies to Latif et al. [11], who applied a filter to remove noise, followed by feature extraction. The important features were then selected using the Grey Wolf Optimizer, and classification was performed using several classifiers. The results showed that the feature selection process combined with the SVM classifier achieved the highest accuracy. Singh et al. [12] segmented the optic disc using preprocessing, after which the HOG features were extracted, and found ANN to have the highest performance as compared to several other classifiers. Similarly, Rahouma et al. [14] used a similar ANN based on statistical texture features extracted with both GLCM and GLRLM, obtaining a very impressive accuracy rate of 99 percent on early glaucoma detection. This goes to show the potential of ANN on glaucoma detection.

Recent research still finds that an important value lies in the combination of hand-crafted feature extraction and ensemble classifiers. The modified Pelican Optimization Algorithm was coupled with Extreme Learning Machine (ELM) by Muduli et al. [15] and combined with curvelet-based feature extraction and dimensionality reduction (FDCT-WRP + PCA/LDA). The method resulted in high accuracy (93.25% and 96.75% on G1020 and ORIGA, respectively) and was able to use explainable AI methods as a means of model explanation. Parallel to it, Zhang et al. [16] have developed an efficient feature selection-based clinical decision support system that, besides increasing the classification accuracy, offered lower training costs, which reveals the relevance of the dimensionality reduction in glaucoma detection. On the whole, there is a definite evolution evident in these studies: the initial dependence on the structural parameters (CDR, RDR), then usage of handcrafted features (HOG, texture descriptors), and, finally, ANN-based specialized methods with feature selection and optimization algorithms to produce a robust and efficient glaucoma detection.

3 PROPOSED METHODOLOGY

A number of important steps will be included in the suggested strategy. The principal five steps of the proposed method are depicted in Fig. 1.

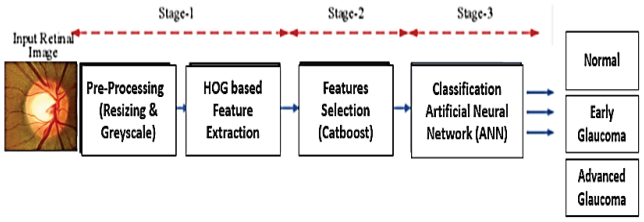


Figure 1 The Main Steps of my Proposed Process

3.1 Input Image

Two types of retinal images have been collected from different datasets. Each of the input variables is in RGB colour. Herein, data collection and pre-processing steps may be singled out in Section 3.

3.2 Image Pre-Processing

The preprocessing step acts as a key stage in image processing and involves two components. Color images are resized to 128 × 128 to match the image's pixel resolution and the resized images are then converted from RGB to grayscale.

3.3 Features Extraction

It is a technique that aims to find important data from necessary data. These features play an important part in distinguishing different classes. However, this method selects the most important features or things that make one sample distinct from others.

3.3.1 The Process of Computing the Histogram of Oriented Gradient (HOG) Feature

In the suggested system, the HOG has been used for the feature extraction from the image. The histogram of oriented gradients is technique is counting the number of times that the sign of the gradient is changing in such localized regions of the picture. The picture will consist of the small groups of cells which are connected and, for the pixels of each cell the histogram of the gradient directions is made. After that the feature vector is eventually the concatenation of these histograms. Fig. 2 demonstrates the process of computing the Histogram of Oriented Gradient (HOG) Feature [17].

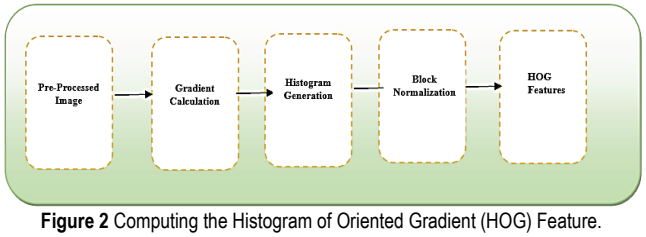


Figure 2 Computing the Histogram of Oriented Gradient (HOG) Feature.

1) Gradient calculation: The gradient is formed out of the picture magnitude and angle. $G(x)$ and $G(y)$ are estimated at each of the pixels based on a block of dimension 3 × 3. The calculation is made using the Eq. (1) and is given below per pixel value [12].

$$G_x(y,x) = H(y,x+1) - H(y,x-1)$$
$$G_y(y,x) = H(y-1,x) - H(y+1,x) \quad (1)$$

Where $G_x(y, x)$ represents the horizontal gradient of pixel (y, x) and $G_y(y, x)$ represents the vertical gradient of pixel (y, x), H represents the image, y represents rows and x represents coloums. After calculating G_x and G_y, magnitude and angle of each pixel is calculated using the Eq. (2) mentioned below [12].

$$Magnitude(\mu) = \sqrt{G_x^2 + G_y^2}$$
$$Angle(\theta) = \left| \tan^{-1}\left(\frac{G_y}{G_x}\right) \right| \quad (2)$$

2) Histogram generation: For every cell, a histogram is formed using that magnitude and direction information previously acquired.The vector gradients are then divided into groups and all the necessary groups are obtained, as well as the cells selected from the cluster are precisely collected so that the vectors are evenly produced.

3) Block normalization: In this step, the data resulting from the histograms for each block are normalized, which is predicated on the variations in the lighting and background of the image. But by normalizing the gradients, we can lessen this disparity in illumination. Therefore, after feature extraction is finished. The normalization of the feature vector, $\boldsymbol{v}$, is done through the L2-norm method which is standardized to normalize the blocks and calculated using the Eq. (3) mentioned below.

$$\boldsymbol{v} \rightarrow \frac{v}{\sqrt{\|\boldsymbol{v}\|_2^2 + \varepsilon^2}} \quad (3)$$

where ε is the small constant is mostly 1and $\boldsymbol{v}$ is a normalized vector.

3.4 Features Selection

In this section, feature importance was calculated for the data set used by a tree-based classification model such as Cat Boost. All the classifiers use various features that are pre-calculated and different thresholds to maximize accuracy.

3.4.1 Finding the Important Features Using CatBoost

CatBoost is a machine learning algorithm used to categorize data by means of a certain category [18]. It also gives details on the relevance of any given feature in decision-making [19]. The data gave the best accuracy after only 1000 iterations. Here, Cat Boost may be described as follows:

Assuming we have a training dataset containing N samples and M features as (x_i, y_i), since x_i is a vector of M features and y_i is the target variable, then the goal of CatBoost is to learn the function which is computed using the Eq. (4) mentioned below.

$$G(x) = G_0(x) + \sum_{m=1}^{M} \sum_{i=1}^{N} g_m(x_i) \quad (4)$$

There is function referred to as $G(x)$ that Cat Boost Try's to Learn Fully. This function accepts an input vector x and results in the target variable y, $G_0(x)$ refers to the first guess or initial assumption. In this case, the guess is typically selected as the median value of the target variable over the training dataset. $\sum_{m=1}^{M}$ stands for the summation throughout the ensemble of trees. M denotes the number of trees. $\sum_{i=1}^{N}$ denotes the sum of the training samples, N = the number of training samples, $g_m(x_i)$ represents is the predicted value of m tree for i training sample and Each tree inside the forest predicts the label for each training set and combines its results.

The equation is a summary of how the final prediction $G(x)$ can be calculated as a summation of the initial guess $G_0(x)$ and the predictions of each tree, $g_m(x_i)$ for each training sample. This average value is computed for all trees and on all training samples.

3.5 Classification

Machine learning can be subdivided into two folds: the first being supervised learning and the second one is called unsupervised learning. The supervised-learning method is the notion of the output formation from the inputs. It can facilitate the process of classification and help in the complicated diagnosis of cases. Being more accurate, supervised learning is more widely applied in research studies Classification from our standpoint has been done by ANN method, supervised learning tool.

3.5.1 Artificial Neural Network (ANN)

The suggested framework includes the input layer, three fully connected layers and the output layer [12, 20]. Features of the dataset would be represented as the input layer, whereas the three hidden layers would have 512, 256, and 128 neurons. The first fully connected layer employs the ReLU (Rectified Linear Unit) activation function to hasten the training procedure. ReLU takes values of inputs that are negative into zero whereas values that are positive are not altered and combined via Eq. (5) below [21].

$$f(x) = \max(0, x) \quad (5)$$

where x is the input value.

The output layer contains a number of neurons equal to the number of classes in the dataset, using the Softmax activation function for multi-class classification [22] was used in this research, as given in Eq. (6) below:

$$P = \frac{e^{zi}}{\sum_{j=1}^{n} e^{zj}} \tag{6}$$

In which P is a probability, e^{zi} is exponent of input vector, and e^{zj} is the exponent of the output vector.

The network uses Categorical-Cross Entropy loss to reduce the difference between predicted and actual distribution and improves accuracy [22]. Loss function provided in Eq. (7) below.

$$L = -\frac{1}{N}\sum_{i=1}^{n}\sum_{j=i}^{C} y_{i,j} \cdot \log(P_{i,j}) \tag{7}$$

Where N represents the number of samples, C denotes the number of classes, $y_{i,j}$ represents the true distribution of sample i belonging to class j, $P_{i,j}$ represents the predicted distribution by the model that sample i belongs to class j, and L denotes the loss function. The classification layer determines the number of classes from the size of the preceding layer, with the fully connected layer output predicting the final classification as $[1\times1\times N]$, where N is the number of classes.

In order to optimize the model, the Adaptive Moment Estimation (Adam) optimizer was employed on account of its success and memory efficiency in complicated tasks. Adam uses momentum and scaling in combination with dynamic learning rates [22]. The weight update rule for Adam is given by Eq. (8) below.

$$w_t = w_{t-1} - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \varepsilon} \tag{8}$$

where $\hat{m}_t$ is the first moment mean of gradients, $\hat{v}_t$ is the second moment mean of squared gradients, ε prevents division by zero, and η is the learning rate for minimizing the cross-entropy loss.

4 RESULTS

4.1 Experimental Settings

Here, we shall discuss the experimental findings and compare the performance of our basically recommended model with that of some other methods.

4.1.1 Datasets

The initial images of the Harvard dataset were obtained from Kim's Eye Hospital [23]. The dataset includes 788 normal images, 289 early glaucoma images, and 467 advanced glaucoma images. The data is already preprocessed. Some images of this database are shown in Fig. 3.

The RIM-ONE DB, on the other hand, has approximately 169 images. The pictures are assigned to different classes: all 118 normal images, 12 early glaucoma images, 14 moderate glaucoma images, 14 advanced glaucoma images, and 11 ocular hypertension (OHT) images. Fig. 4 presents photographs of healthy individuals and glaucoma patients with early, moderate, and severe cases of the glaucoma as examples of the RIM-ONE dataset [24].

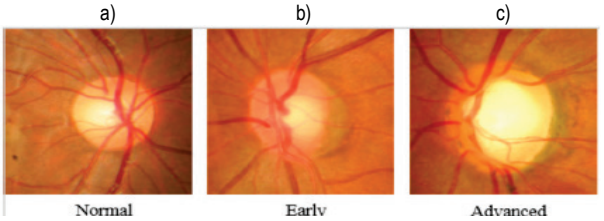


Figure 3 The normal image (a), early glaucoma (b), and advanced glaucoma (c) from the Harvard dataset

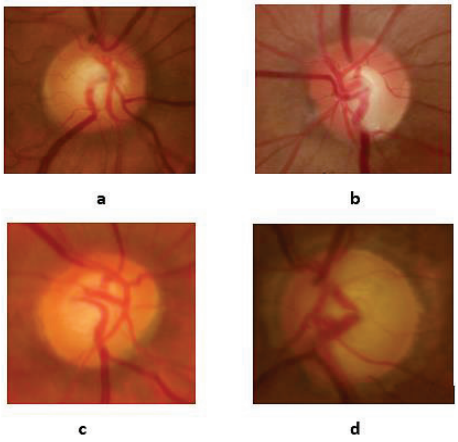


Figure 4 The normal image (a), mild glaucoma (b), moderate glaucoma (c), and severe glaucoma (d) from the RIM-ONE dataset

4.1.2 Implementation Details

It trained the model with the batch size of 32, and all images included in the dataset were resized to 128 × 128 pixels. The test set detection accuracy was registered at 100 epochs of model training.

4.1.3 Evaluating the performance of Machin Learning

The described models' performance is evaluated using accuracy, sensitivity, specificity. Diagnosis model accuracy is called the ability of the model to precisely distinguish between patients with and without disease status Eq. (9) shows how to obtain accuracy [12].

$$Accuracy\,(ACC) = \frac{TP + TN}{TP + TN + FP + FN} \tag{9}$$

The sensitivity of a diagnosis model is interpreted as the probability that the model will correctly recognize the sick among all sick patients. This ratio is calculated as shown by Eq. (10) [12].

$$Sensitivity\,(SN) = \frac{TP}{TP + FN} \tag{10}$$

Specificity is the ability to differentiate between tested persons without disease. This ratio is calculated as shown by Eq. (11) [12].

$$Specificity\,(SP) = \frac{TN}{TN + FP} \tag{11}$$

4.2 Multi-stage Glaucoma

During our experiments, the data is divided into two sets—a training set and a testing set. The training set is comprised of a round of 80 percent, with the rest being the testing set of 20 percent.

In the case of the model, we used the Adam optimizer to tune the parameters based on the suggested default values, i.e., the parameters were fixed as $\beta_1 = 0.9$ and $\beta_2 = 0.999$ in the training of the artificial neural network (ANN). Also, the learning rate was endowed with 0.001 in order to arrive at the ideal value of categorical cross-entropy loss [21] to the ANN.

The training and validation accuracy curves for the Harvard and RIM-ONE datasets were plotted as shown in Figs. 5 and 6, respectively

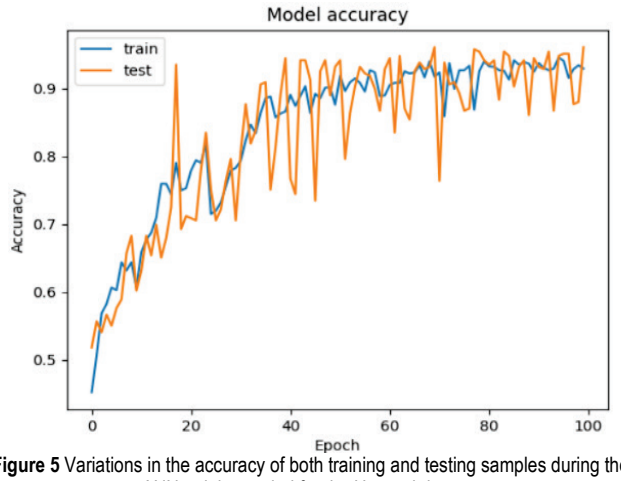


Figure 5 Variations in the accuracy of both training and testing samples during the ANN training period for the Harvard dataset

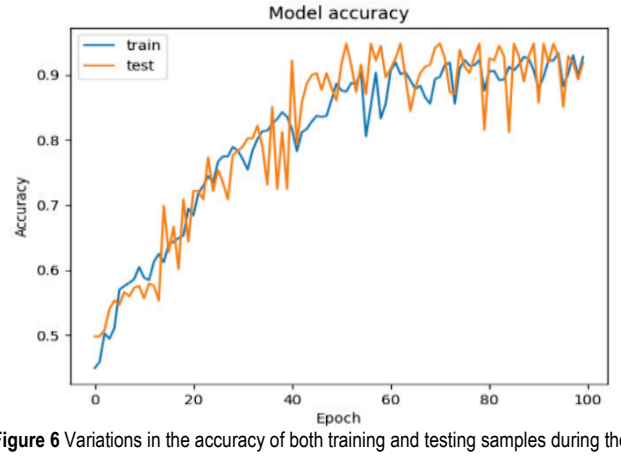


Figure 6 Variations in the accuracy of both training and testing samples during the ANN training period for the RIM-ONE dataset

For the Harvard dataset, the training accuracy achieved was 96.12%, while for the RIM-ONE dataset, the training accuracy achieved was 91.67%.

At the beginning of training, accuracy values are low but gradually improve as training progresses in classifying the stages of glaucoma (multi-class classification) for both datasets. In the early stages, the test accuracy is higher than the training accuracy, then the values converge as training continues.

In the case of the Harvard data (Fig. 5), training and test accuracy were low initially and improved as the epochs progressed, ending with an accuracy of 96.12% during training. The accuracy of the testing was close to the training curve, showing a high level of generalization.

Training performance and test performance trained on the RIM-ONE data set Fig. 6 were initially low but continued to increase consistently across the durations of the epochs. Though the test curve had more convulsions as the sample was small, the overall pattern fitted close to the training curve, with a convergence and high degree of generalization as pertains to the case of the Harvard dataset .The loss and validation curves for the Harvard and RIM-ONE datasets were plotted as shown in Figs. 7 and 8, respectively.

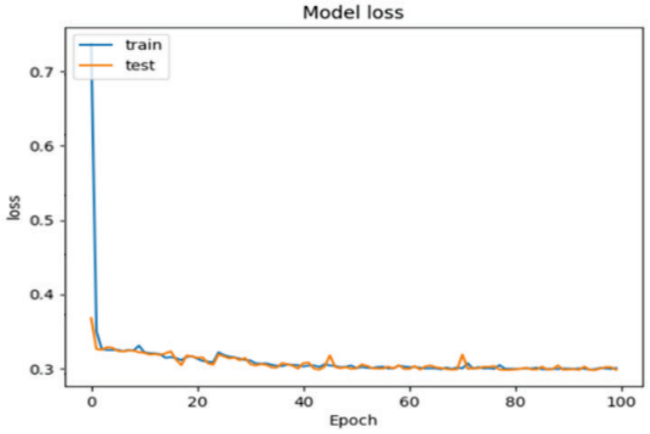


Figure 7 Variations in the loss of both training and testing samples during the ANN training period for the Harvard dataset

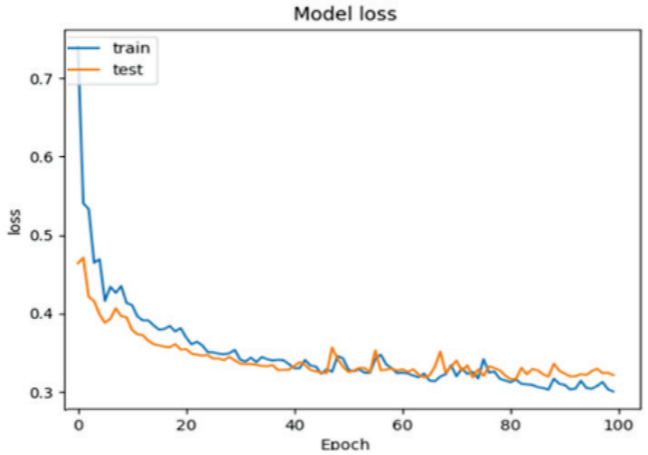


Figure 8 Variations in the loss of both training and testing samples during the ANN training period for the RIM-ONE dataset

Concerning the loss function, it begins by taking a high value before reducing at a sharp rate, depicting that the model

initially had a tricky time with the classification, and then it began to act better. The loss values of the test samples are less than that of the training samples and then tend to come closer and closer with training.

For the Harvard dataset Fig. 7, the training loss started at a high value and gradually decreased as the epochs progressed. The loss values of the test samples were initially lower than those of the training samples, and with further training, the two curves converged, indicating that the model was successfully minimizing the categorical cross-entropy loss while maintaining generalization.

In the case of RIM-ONE data, Fig. 8 the training loss exhibited an initial steep decline as the model takes time to pick up on its accuracy. The losses of the test samples were converted more strongly than the training samples, but as the training proceeded, the two curves slowly approached, which is another evidence that ANN was capable of the generalization encountered even though the size of the dataset was small.

To evaluate the model's performance and its ability to separate and distinguish between categories, the following metrics were used: Sensitivity, Specificity and Accuracy. In Tabs. 1 and 2 Sensitivity, Specificity and Accuracy rates for the proposed model.

Table 1 Multiple classification on the Harvard dataset using the proposed model

The datasets (Classes)		Sensitivity	Specificity	Accuracy
Harvard	Advanced	94.95%	99.52%	96.12%
	Early	96.72%	95.97%	
	Normal	96.64%	99.38%	

Harvard dataset results show a high performance of the model in all glaucoma stages. Sensitivity and specificity are always high and reach their maximum in the Advanced as well as Normal category (96.12 percent), showing that they are highly classified. Markedly, also the model got a high sensitivity at the early stage (96.72%), an aspect that holds its potential in effective identification of glaucoma at an early stage.

Table 2 Multiple classification on the RIM-ONE dataset using the proposed model

The datasets (Classes)		Sensitivity	Specificity	Accuracy
RIMONE	Advanced	100%	100%	91.67%
	Early	100%	97.14%	
	Normal	100%	81.82%	
	Moderate	33.33%	100%	
	OHT	66.67%	100%	

The performance on the RIMONE dataset indicates the above model makes perfect sensitivity (100%) and high specificity in detecting advanced, early, and normal stages of glaucoma. Still, sensitivity on the Moderate and OHT stages is significantly lower (33.33% and 66.67%, respectively), which may be explained by fewer samples and the less obvious nature of these stages; that is why they can be detected less precisely.

4.3 Confusion Matrix

The confusion matrix delivers a comprehensive assessment of the classification accuracy in terms of the comparison of actual labels and predicted classes. It provides a glimpse into the effectiveness with which the various glaucoma levels and normal scenarios are discriminated by the proposed system. Confusion matrices for the Harvard and RIM-ONE datasets are attached as Figs. 9 and 10, respectively.

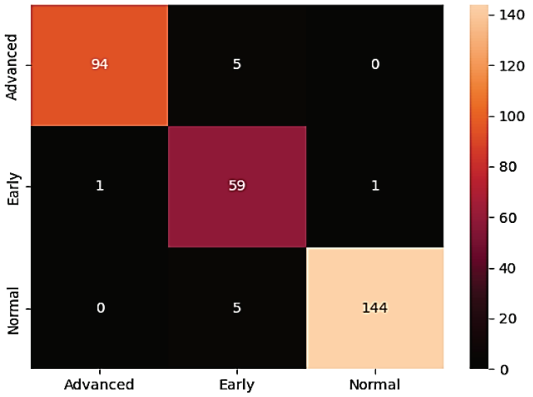


Figure 9 Confusion matrix for identifying types of glaucoma using ANN for the Harvard dataset

94 advanced cases are precisely classified as being advanced, whereas a few cases are misclassified between early and normal, which shows that the system is robust with little confusion in borderline cases.

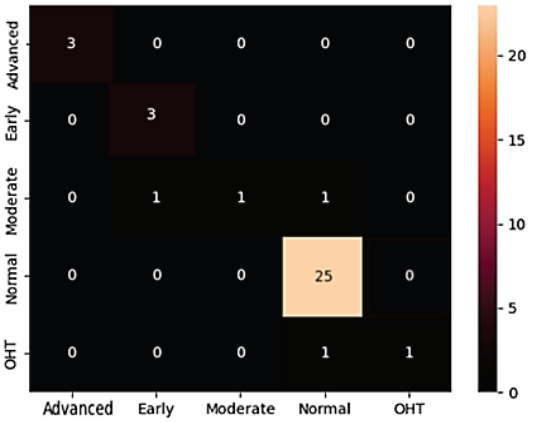


Figure 10 Shows the confusion matrix for identifying types of glaucoma using ANN for the RIM-ONE dataset

The model has shown acceptable classification even with a reduced data size, but only some normal cases (25) can be classified accurately, and there is minor confusion with mild or normal classes. This implies that although the model has high generalization ability, sensitivity as a differentiator of faint glaucoma stages may be hampered by the small quantity of samples in some of the categories.

4.4 Performance Comparison

In our search, we classified the RIM ONE database into five stages of glaucoma (advanced, early, moderate, healthy, and ocular hypertension). While we did not find any research with such a classification to compare the results of our research with.

We compared glaucoma identification with existing methods, taking into account the Harvard data set, and the results are presented in Tab. 3.

Different CNN models [25] : such as pre-trained CNN networks (ResNet50, AlexNet, VGG19, DenseNet-201, and Inception-ResNet-v2), were used for feature extraction and classification.These models achieved moderate performance, with accuracy up to 85.43%.They did not do feature selection but directly employed deep features of the fundus images.

U-Net and Xcep-Dense [26]: networks were combined, with U-Net being used for segmenting the optic disc region and Xcep-Dense utilized for both feature extraction and classification, reaching 87.58% accuracy.

the GS-Net [27]: GS-Net is a CNN that applied a Global Self-Attention Module (GSAM). The GSAM assists the model to concentrate on those most significant aspects of the retina only to consider the features of the picture about them and also to locate them so that the model could see the whole picture of the retina. Conversely, our model is based on HOG features, which capture the distribution of edge orientations and gradients in the retinal images, emphasizing the most informative aspects of the images and achieving high accuracy in glaucoma stage classification.

HOG_ANN: Initially, all extracted HOG features were used for classification without selecting the most important ones, resulting in an accuracy of 84.14%.

The proposed model: The model uses HOG features with CatBoost for selecting only the most important features, followed by classification with ANN. the model achieved improved performance with 96.12% accuracy, 97.10% sensitivity, and 98.29% specificity, showing the importance of feature selection in improving the outcome of the classification process.

Table 3 Comparative for glaucoma classification in Harvard dataset

Method	SP (%)	SEN (%)	ACC (%)
Different CNN models [25]	92,14%	80,80%	85,43%
U-Net and Xcep-Dense [26]	-	-	87.58%
GS-Net [27]	-	-	84.91%
HOG_ANN	92.01%	86%	84.14%
Proposed model	**98.29%,**	**97.10%**	**96.12%**

The results of the comparison make clear that the proposed solution is superior to those that can be implemented in practice, especially thanks to the effective feature selection step, as it added to the overall enhancement of the performance. As evidenced in the comparison table, the combination of the CatBoost algorithm during the selection of the most relevant features enabled the model to concentrate on important information, thus making it less redundant and offering it a better chance of distinguishing between various glaucoma stages.

The proposed method reveals a more focused and effective use of the given information as compared to its counterparts that tended to rely on greater numbers of raw features. Such augmentation further emphasizes the role of a combination of advanced feature selection methods with powerful models of classification, the provision of a more obvious and less subjective framework of glaucoma-stage provisions, and the entering of the possible prospects of clinical practice.

5 DISCUSSION

The findings of the study illustrated that feature extraction resulted in accuracy, sensitivity, and specificity values higher than 96%. The experiments in this study also show that the feature extraction technique and the selection of significant features for disease diagnostics give good results. There was a comparison made with the results of this study and other literature on glaucoma screening. The outcomes are contrasted with the help of all the information presented in the tables above. The same datasets are used in the study we discuss. It can be concluded, that our approach is more accurate compared to other studies. The benefit of this research is that it uses traditional machine learning methods like ANN, but at the same time, it provides an accuracy, sensitivity, and specificity of more than 96%.

Here, we apply the CatBoost algorithm, which is critical in extracting features that distinguish the stages of glaucoma.

6 CONCLUSION AND FUTURE WORK

In this paper, we proposed an AI-based framework for the automatic detection of glaucoma using Histogram of Oriented Gradients (HOG) for feature extraction, CatBoost-based feature selection, and Artificial Neural Networks (ANN) for classification. The method was tested against two test data sets, Harvard (1,544 images) and RIM-ONE (169 images).

The results of our method were also quite good, with the accuracy, sensitivity, and specificity being 96.12%, 97.10%, and 98.29% on the Harvard dataset and 91.67%, 80.00%, and 95.79%, respectively, on the RIM-ONE dataset. Such research evidence shows that the problem of detecting glaucoma automatically, using feature extraction, can prove to be extremely useful for researchers and doctors in the future.

The principal asset of this research is in the combination of the efficient feature selection and ANN classification technique, which minimizes irrelevant data and maintains the accuracy of the diagnostic data. In contrast to most of the previous efforts, the framework demonstrates a high generalization potential in several collections of information that have various properties.

In the future, there are a few directions that can be taken that would augment this study. First, the proposed model can be put into practice in real-life clinical trials, which will confirm its appropriateness in daily medical practice. Second, by growing the datasets to an increased number and variety of populations, across imaging modalities and acquisition settings, one would strengthen robustness and generalization. Thirdly, the combination of the multi-modal data (i.e., OCT, fundus photography, and patient clinical records) might present a broader diagnosis tool. Lastly, the framework could be generalized to identify and classify other diseases of the eye, like diabetic retinopathy or macular degeneration, to increase the clinical significance. Overall, such work is also important to the science of computer-assisted diagnosis that identifies glaucoma and introduces the robust fundamentals to develop AI-assisted ophthalmology devices that could be deployed in the practical world of health care.

7 REFERENCES

[1] Tham, Y. C., Li, X., Wong, T. Y., Quigley, H. A., Aung, T., & Cheng, C. Y. (2014). Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology, 121*(11), 2081-2090. https://doi.org/10.1016/j.ophtha.2014.05.013

[2] Sommer, A., Tielsch, J. M., Katz, J., Quigley, H. A., Gottsch, J. D., Javitt, J., & Singh, K. (1991). Relationship between intraocular pressure and primary open angle glaucoma among white and black Americans: the Baltimore Eye Survey. *Archives of ophthalmology*, *109*(8), 1090-1095 https://doi.org/10.1001/archopht.1991.01080080050026

[3] Shabbir, A., Rasheed, A., Shehraz, H., Saleem, A., Zafar, B., Sajid, M., ... & Shehryar, T. (2021). Detection of glaucoma using retinal fundus images: A comprehensive review. *Mathematical Biosciences and Engineering, 18*(3), 2033-2076. https://doi.org/10.3934/mbe.2021106

[4] Chan, Y. M., Ng, E. Y. K., Jahmunah, V., Koh, J. E. W., Lih, O. S., Leon, L. Y. W., & Acharya, U. R. (2019). Automated detection of glaucoma using optical coherence tomography angiogram images. *Computers in biology and medicine*, *115*, 103483. https://doi.org/10.1016/j.compbiomed.2019.103483

[5] Hatanaka, Y., Fukuta, K., Muramatsu, C., Sawada, A., Hara, T., Yamamoto, T., & Fujita, H. (2009). Automated measurement of cup-to-disc ratio for diagnosing glaucoma in retinal fundus images. In *World Congress on Medical Physics and Biomedical Engineering*, September 7-12, 2009, Munich, Germany: Vol. 25/11 *Biomedical Engineering for Audiology, Ophthalmology, Emergency & Dental Medicine* (pp. 198-200). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-03891-4_53

[6] Parashar, D., & Agrawal, D. K. (2020). Automated classification of glaucoma stages using flexible analytic wavelet transform from retinal fundus images. *IEEE Sensors Journal*, *20*(21), 12885-12894. https://doi.org/10.1109/JSEN.2020.3001972

[7] Cheng, J., Liu, J., Xu, Y., Yin, F., Wong, D. W. K., Tan, N. M., ... & Wong, T. Y. (2013). Superpixel classification based optic disc and optic cup segmentation for glaucoma screening. *IEEE transactions on medical imaging*, *32*(6), 1019-1032. https://doi.org/10.1109/TMI.2013.2247770

[8] Mohamed, N. A., Zulkifley, M. A., Zaki, W. M. D. W., & Hussain, A. (2019). An automated glaucoma screening system using cup-to-disc ratio via simple linear iterative clustering superpixel approach. *Biomedical Signal Processing and Control, 53*, 101454. https://doi.org/10.1016/j.bspc.2019.01.003

[9] Pathan, S., Kumar, P., Pai, R. M., & Bhandary, S. V. (2023). An automated classification framework for glaucoma detection in fundus images using ensemble of dynamic selection methods. *Progress in Artificial Intelligence*, *12*(3), 287-301. https://doi.org/10.1007/s13748-023-00304-x

[10] Alice, K., Deepa, N., Devi, T., BeenaRani, B. B., & Nagaraju, V. (2023). Effect of multi filters in glucoma detection using random forest classifier. *Measurement: Sensors, 25*, 100566. https://doi.org/10.1016/j.measen.2022.100566

[11] Latif, J., Tu, S., Xiao, C., Bilal, A., Ur Rehman, S., & Ahmad, Z. (2023). Enhanced Nature Inspired-Support Vector Machine for Glaucoma Detection. *Computers, Materials & Continua, 76*(1). https://doi.org/10.32604/cmc.2023.040152

[12] Singh, L. K., Garg, H., & Khanna, M. (2022). Histogram of Oriented Gradients (HOG)-Based Artificial Neural Network (ANN) Classifier for Glaucoma Detection. *International Journal of Swarm Intelligence Research (IJSIR), 13*(1), 1-32. https://doi.org/10.4018/IJSIR.309940

[13] Al-Sarem, M., Saeed, F., Boulila, W., Emara, A. H., Al-Mohaimeed, M., & Errais, M. (2021). Feature selection and classification using CatBoost method for improving the performance of predicting Parkinson's disease. In *Advances on Smart and Soft Computing: Proceedings of ICACIn 2020* (pp. 189-199). Springer Singapore. https://doi.org/10.1007/978-981-15-6048-4_17

[14] Rahouma, K. H., Mohamed, M. M., & Hameed, N. S. A. (2019). Glaucoma Detection and Classification Based on Image Processing and Artificial Neural Networks. *Egyptian Computer Science Journal, 43*(3).

[15] Muduli, D., Kumari, R., Akhunzada, A., Cengiz, K., Sharma, S. K., Kumar, R. R., & Sah, D. K. (2024). Retinal imaging based glaucoma detection using modified pelican optimization based extreme learning machine. *Scientific Reports, 14*(1), 29660. https://doi.org/10.1038/s41598-024-79710-7

[16] Singh, L. K., Khanna, M., Garg, H., & Singh, R. (2024). Efficient feature selection based novel clinical decision support system for glaucoma prediction from retinal fundus images. *Medical Engineering & Physics, 123*, 104077. https://doi.org/10.1016/j.medengphy.2023.104077

[17] Shamna, N. V., & Aziz Musthafa, B. (2023). Feature extraction method using hog with ltp for content-based medical image retrieval. *International journal of electrical and computer engineering systems*, *14*(3), 267-275. https://doi.org/10.32985/ijeces.14.3.4

[18] Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. *Advances in neural information processing systems, 31*.

[19] Dorogush, A. V., Ershov, V., & Gulin, A. (2018). CatBoost: gradient boosting with categorical features support. arXiv preprint arXiv:1810.11363.

[20] Gershenson, C., (2003), Artificial neural networks for beginners. *arXiv preprint cs/0308031*.

[21] Nair, V. & Hinton, G. E. (2010). Rectified linear units improve restricted Boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pp. 807-814.

[22] Terven, J., Cordova-Esparza, D. M., Ramirez-Pedraza, A. & Chavez-Urbiola, E. A. (2023). Loss functions and metrics in deep learning. *arXiv preprint arXiv:2307.02694*.

[23] Kim, U. (2018). Machine learn for glaucoma. *Harvard Dataverse*.

[24] Fumero, F., Alayón, S., Sanchez, J. L., Sigut, J., & Gonzalez-Hernandez, M. (2011, June). RIM-ONE: An open retinal image database for optic nerve evaluation. In *24th IEEE international symposium on computer-based medical systems (CBMS2011)* (pp. 1-6). https://doi.org/10.1109/CBMS.2011.5999143

[25] Velpula, V. K., & Sharma, L. D. (2023). Multi-stage glaucoma classification using pre-trained convolutional neural networks and voting-based classifier fusion. *Frontiers in Physiology*, *14*, 1175881. https://doi.org/10.3389/fphys.2023.1175881

[26] Desiani, A., Priyanta, S., Ramayanti, I., Suprihatin, B., Halim, M. R., Geovani, D., & Rayani, I. (2023). Multi-Stage CNN: U-Net and Xcep-Dense of Glaucoma Detection in Retinal Images. *Journal of Electronics, Electromedical Engineering, and Medical Informatics, 5*(4), 211-222. https://doi.org/10.35882/jeeemi.v5i4.314

[27] Das, D., & Nayak, D. R. (2023, October). Gs-net: Global self-attention guided cnn for multi-stage glaucoma classification. In *IEEE International Conference on Image Processing (ICIP2023)* (pp. 3454-3458). https://doi.org/10.1109/ICIP49359.2023.10222689

Authors' contacts:

Dimah Alkasem, Master's degree student
(Corresponding author)
Department of Computer Engineering,
College of Electrical and Electronic Engineering, University of Aleppo,
12212 Aleppo, Syria
dimaalkasem581@gmail.com

Hazem Issa, PhD
Department of Computer Engineering,
College of Electrical and Electronic Engineering, University of Aleppo,
12212 Aleppo, Syria
hazemisaa17@gmail.com

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250213103422

Original scientific paper
Received: 2025-02-13, Accepted: 2025-03-12

Experimental Study and Performance Analysis of the Photovoltaic System Output under Real Weather Conditions

Danijela Kardaš Ančić*, Mirko Komatina, Ante Čikić, Petar Gvero*, Simon Muhič

Abstract: The aim of this study is to evaluate PV output, considering PV module temperature and real meteorological parameters. An experimental investigation was performed in warmest month, July, and the following parameters were measured: PV module temperature, PV voltage and current, solar irradiation, air temperature, and wind velocity. It was shown that increasing PV module temperature above 25 °C leads to voltage drops and decreased electrical efficiency. The minimum measured average daily PV module temperature was 40.86 °C and the maximum was 42.59 °C. It is shown that when the PV module temperature exceeds 25 °C, the electrical efficiency starts to decrease at a rate of 0.449 – 0.455%/°C. It can be concluded that for approximately 2.2 °C above 25 °C leads to a 1 % reduction in the PV module's electrical efficiency. These results are significant for designing PV systems and evaluating energy yield considering the climate parameters in which they will operate.

Keywords: experimental study; photovoltaic output; temperature dependence; weather conditions

1 INTRODUCTION

Climate changes and secure energy supply chains are key factors in the transformation processes of energy systems. The sixth Intergovernmental Panel on Climate Change report concludes that global greenhouse gas emissions must be reduced by 43% by 2030 to achieve the goal of limiting global warming to 1.5 °C [1]. To fight with all these issues, different scenarios to use various types of renewable energy sources by many researchers have been suggested [2, 3]. These scenarios propose how to reduce fossil fuels and to use more renewable energy sources. Since 1990, the adoption of renewable energy has been rising in both developed and developing countries. This shift is primarily driven by technological progress, lower costs, government support for renewable energy, climate change concerns, and a desire to reduce dependence on foreign energy [4, 5]. European Union proposes to increase the target in the Renewable Energy Directive [6]**.** In 2023, renewable energy capacity globally increased by almost 50% and this is fastest growth rate in the past two decades [7]. According to STEPS (Stated Policies Scenario), by 2030, 80 % of new capacity will be renewable energy sources. Solar PV alone are set for more than half, solar PV alone are set [8]. Over the coming five years, renewable energy milestones that expected to be achieved are: renewable energy sources will overtake coal and will become the leading source of electricity generation in 2025 and solar PV will overtake nuclear electricity generation in 2025 and 2026 respectively [7]**.** In addition to the standard benefits of reducing greenhouse gas emissions, improving air quality, and protecting natural resources, the transition to renewable energy sources stimulates economic growth, creates new jobs, and contributes to a more sustainable and environmentally friendly energy system [9].

The photovoltaic (PV) technology is the most used technologies available for solar energy utilization. Despite all the challenges, solar energy is one of the most favourable renewable energy resources [7, 10, 11]. Solar energy provides many advantages, including great potential, the affordability of the prices per individual system, zero or decreased greenhouse gas emissions, no noise emission and no need for fossil fuels [12]. The potential of solar energy is unlimited, because it is available virtually everywhere. For example, in less than 9 minutes, the Earth is illuminated with as much solar energy as humanity needs in a year [13]. Solar energy is a renewable resource capable of replacing large amounts of conventional energy, but its utilization is restricted by technological, economic, and geographical limitations [14]. In 2023, solar PV alone took three-quarters of renewable capacity worldwide [7]. As part of the REPowerEU plan, EU Solar Energy Strategy aims to bring over 320 GW of solar photovoltaic by 2025 and almost 600 GW by 2030 [15]. Solar PV and wind energy will take 95% of worldwide renewable capacity by 2028.

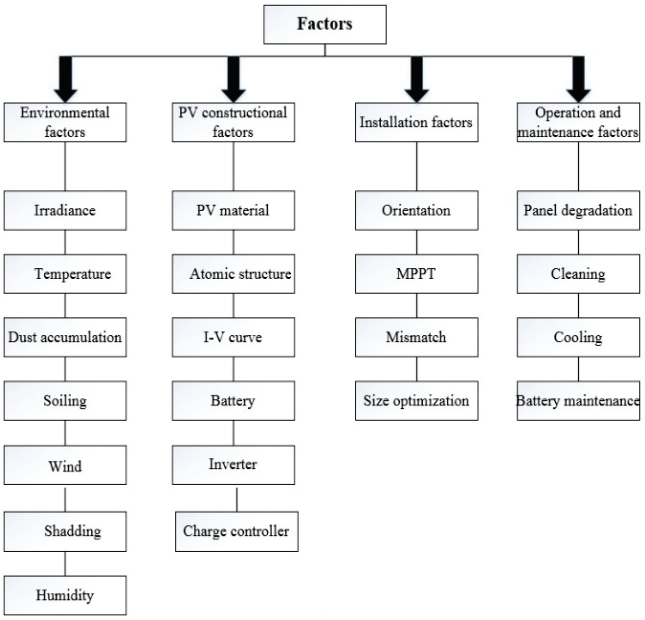


Figure 1 PV panel efficiency degradation factors

1.1 Overview: Solar PV Efficiency

Typical PV module convert up to 23% of the solar energy into electricity while the rest is mainly converted into

heat, depending on the different factors [16-19]. PV efficiency depends on many parameters (Fig. 1). There are developed standard test procedure to determine the characteristics of photovoltaic modules from different manufacturers under equal conditions. These conditions are known as standard test conditions (STC). The standard conditions determine that the module testing is performed at a total irradiance of 1000 W/m^2, air mass AM (defined by standards ISO 9845-1:2022 and ASTM G173-03) 1.5, and that the temperature of the cells is 25 °C. Under these conditions, all the operating parameters of the PV panel, as well as its electrical efficiency, are specified. In real conditions, the irradiance is lower and the temperature of the cells inside the module is higher than the values of STC. The authors [20] concluded that approximately 50% of the analysed articles focused on the impact of environmental factors on PV performance. The most important environmental factors are: solar irradiance, wind velocity, air temperature, dust accumulation, soiling and effect of shading [20-25].

In [26] it is found that water droplets on the PV panel improved the PV power output by at least 5.6% while birds fouling the PV module surface reduce the PV system performance by about 7.4%. The PV panel output significantly depends on solar irradiance. Irradiance varies due to the meteorological parameters such as weather, geographical location, seasonal changes, sun position in the sky and period of the day. As the irradiance drops below 1000 W/m^2, the *I-V* characteristic of the PV panel changes. The short - circuit current (I_{sc}) decreases in direct proportion to the decrease in irradiance. The short-circuit current is the current in solar cell at the zero voltage across the solar cell. The open circuit voltage (V_{oc}) also decreases, but its percentage change is far less compared to the change in current. The open circuit voltage is the maximum voltage available from the solar cell, and this occurs when the net current thorough device is zero. Temperature is also one of the physical quantities that have a significant impact on the performance of photovoltaic modules. The open - circuit voltage (V_{oc}) decreases with increasing temperature, while the short – circuit current (I_{sc}) increases. As a superposition of temperature variations of current and voltage, the power produced by PV modules also changes depending on temperature. A more intense decrease in voltage dominates compared to a milder increase in current, resulting in the fact that the power decreases as the temperature increases [27]. Temperature of PV cell higher than 25 °C negatively affect the PV electrical efficiency [28, 29]. It is an interesting phenomenon that the photovoltaic panel produces more power when the day is clear and cold than when it is clear and hot. The heat flux generated will increase the module temperature and decrease the electrical efficiency of the PV module. There is a strong relationship between module temperature and the bandgap energy of the PV cell material. When the operating temperature soars, the energy bandgap in materials tends to shrink. This shift allows the cell to soak up photons with longer wavelengths, which can be quite beneficial. This rise in temperature often extends the lifespan of those minority carriers, enhancing overall efficiency. Different mounting systems can have influence on the airflow and ventilation around the PV module. This can affect the heat dissipation of the panel and the module temperature. Tab. 1 gives range of annual PV performance loss due to temperature effect for the different mounting systems [30]. The PV, which are mounted on the sloped roof, have the largest annual loss. Reason for this is the little air circulation between the roof and the panel. This leads in rise of the module temperature. A free - standing and flat roof mounting system allow airflow around the panel which gives a better cooling of the panel, and more heat dissipation occurred.

Table 1 Annual PV performance loss due to temperature effect for the different mounting systems

	Temperature rise (K)	Annual loss due to temperature effect (%)
Sloped roof	20 - 55	1.7 - 11.3
Facade integrated	46 - 52	5 - 7
Free-standing	20 - 26	1.7 - 5
Flat roof	25 - 28	2.5 - 4.0

2 MATERIALS AND METHODS

2.1 PV Module Efficiency as a Function of the Operating Temperature of a Photovoltaic Cell/Module

Operating temperature of a photovoltaic (PV) cell/module has significant impact on its electrical efficiency [31]. There are many correlations, which are analysing the PV cell temperature (T_c) as a function of meteorological variables: the ambient temperature (T_a), solar radiation (G_T), local wind velocity (v), as well as material and system dependent properties [32-36]. The temperature of the back surface of the photovoltaic module (T_m) and the temperature of the photovoltaic cell (T_c) can differ for high intensities of solar radiation [37]. At solar radiation of 1000 W/m^2, this temperature difference is typically 2 – 3 °C for flat - plate modules. This temperature difference can be assumed to be zero for flat – plate with a thermally insulated back surface [37]. As the maximum measured average daily solar radiation for the given location is less than 700 W/m^2, these differences were ignored and $T_c = T_m = T_{PV}$ was adopted. There are wide number of correlations which are analysing temperature dependence of the PV module's electrical efficiency [38]. The effect of temperature on the electrical efficiency of a PV cell/module can be define by using Eq. (1) [38-42]:

$$P_m = I_m \cdot V_m = (FF) \cdot I_{sc} \cdot V_{oc} \tag{1}$$

In this expression, subscript "m" refers to the maximum power point in the module's *I–V* curve. Both the open circuit voltage and the fill factor decrease substantially with temperature while the short-circuit current increases. This leads to a linear relation for the electrical efficiency of a PV given in Eq. (2) [38 - 39]:

$$\eta_{el} = \eta_{STC} \cdot \left[1 - \beta \cdot \left(T_c - T_{STC}\right) + \gamma \log_{10} G(t)\right] \tag{2}$$

which η_{STC} is the module's electrical efficiency at the reference temperature, T_{STC} = 25 °C, and at solar radiation flux of 1000 W/m^2 [39]. The temperature coefficient, β, and the solar radiation coefficient, γ, are mainly material

properties, having values of about 0.004 K^{-1} and 0.12, respectively, for crystalline silicon modules [38-40]. The last part of the Eq. (2), however, is usually taken as zero [38-40], and Eq. (2) reduces to Eq. (3):

$$\eta_{el} = \eta_{STC} \cdot \left[1 - \beta \cdot \left(T_c - T_{STC}\right)\right] \quad (3)$$

This equation represents the traditional linear expression for the PV electrical efficiency [39 - 42]. Reduction in efficiency of PV module can be estimated as follows in Eq. (4):

$$Reduction\ in\ efficiency = \frac{\eta_{STC} - \eta_{calculated}}{\eta_{STC}} \cdot 100\% \quad (4)$$

2.2 Experimental Setup

Since PV module electrical efficiency depends significantly of geographical areas and climate, an own experimental installation was designed, constructed and built in order to be able to measure the temperature of PV module, voltage and current in real time taking into account relevant meteorological parameters such as air temperature, solar radiation and wind velocity. With these real measured parameters, PV electrical efficiency was calculated. Fig. 2 gives different components used in the experiment. Experimental PV installation used for this study is located outside of the building of the Faculty of Mechanical Engineering, University of Banja Luka (44.46282 N, 17.11502 E) in Bosnia and Herzegovina. The meteorological station is installed on the rooftop of the building. The PV module used for the experiment is a monocrystalline type (SZ-100-36M). Technical specification of the PV module are given in more details in Tab. 2. Temperature of PV module was measured by PT100 thermocouple. These thermocouples were placed on the rear of the PV module in two points. These locations were chosen to determine if there was a difference in the temperature distribution across the panel. For this reason, two diametrically different points were chosen for positioning the thermocouples (left corner down and upper right, Fig. 2c). The average value of these two temperatures was used for analysis. Accuracy of PT100 is ±(0.15 + 0.002×|T|) °C for validity range from –50 °C to 300 °C.

When designing the installation itself, it was decided to place the panels on a metal pillar. The reason for this is to ensure an undisturbed flow of wind around the panel and thus to obtain the most relevant data on its influence on the temperature of the panel. Meteorological data including solar radiation, ambient temperature and wind velocity were measured by the meteorological station Luft WS10. Technical data for meteorological station are given in Tab. 3.

The datasets were collected during the period of 3 days (from July 15 to July 17, 2023) so that measurements results could be monitored and evaluated. Month July was selected for these analysis as a warmest month in the year for analysed geographical area. Measurements were performed from 10 AM to 4 PM with a 10-minute recording interval and stored via EASY-E4-UC-12RC1 data acquisition.

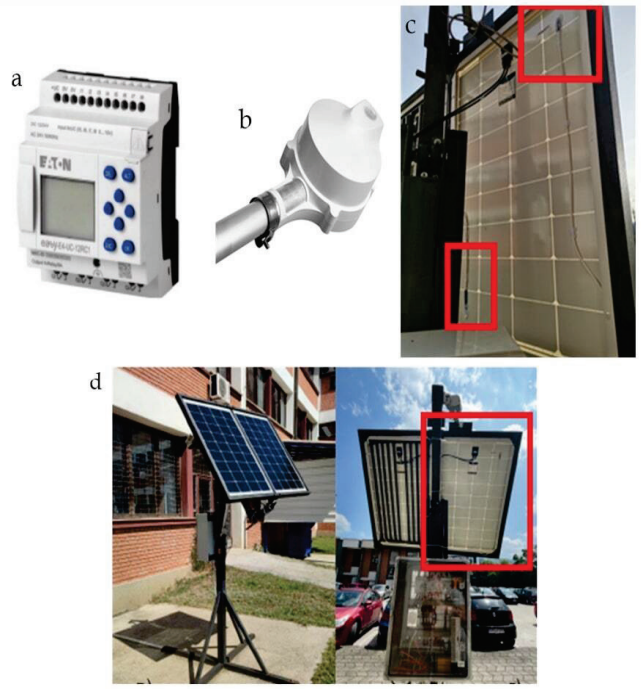


Figure 2 Different outdoor experimental facilities (a - EASY-E4-UC-12RC1 data acquisition, b - meteorological station Luft WS10, c - positions of thermocouples for measuring the PV module temperature, d - experimental setup)

Table 2 Technical specification of the PV module

Standard Test Conditions AM = 1.5, E = 1000 W/m², T_C = 25 °C	
Nominal Power	100 W
Cell Type/Efficiency	Monocrystalline / 17.66 %
Maximum Power Current (I_m)	5.41 A
Maximum Power Voltage (V_m)	18.50 V
Open-circuit Voltage (V_{oc})	22.50 V
Short-circuit Current (I_{sc})	5.92 A
Working Temperature	–40 °C to +80 °C
Module Dimensions	1020 × 670 × 35 mm

Table 3 Technical specification of the meteorological station Luft WS10

Air temperature	PTC, –40 °C to +60 °C (±1.0 °C)
Irradiance	Silicium - Pyranometer, 0 ÷ 1500 W/m² (±10 % or ±120 W/m²)
Wind velocity	0 - 40 m/s (±1 m/s or 5 %)

3 RESULTS

The maximum and minimum values of meteorological data recorded during experiment are presented in the Tab. 4. The analyzed data showed that solar radiation (G_T), ambient temperature (T_a) and wind velocity (v) ranged from 457 - 822.3 W/m², 30.7 - 40.8 °C and 1.3 - 13.5 km/h, respectively.

Table 4 The maximum and minimum values of meteorological data recorded during experiment

Day		G_T (W/m²)	Ta (°C)	v (km/h)
15.7.2023	Max	822.3	37.9	10.3
	Min	457	30.7	1.3
16.7.2023	Max	810.4	39	13.5
	Min	453.3	32.7	1.6
17.7.2023	Max	806.7	40.8	9.3
	Min	461	34.8	1.6

Experimental measured PV module temperature was above 25 °C for whole period of measurement. Average values of meteorological data (air temperature, wind velocity and irradiance) for the whole period of measurement as well as daily average measured PV module outputs are in the Tab. 5.

Table 5 Average values of meteorological data as well as measured PV module outputs in period of measurement on July 15-17, 2023

	On July 15, 2023	On July 16, 2023	On July 17, 2023
Daily average measured air temperature, °C	34.71	36.99	38.72
Daily average measured wind velocity, km/h	5.10	5.34	4.67
Daily average measured irradiance, W/m	703.62	698.15	686.63
Daily average measured PV module temperature, °C	40.86	42.43	42.59
Daily average current, A	4.55	4.40	4.47
Daily average voltage, V	15.41	15.38	15.13
Daily average PV power output, W	70.27	67.67	67.60
Daily average PV electrical efficiency	0.164	0.162	0.161
Daily average reduction in electrical efficiency, %	7.13	7.91	8.00

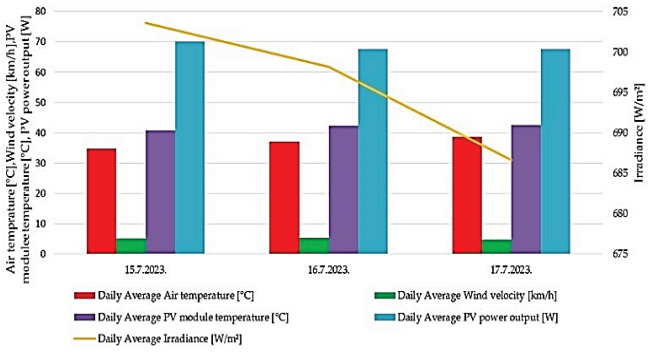


Figure 3 PV power output and average daily meteorological parameters for all three days of measurements

Maximum average daily air temperature was on July 17 (2023) with 38.72 °C and minimum was on July 15 (2023) with 34.71 °C. Maximum average daily PV module was on July 17 (2023) with 42.59 °C and minimum on July 15 (2023) with 34.71 °C. On July 16 (2023), average daily wind velocity had maximum value compared to other two days when experiment was run. Maximum PV power output was on July 15 (2023) with 70.27 W output and minimum PV power output was on July 17 (2023) with 67.60 W output. Fig. 3 presents PV power output and average daily meteorological parameters. Maximum daily average reduction in electrical efficiency corresponds to the warmest day for the whole period of measurement on July 17, 2023 with reduction of 8%.

Minimum daily average reduction is on July 15 (2023) with reduction of 7.13%. On Figs. 4, 5 and 6 are given graphical representation of measured parameters during whole three days of measurement.

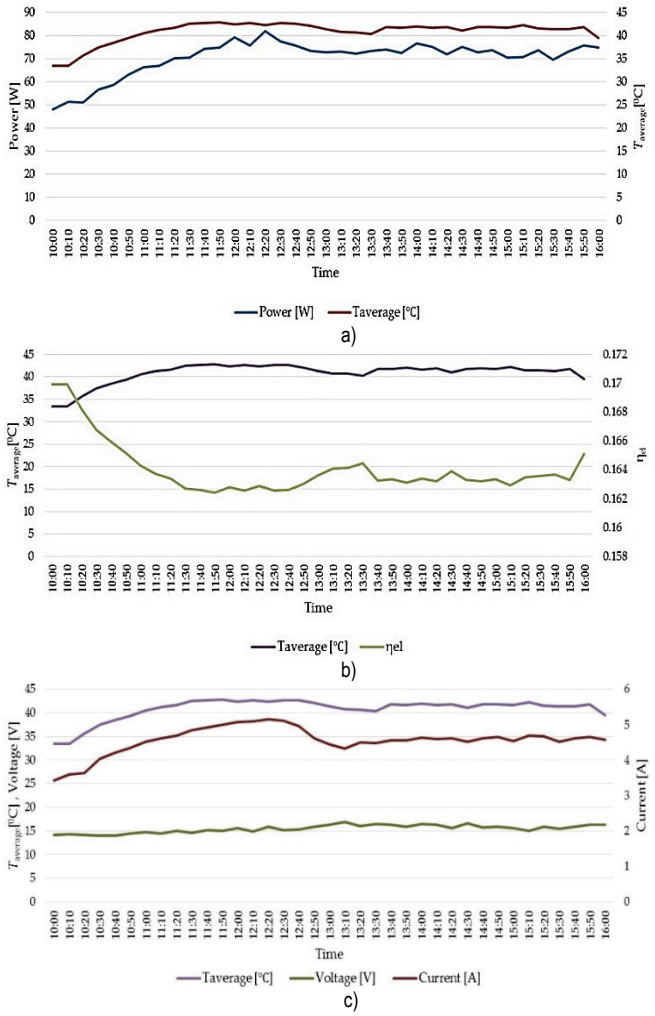


Figure 4 a) Dependence of power on PV module temperature; b) Dependence electrical efficiency on PV module temperature; c) Dependence of voltage and current on PV module temperature (measured on July 15, 2023)

4 DISCUSSION

Analyzing results in Tab. 6 and Figs. 4, 5 and 6, it can be seen that when air temperature increase also increases PV module temperature. Maximum and minimum average daily PV module temperature correspond with the maximum and minimum average daily air temperature. Increasing PV module temperature average daily PV voltage decrease as well as average daily current and power output. Maximum PV power output for the whole period of measurement correspond to the day with minimum average daily air temperature and PV module temperature but also maximum daily average solar irradiance. In addition, on that day, maximum value of PV electrical efficiency for the entire measurement period was recorded. Minimum PV power output correspond to the warmest day for the whole period of measurement with also maximum PV module temperature on the same day. It can be concluded that PV power output is reduced with PV module temperature increase. The same is for PV electrical efficiency. Increasing PV module temperature leads to the PV electrical efficiency decrease as expected.



Figure 5 a) Dependence of power on PV module temperature; b) Dependence electrical efficiency on PV module temperature; c) Dependence of voltage and current on PV module temperature (measured on July 16, 2023)

It is shown that an increase in the PV module temperature above 25 °C leads to a 0.455 - 0.449% reduction in electrical efficiency for each 1°C rise. In Figs. 4, 5, 6, a), it can be seen that with increasing PV module temperature, PV power output is reduced. Drops in PV module temperature corresponds to increases in PV power output. Analyzing influence of PV module temperature on electrical efficiency it can be concluded that with increasing PV module temperature PV electrical efficiency decrease (Figs. 4, 5, 6, b). Measurements over three days show a positive correlation between PV module temperature and current. As it can be seen on Figs. 4, 5, 6, c) current curve had a similar distribution as PV module temperature. With increasing PV module temperature voltage decrease. Maximum values of voltage correspond to the drops of PV module temperature (Figs. 4, 5, 6, c).

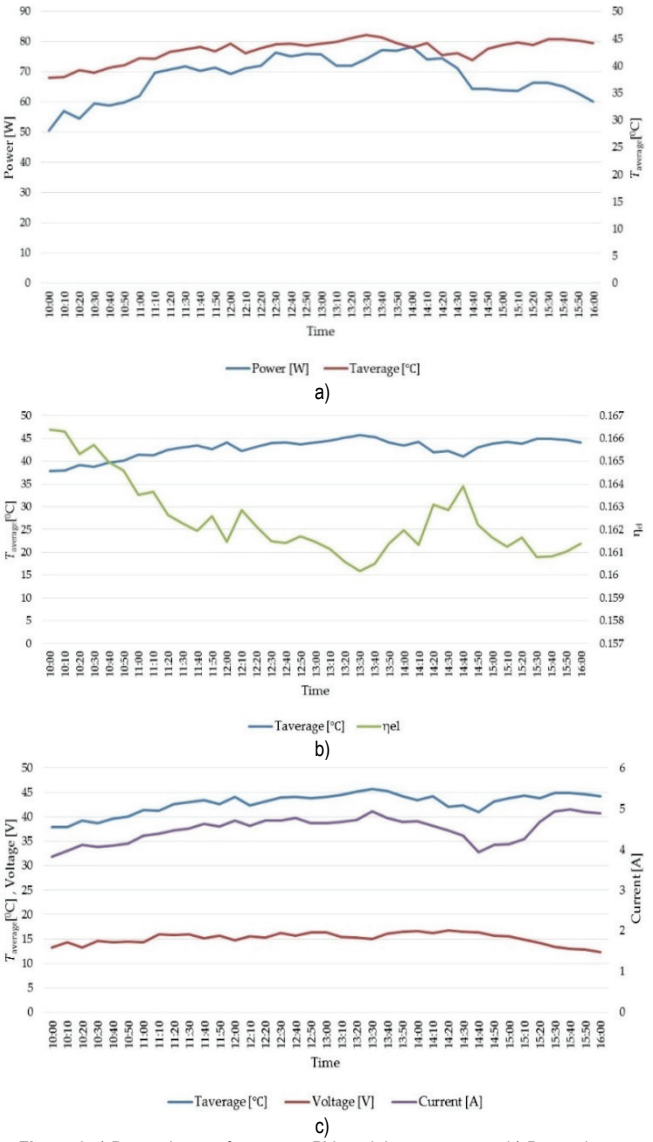


Figure 6 a) Dependence of power on PV module temperature; b) Dependence electrical efficiency on PV module temperature; c) Dependence of voltage and current on PV module temperature (measured on July 17, 2023)

5 CONCLUSION

Solar radiation that reaches the surface of the photovoltaic (PV) module is partly transformed into electrical energy. The most of radiation is transformed into heat. The temperature of the photovoltaic module increases which causes a decrease in their electrical efficiency particularly at temperature above 25 °C. In this study, an evaluation of PV output taking into account real weather conditions and PV module temperature is presented. Experimental investigation was performed in summer time, in July in period of three days. Considering all the analyses, it can be concluded that weather conditions—such as sunlight exposure, air temperature, and wind speed—greatly affect the temperature of PV modules and, as a result, their performance. Maximum and minimum average daily PV module temperature correspond with the maximum and minimum average daily air temperature for the whole period of measurement. Measurement analysis shows that higher PV

module temperatures correlate with lower PV voltage, average power, and efficiency. It was shown that an increase in PV module temperature in range 2.199 - 2.224 °C above 25 °C leads to an approximately 1% reduction in electrical efficiency. These results are significant because they showed how the real operating conditions influence on the PV module. In addition, these results contribute to evaluating energy yield considering the climatic parameters. This is very important for some further strategic energy planning in the field of solar energy. Extremely high air temperatures and solar radiation, causes increase of the temperature of the PV module and it can be expected to exceed 25 °C for almost half of the year in analysed geographical area. Accordingly, it is necessary to analyse the cooling possibilities of the PV modules duo to decrease PV module temperature.

Acknowledgments

The authors acknowledge the support of the European Research Executive Agency (REA) for funding this research under the project "ENPOWER – Enhancing Scientific Capacity for Energy Poverty (101160253 – ENPOWER – HORIZON – WIDERA – 2023 – ACCESS – 02)".

The authors acknowledge the support of the Science Fund of the Republic of Serbia, Grant No. 4344, "Forward-Looking Framework for Accelerating Households" Green Energy Transition – FF GreEN and by the Ministry of Science, Technological Development and Innovation of the Republic of Serbia; grant number 451-03-137/2025-03/200105.

Disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

6 REFERENCES

[1] IPCC. Summary for Policymakers. In: Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.

[2] Elqadhi, M. E. R., Škrbić, S. V., Mohamoud, O. A., & Ašonja, A. N. (2024). Energy integration of corn cob in the process of drying the corn seeds. *Thermal Science, 28,* 3325-3336. https://doi.org/10.2298/TSCI231223104E

[3] Østergaard, P. A., Duic, N., Noorollahi, Y., & Kalogirou, S. A. (2021). Recent advances in renewable energy technology for the energy transition. *Renewable Energy*, *179*, 877-884. https://doi.org/10.1016/j.renene.2021.07.111

[4] Aksoy, M. H., & Ispir, M. (2023). Technoeconomic feasibility of different photovoltaic technologies. *Applied Engineering letters*, *8*, 1-9. https://doi.org/10.18485/aeletters.2023.8.1.1

[5] Shanga, Y., Han, D., Gozgor, G., Mahalik, M. K., & Sahoo, B. K. (2022). The Impact of climate policy uncertainty on renewable and non-renewable energy demand in the United States. *Renewable Energy. 197(1)*. 654-667. https://doi.org/10.1016/j.renene.2022.07.159

[6] European Commission. REPowerEU Plan. 2022

[7] International Energy Agency. Renewables 2023 - Analysis and forecast to 2028, 2024

[8] International Energy Agency. World Energy Outlook, 2023

[9] Lungadee, T., Tippayawong, Y. K., Tangparitkul, S., Jaroenkhasemmeesuk, C., Chaichana, C., Wongspai, W., & Tippayawong, N. (2024) Decarbonizing a Thai coal plant: effect of flue gas loads on carbon capture performance and economics. *Applied Engineering letters*, *9*, 12-21. https://doi.org/10.46793/aeletters.2024.9.1.2

[10] Akrami, E., Gholami, A., Ameri, M., & Zandi, M. (2018). Integrated an innovative energy system assessment by assisting solar energy for day and night time power generation: exergetic and Exergo-economic investigation. *Energy Conversion and Management*, *175*, 21-32. https://doi.org/10.1016/j.enconman.2018.08.075

[11] Aslan, G., Mohammad, A., Majid, Z., Roghayeh, G. G., Saeed, J. G., Hussein, A. K., & Ali, H. A. A. (2023). Impact of harsh weather conditions on solar photovoltaic cell temperature: Experimental analysis and thermal-optical modeling. *Solar Energy*, *252*, 176-194. https://doi.org/10.1016/j.solener.2023.01.039

[12] Ahmed, H., Abdelgalil, E., & Hesham, M. E. (2023). Experimental and numerical investigation for PV cooling by forced convection. *Alexandria Engineering Journal*, *64*, 427-440. https://doi.org/10.1016/j.aej.2022.09.006

[13] Ašonja, A., & Vuković, V. (2018). The potentials of solar energy in the Republic of Serbia: current situation, possibilities and barriers. *Applied Engineering letters*, *3,* 90-97. https://doi.org/10.18485/aeletters.2018.3.3.2

[14] Blagojević, N., Dihovični, Đ., & Krunić, V. (2022) Analysis of the solar plant for the production of hot consumption water for the hospital. *Advanced Engineering letters*, *1,* 23-27. https://doi.org/10.46793/adeletters.2022.1.1.4

[15] European Commission. EU Solar Energy Strategy 2022.

[16] Odeh, S. (2018). Analysis of the Performance Indicators of the PV Power System. *Journal of Power and Energy Engineering, 6*, 59-75. https://doi.org/10.4236/jpee.2018.66005

[17] Chong, Y. B. (2023). Optimal orientation and tilting angle of PV panels considering shading and temperature effects. Lee Kong Chian Faculty of Engineering and Science University Tunku Abdul Rahman.

[18] Rahman, M. M., Hasanuzzaman, M., & Rahim, N. A. (2017). Effects of operational conditions on the energy efficiency of photovoltaic modules operating in Malaysia. *Journal of Clean Production, 164*, 1474-1485. https://doi.org/10.1016/j.jclepro.2016.12.029

[19] Browne, M. C., Norton, B., & McCormack, S. J. (2016). Heat retention of a photovoltaic/thermal collector with PCM. *Solar Energy, 133*, 533-548. https://doi.org/10.1016/j.solener.2016.04.024

[20] Fouad, M. M., Shihata, L. A., & Morgan, E. I. (2017). An integrated review of factors influencing the performance of photovoltaic panels. *Renewable and Sustainable Energy Reviews, 80*, 1499-1511. https://doi.org/10.1016/j.rser.2017.05.141

[21] Nezamisavojbolaghi, M., Davodian, E., Bouich, A., Tlemçani, M., Mesbahi, O., & Janeiro, F. N. (2023). The Impact of Dust Deposition on PV Panels' Efficiency and Mitigation Solutions: Review Article. *Energies, 16*(24), 2-19. https://doi.org/10.3390/en16248022

[22] Gallardo-Saavedra, S., & Karlsson, B. (2018). Simulation, validation and analysis of shading effects on a PV system. *Solar Energy, 170*(6)*,* 828-839. https://doi.org/10.1016/j.solener.2018.06.035

[23] Conceição, R., González-Aguilar, J., Merrouni, A. A., & Romero, M. (2022). Soiling effect in solar energy conversion systems: A review. *Renewable and Sustainable Energy Reviews, 162*. https://doi.org/10.1016/j.rser.2022.112434
[24] Korab, R., Połomski, M., Naczyński, T., & Kandzia, T. (2023). A dynamic thermal model for a photovoltaic module under varying atmospheric conditions. *Energy Conversion and Management, 280* (116773). https://doi.org/10.1016/j.enconman.2023.116773
[25] Shadid, R., Khawaja, Y., Bani-Abdullah, A., Akho-Zahieh, M., & Allahham, A. (2023). Investigation of weather conditions on the output power of various photovoltaic systems. *Renewable Energy, 217*. https://doi.org/10.1016/j.renene.2023.119202
[26] Mustafa, R. J., Gomaa, M. R., Al-Dhaifallah, M., & Hegazy, R. (2020). Environmental Impacts on the Performance of Solar Photovoltaic Systems. *Sustainability, 12*(2). https://doi.org/10.3390/su12020608
[27] Chander, S., Purohit, A., Sharma, A., Nehra, S. P., & Dhakaa, M. S. (2015). Impact of temperature on performance of series and parallel connected mono-crystalline silicon solar cells. *Energy Reports, 1*, 175-180. https://doi.org/10.1016/j.egyr.2015.09.001
[28] Aoun, N. (2022). Methodology for predicting the PV module temperature based on actual and estimated weather data. *Energy Conversion and Management: X; 14*(100182). https://doi.org/10.1016/j.ecmx.2022.100182
[29] Santiago, I., Trillo-Montero, D., Moreno-Garcia, I. M., Pallares-Lopez, V., & Luna-Rodríguez, J. J. (2018). Modeling of photovoltaic cell temperature losses: a review and a practice case in South Spain. *Renewable and Sustainable Energy Reviews, 90*, 70-89. https://doi.org/10.1016/j.rser.2018.03.054
[30] Nordmann, T., & Clavadetscher, L. (2003). Understanding temperature effects on PV system performance. In *3rd World Conference on Photovoltaic Energy Conversion – Proceedings*, Japan, 3, 2243-2246.
[31] Dubey. S., Sarvaiya, J. N., & Seshadri B. (2013) Temperature Dependent Photovoltaic (PV) Efficiency and Its Effect on PV Production in the World - A Review. *Energy Procedia, 33*(311-321). https://doi.org/10.1016/j.egypro.2013.05.072
[32] Ross, R. G. (1976). Interface design considerations for terrestrial solar cells modules. In *Proceedings of the 12th IEEE photovoltaic specialists conference, Vol. 12*, 801-806.
[33] Lasnier, F., & Ang, T. G. (1990). Photovoltaic engineering handbook. New York: Taylor & Francis.
[34] King, D. L., Boyson, W. E., & Kratochvill, J. A. (2004). *Photovoltaic Array Performance Model*. Sandia National Laboratories. https://doi.org/10.2172/919131
[35] Skoplaki, E., Boudouvis, A. G., & Palyvos, J. A. (2008). A simple correlation for the operating temperature of photovoltaic modules of arbitrary mounting. *Solar Energy Materials and Solar Cells, 92*(11), 1393-1402. https://doi.org/10.1016/j.solmat.2008.05.016
[36] Kamuyu, C. L. W., Lim, J. R., Won, C. S., & Ahn, H. K. (2018). Prediction model of photovoltaic module temperature for power performance of floating PVs. *Energies, 11*(2). https://doi.org/10.3390/en11020447
[37] Zouine, M. (2018). Mathematical Models Calculating PV Module Temperature Using Weather Data: Experimental Study. In *Proceedings of the 1st International Conference on Electronic Engineering and Renewable Energy (ICEERE2018)*. Springer Singapore. https://doi.org/10.1007/978-981-13-1405-6_72
[38] Skoplaki, E., & Palyvos, J. A. (2009). On the temperature dependence of photovoltaic module electrical performance: A review of efficiency/power correlations. *Solar Energy, 83*, 614-624. https://doi.org/10.1016/j.solener.2008.10.008
[39] Evans, D. L. (1981). Simplified method for predicting photovoltaic array output. *Solar Energy, 27*, 555-560. https://doi.org/10.1016/0038-092X(81)90051-7
[40] Notton, G., Cristofari, C., Mattei, M., & Poggi, P. (2005). Modelling of a double-glass photovoltaic module using finite differences. *Applied Thermal Engineering, 25*, 2854-2877. https://doi.org/10.1016/j.applthermaleng.2005.02.008
[41] Limane, B., Ould-Lahoucine, C., & Diaf, S. (2023). Modeling and simulation of the thermal behavior and electrical performance of PV modules under different environment and operating conditions. *Renewable Energy, 219*(119420). https://doi.org/10.1016/j.renene.2023.119420
[42] Ahmad, F. A., Sopian, K., Yazdi, M. H., Ruslan, M. H., Ibrahim, A., & Kazem, H. A. (2014). Performance analysis of photovoltaic thermal (PVT) water collectors. *Energy Conversion and Management, 78*, 641-651. https://doi.org/10.1016/j.enconman.2013.11.017

Authors' contacts:

Danijela Kardaš Ančić, Msc, Teaching Assistant
(Corresponding author)
University of Banja Luka, Faculty of Mechanical Engineering,
Stepe Stepanovića 71, 78000 Banja Luka, Bosnia and Herzegovina
danijela.kardas@mf.unibl.org

Mirko Komatina, PhD, Full Professor
University of Belgrade, Faculty of Mechanical Engineering,
Kraljice Marije 16. 11120, Belgrade, Serbia
mkomatina@mas.bg.ac.rs

Ante Čikić, PhD, Full Professor
University North, Department of Mechatronics,
104. brigade 3, 42000 Varaždin, Croatia
acikic@unin.hr

Petar Gvero, PhD, Full Professor
(Corresponding author)
University of Banja Luka, Faculty of Mechanical Engineering,
Stepe Stepanovića 71, 78000 Banja Luka, Bosnia and Herzegovina
petar.gvero@mf.unibl.org

Simon Muhič, PhD, Full Professor
Faculty of Industrial Engineering Novo mesto,
Šegova ulica 112, 8000 Novo mesto, Slovenia
Rudolfovo - Science and Technology Centre Novo mesto,
Podbreznik 15, 8000 Novo mesto, Slovenia
INOVEKS, d.o.o., Cesta 2. Grupe odredov 17,
1295 Ivančna Gorica, Slovenia
simon.muhic@rudolfovo.eu

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250228210232

Original scientific paper
Received: 2025-02-28, Accepted: 2025-03-17

Technical Evaluation Pre-Dehydration Stage under Off-Solar Duration Equipped with Thermal Backup and Cyclic Fan Operation

Nursiwan, Samsul Arifin, Budhi Muliawan Suyitno, Reza Abdu Rahman*

Abstract: This work assesses off-solar duration (OSD) dehydration using three volumetric air flows (VAF) which operate under continued, long and short intermittency cycles. VAF and fan continuity is proportional to total water removed of the product with a maximum capacity of 47.3%. However, it costs higher energy consumption, which only results in maximum effective level of fan by 142.4%. Operating the fan under short intermittency at medium VAF achieves a higher effectivity up to 342.5%. The water diffusion of seaweed is a complex and time-consuming process, making the designed operation favorable to enhance effective energy consumption. Employing thermal backup reaches equilibrium moisture quickly under continuous cycle. However, it has the lowest energy effectiveness (only 6.9%), while setting the target to achieve critical moisture offers higher value at 22.7% using short-intermittency cycle. It shows OSD offers reliable results to reduce the midday cycle duration and ensure minimum losses rather than open dehydration.

Keywords: efficiency; heat storage; intermittency; seaweed; vaporization

1 INTRODUCTION

Population escalated significantly, requiring tremendous amount of food to sustain human life. The process is directed more sustainably to ensure it gives little impact on the environment. It comes as global goal to achieve responsible consumption and production as part of SDGs. The sector is accompanied by rapid development and technological improvement to harvest more clean energy from renewable sources [1–3]. Food processing has become more complex in modern era to accommodate the high variation and global trend, causing high energy consumption in this sector. Specifically, one good example comes from the postharvest handling of aquaculture product (seaweed). The harvested product is handled for dehydration process to extend the lifetime of product and facilitate the next production step. However, dehydration is an extensive energy process due to high latent heat of vaporization for water, approximately 40 kJ/mol [4]. From local perspective, the coastal region from major country producers demands specific technical consideration, which makes industrial-grade drying equipment commonly not used for local farmers [5]. It results in conventional drying by exposing the product directly to solar radiation, which is easier, requires no additional equipment, and takes no additional processing cost.

Direct dehydration during midday supports clean energy utilization and meets the base criteria for sustainable production. Nevertheless, it leads to different problems, such as product degradation due to direct exposure to dirt. Employing direct dehydration gives additional labor cost and highly depends on solar power, followed by quality drops, which reduces the value of the dried product. Consequently, the process becomes economically unfavorable. Solving the issue is performed using closed greenhouse dehydrator (CGD). It remains harvest solar energy, which contributes in energy mix, and simultaneously performs dehydration process in closed room, reducing soiling factor significantly. Labor costs potentially reduce while the product quality increases, making this approach more economically feasible. The difference in relative humidity and temperature (RH/T) inside the CGD and the outdoor potentially causes natural convection, allowing the air to circulate from outside [6]. For higher rate, forced convection is generally used by forcing air to circulate at designed conditions using fluid machine such fan/blower [7]. The energy for operating the machines can be supplied by various sources such wind and solar energy [8–10]. Thus, the system operates primarily by using renewable source to perform dehydration process.

The main technical problem for utilizing CGD is linked specifically to poor dehydration rate compared to industrial-grade dryers. It is highly associated with low solar intensity per unit area [11], making further modification essential to enhance the CGD capacity. Researcher performed several modifications by equipping the system with solar heat collector (SHC). The work [11] used evacuated tube-SHC, which combined with heat energy material (HEM), allowing the supplied air to reach higher temperature (100 °C) with potential to reduce CO_2 and cost per mass flow (< 1 $/kg). The reliability of the system is enhanced by employing latent HEM to improve the storage capacity and thermodynamic aspect [12-14]. In addition, the HEM was modified through encapsulation to prevent leak, which remains shows desirable performance for heating air up to 82 °C [15]. The ideal combination of SHC-CGD promotes the system to operate maximally and reduce the low dehydration rate. Suitable modification of the combined system reached higher extraction rate of water content with lower energy consumption around 0.34 kg/kWh [16]. Many reports demonstrate that the system (CGD with SHC and HEM) has technical feasibility and better performance than non-combined CGD. However, the process uses fan, which consumes energy and requires additional attention regarding the operation behavior of the unit.

Dehydration of the water content from the product depends not only on heat supply but also on mass transfer. The continued operation flow (COF) of the working fan contributes to high specific energy consumption for the process [17]. Thus, modification of the fan operation is done

by adjusting its cycle. For example, M. Li et al. assessed the intermittent operation (IOF) for infrared dryer to dehydrate jujube, showing the proposed operation model potentially improved the quality of the product [18]. The IOF operation is applicable as well for heat pump dehydrator. Y. Hou et al. conducted detail analysis of IOF operation on dehydration of corn kernels, resulting in positive contribution on all IOF setting compared to COF based on the dehydration rate with maximum value of 19.5% [19]. The study [20] utilized solar-electric technology for onion product and compared COF/IOF operation. It showed a positive influence on the efficiency of the process by using IOF at around 35.7%, which also improved the economy annually. Technically, operating the system with IOF offers positive improvement for dehydration process. However, the trend cannot be generalized since many factors involved in dehydration, particularly water diffusivity and interaction, vary from one product to another [21].

Adjusting the drying operation becomes mandatory to maximize the product quality and amount of energy consumed to achieve the target. COF causes continuous dehydration and exposes the product to heat (for thermal dehydration). Nevertheless, water diffusion requires time to move from internal structure to the surface of product [22]. IOF provides additional relaxation time for the product when the fan shuts off, allowing the product to cool down and accelerating water diffusion to reach moisture equalization [23]. Modification of intermittency cycle is also done for osmotic dehydration and demonstrates various effects [24]. The work [25] gave specific examples for the dehydration of olive using various scenarios (freezing, oven and infrared drying), indicating a certain condition to achieve optimum output for IOF, which depends on the method and processed product. It represents the complexity of dehydration process, which varies from case to case, affected by type of product and chosen method, resulting in different dehydration behavior and quality.

The previous research focused on improving the operation of CGD, showing a serious effort to maximize the applicability of the technology for dehydration process. The efforts mainly focused on enhancing solar radiation input to the system under on-solar duration. Nevertheless, weather is unpredictable, and seasonal changes require additional consideration for the sustainability and continuous dehydration process [26]. Also, limited land available in certain regions comes as another challenge, which makes the processed product need to be arranged specifically, including another technical issue such as shading, which limits the effective daily radiation to CGD [27]. Another factor of the processed product (seaweed) is that it has high moisture content (MC) with strong water molecular force to the structure, which demands a longer dehydration duration [28]. Thus, pre-dehydration can be offered to reduce the given issue, which can be performed under off-solar duration (OSD). It makes further work still required to assess the operation of the system under OSD, offering additional information to maximize the pre-dehydration stage with no solar intensity condition, which is also applicable for seasonal changes such as rain season. It is taken to support and promote a better dehydration cycle under conditions with no solar supply and solely depends on air transfer within the system.

2 MATERIALS AND METHOD

We focused the assessment and evaluation for practical application, hoping the findings is applicable for actual implementation and deliver great impact for society. Thus, the basic model of CGD was designed as even span model considering its common model in the field, which has the most optimum solar radiation input annually [29]. The model schematic is presented in Fig. 1a, including the plan dimension and arrangement of the system assembly. The wall was covered with thin plastic (0.2 mm) UV-HDPE to meet the criteria for on-solar duration at midday [30]. The load (seaweed) was hung inside the CGD and positioned at the center point between intake/exhaust lines. The frame was made from hardwood, while the cladding was attached evenly at the frame to ensure well air circulation from outside to the dehydrator chamber. The system was equipped with twin electric fan (DC) at the exhaust port to force air circulation from the outside to pass the load inside the CGD. The given concept is the most common model of basic operation for CGD and highly relevant to typical actual system. Thus, there is strong correlation between the proposed model and the actual design, making the result from this study applicable to the typical CGD system.

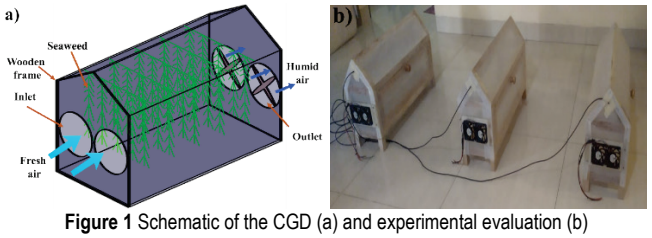


Figure 1 Schematic of the CGD (a) and experimental evaluation (b)

The key parameter for the evaluation is defined according to the mass of the product. It is the most essential point for dehydration process. The initial weight of the load was 0.5 kg (± 4.2%), and it was weighed every hour during the test. The water content (MC) was found at 85% (wet base), showing the typical high MC of aquaculture product. The effective energy is obtained based on the total removed MC from the load. Fig. 1b shows the experimentation unit. It was performed simultaneously for three different parameters. The base variation for evaluation was intake volumetric air flowrate (VAF, $\times 10^{-2}$ m^3/s) which was set at 1.4 (LVAF), 2.2 (MVAF), and 3.1 (HVAF). The fan operation was adjusted based on COF, IOF0.5 (IOF-I) and IOF0.25 (IOF-II). The subscript indicates the cycle schedule of the fan, which was every 30 minutes (IOF-I) and every 15 minutes (IOF-II). For actual comparison, the load was also placed at the outside to determine the baseline of OSD operation without using CGD. The assessment was performed for 10 hours from 07 pm to 05 am (local time) and simultaneously recorded for the RH/T of air inside the CGD and at the ambient.

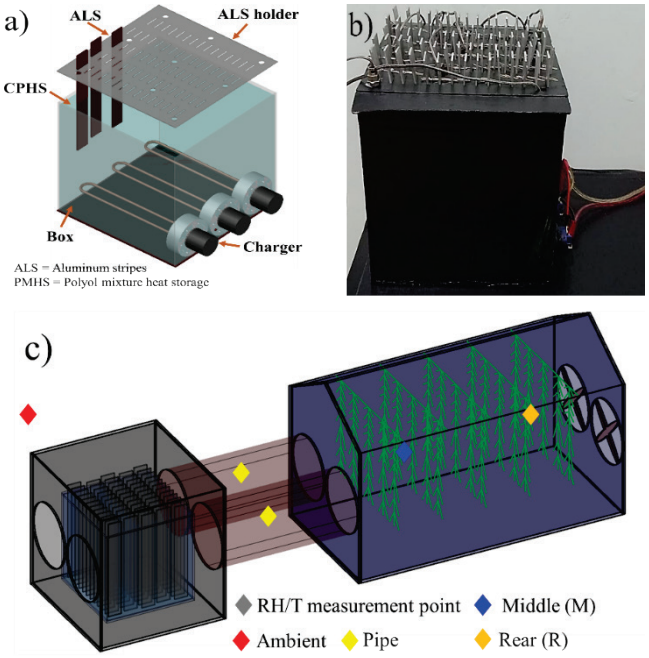


Figure 2 Detail design of thermal box (TBX) (a), its photograph (b) and assembly of system TBX-CGD (c)

The next stage evaluation was conducted according to the best VAF for the system and combined with thermal box (TBX). The TBX is a device that stored thermal energy which commonly absorbs excess heat from sun during optimum solar duration [31–33]. Thus, CGD can be operated by absorbing the stored heat for OSD operation from TBX when its fully charged. The work employed CPHS as material TBX which was made of sorbitol (75%) and mannitol (25%). The polyols offer higher storage density and can operate at elevated temperature, which is suitable for pre-dehydration stage for CGD under OSD operation. Fig. 2a shows the designed TBX for this work. In brief, the model operates using thermal charger from electric power to heat the M-TBX until it reach the designed working temperature. ALS helps accelerate the charge stage by distributing heat properly as the charge process is taken at the lower zone of the assembly. The stored heat was distributed through the body of TBX and the tip of ALS, which directly interacted with the working air before entering CGD. Thus, compact and robust TBX model is obtained, which can be easily developed and implemented for OSD. Fig. 2b presents the photograph of the TBX unit.

The system was evaluated with TBX based on the scheme in Fig. 2c. The fully charged TBX (at 120 °C) was discharged to the OSD through air based on the fan operating scenario. The psychrometric profile for the air was measured at four different regions. The gravimetric capacity of the load was kept the same (0.5 kg). The process provided additional indication of the impact of employing thermal backup based on the designed operation condition for dehydration inside the CGD. Thus, the work assessed two essential OSD operation models, with and without thermal backup, which makes clear interpretation from each operation can be provided to justify the practical condition of CGD operation under OSD.

3 RESULTS AND DISCUSSION

The cyclic operation pattern from the evaluation is observed according to the change of air properties, in this case RH. The pattern is observed since the data was taken every 5 seconds. The obtained data is presented as an arbitrary unit graph (Fig. 3) to maximize the observation of cyclic patterns. The profile demonstrates a clear cyclic pattern as well as the effect of VAF from the process. Under LVAF operation (Fig. 3a), the RH changes inconsistently for the IOF-I and IOF-II. The vast initial water content followed by low VAF results in steep fluctuation of RH value due to high absorption of water molecule by the air. COF shows opposite profile from ambient, which signifies the continuous air evacuation process. It was raining outside the building during the evaluation, which caused the RH ambient to drop dramatically. It is affected by the movement of moist air in the atmosphere, causing the ambient humidity to decrease [34]. The changing pattern for IOF operation demonstrates delayed profile once the fan shuts off, causing the RH to increase slightly before reaching the maximum point when fan sets on.

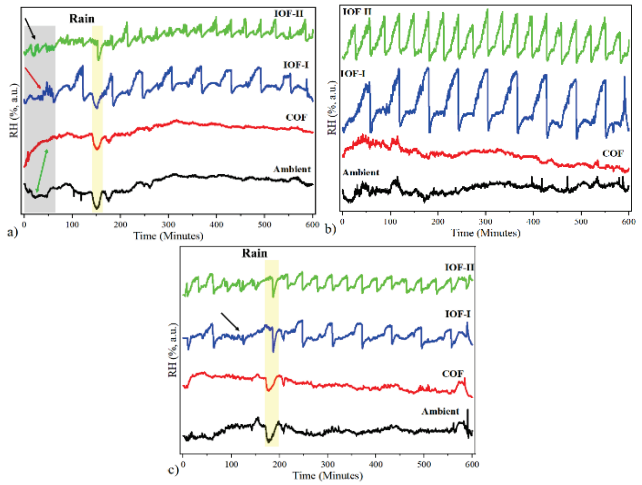


Figure 3 Variation of RH (%) at LVAF (a), MVAF (b) and HVAF (c)

The evaluation of MVAF was conducted on clear night without rain outside the building. Interestingly, the delayed RH profile becomes more apparent at the transition fan operation for IOF-II (Fig. 3b). Higher VAF probably improves the average evacuation rate of moist air, making a higher RH gradient seen distinctly. However, adjusting suitable VAF relative to the volume of CGD is important to consider since the RH value becomes lower at the active cycle (Fig. 3c). Suitable cyclic pattern for MVAF indicates steady operation on the exchange process between the air and water removal from the product. The presence of resting time causes various diffusion process throughout the structure body of the product. Increasing the air velocity may result in slower evaporation, as observed with insignificant RH increment at the initial phase of HVAF operation (black arrow Fig. 3c). The cyclic pattern tends to occur after certain period, showing the time as essential variable for evaporation process of water content from the product. The arbitrary graph allows for a better observation to analyze the impact of each evaluated parameter of the assessment process.

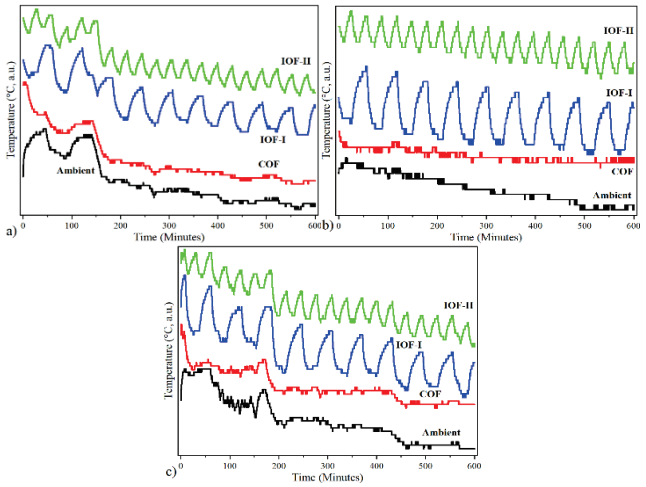


Figure 4 Variation of T (°C) at LVAF (a), MVAF (b) and HVAF (c)

Another psychrometric aspect of air (T, °C) is plotted in Fig. 4. The profile has identical pattern where it tends to decrease as the process continues throughout the night. It implies that evaporative cooling [35] takes place inside the dehydrator. It proves the water removal process from the product to the air. The impact of cyclic process is observed based on T indicator of the process, which is less observed for LVAF operation (Fig. 4a). In contrast, operating the system at MVAF allows to achieve of stationary temperature between active-resting period (Fig. 4b). It causes low-temperature decrement which observed insignificantly between each period. The HVAF operation indicates partial decrement on T at the transition active-resting period (Fig. 4c). It implies the zone experiences further evaporation under free convection mode. It emphasizes there is strong factor which comes from the processed product regarding the water movement inside the cellular structure of product, resulting in different evaporation behavior inside the CGD, particularly under operating without thermal source such as in OSD.

Fig. 5 summarizes the psychrometric value from the evaluation. The presented scale clearly represents the changes in psychrometric aspect (RH/T) throughout the process under different fan operations and speeds. The ambient RH is generally lower than inside the CGD, allowing for possible evaporation process inside CGD under OSD (Fig. 5a). It is justified based on the tendency of higher RH value inside CGD demonstrating water liberation from product to the working fluid, which causes the air becomes more humid. One interesting result is shown for IOF-I profile, which increases by 87.1% (LVAF) and triple times for MVAF. It decreases once the system operates under HVAF. Accordingly, intermittency operation and volumetric speed strongly correlate with the changes in RH value. Moreover, shorter cyclic with IOF-II causes the RH limit to becomes lower than IOF-I and COF (only for HVAF). The changes in RH are simultaneously followed by temperature drop inside CGD (Fig. 5b). However, the temperature variation for all process with different parameter demonstrate identic value, ranging between 0.7 – 3.4 °C. The insignificant temperature variation shows the technical limitation for operating dehydration under OSD, which limits the evaporation cooling rate based on the designed dimension and VAF.

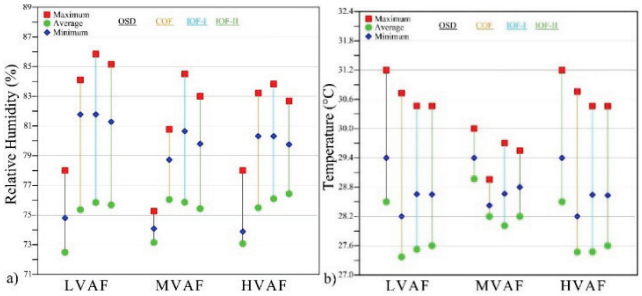


Figure 5 Scaling the RH-T variation of the operation: a) RH map and b) T Map

The changes in RH/T inside CGD indicate the water liberation from product, causing the product mass to decrease along the process. The total mass loss is presented in Fig. 6a and taken as key parameter to compare each evaluation scenario. At this point, locating the product outside (defined as OSD) still causes dehydration since there is air movement due to wind (even though insignificantly), variation of RH and T of the ambient air, and method to place the product (hanging) which increases the contact area with air and accelerate gravitational effect. Unfortunately, it results in insignificant dehydration rate since the process occurs as free convection without involving any thermal source. In contrast, employing fan to accommodate force convection inside CGD contributes to the higher total mass loss. It is also observed that increasing the VAF and continue operation yield higher total mass loss around 47.3%. The impact of changing the intermittency is observed distinctly as the VAF increases, indicating that evaporation can be accelerated as the average mass transfer improves within the system.

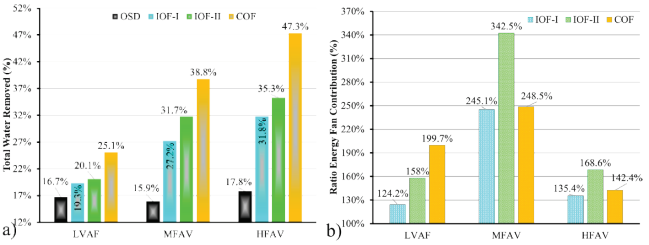


Figure 6 Total mass loss (a) and ratio useful energy from fan (b)

The primary goal for operating the system at OSD is pre-dehydration stage, making the consumed energy the most crucial factor to consider. Each process (Fig. 6b) shows different results, which implies that additional points need to be considered regarding the effectiveness of OSD operations. The amount of liberated water from product under CGD operation is subtracted by total water removal from OSD (baseline). The net liberated water is multiplied by standard latent heat of vaporization and compared by the energy consumed by the component (fans). It is plotted as ratio of energy fan contribution. As observed, it has positive value (higher than 100%), which indicates that the total water removed is higher than energy consumed by the fan.

The LVAF requires suitable fan operating, resulting in higher energy effectiveness for continuous cycle and shorter intermittency. However, dehydration takes time for the water molecule to diffuse throughout the product, making continuity and increasing the fan speed not always proportional in terms of effective energy consumption. It is shown notably for the MVAF, where IOF-I (30 minutes intermittency) and COF (continue) have identic ratio. In contrast, IOF-II (15 minutes intermittency) allows the system to achieve the highest effectiveness from the energy input to operate fans. Increasing the operation speed (HVAF) decreases the effective energy input since the fan consumes more energy. The major finding of this work is that operational parameters are essential for optimum OSD operation and can be modified according to the system and product conditions to achieve the highest effective energy ratio throughout the process.

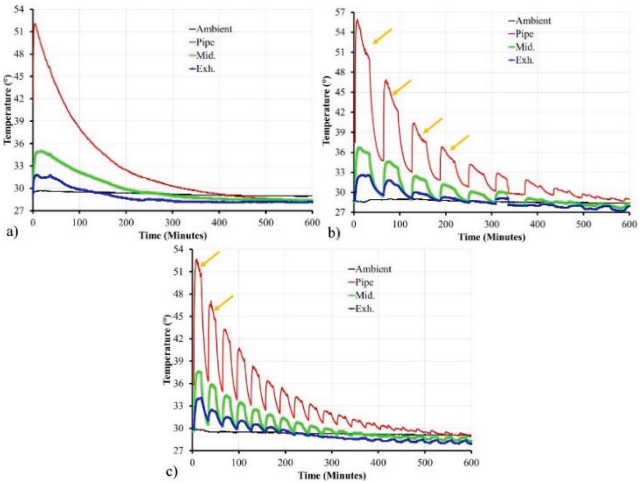


Figure 7 Changing of temperature of the assembly TBX-CGD for a) COF, b) IOF-I, and c) IOF-II

The system is tested by combining TBX as thermal backup for CGD under speed MVAF. The observed RH/T value for each location of the assembly (Fig. 2c) is presented in Fig. 7. All zone shows indications for temperature increment as a result of the addition TBX, where the value becomes lower as the measurement farther from the heat source. The effect of operation setting for the CGD changes the temperature pattern. For example, COF shows consistent temperature decrement from the pipe to the inside CGD (Fig. 7a). The temperature initially increases by 75.4% at the pipe zone and steadily decreases to 7% at the exhaust region. The TBX has a finite capacity, making the effective operation only 400 minutes according to the plotted temperature.

Fig. 7b presents the cyclic pattern of IOF-I, allowing the discharge stage to be extended until the final evaluation. It achieves a higher increment on the temperature at 93.1%, with the lowest point at 12.4%. A short delay region (brown arrow) for the discharged heat indicates that TBX takes time to perform heat exchange with air. It is probably advantageous for dehydration to prevent the product from overheating due to sudden high-temperature exposure. Shorter-cyclic (IOF-II, Fig. 7c) shows a relatively similar pattern with IOF-II. Even though the maximum increment is lesser than IOF-II (only 76.3%), the exhaust zone temperature maintained a higher value at 34.1 °C.

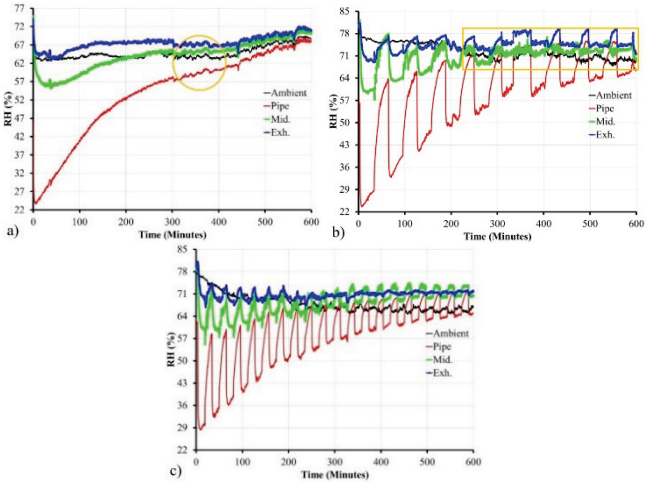


Figure 8 Changing of RH of the assembly TBX-CGD for a) COF, b) IOF-I, and c) IOF-II

TBX transfers heat to the air, causing it to expand and resulting in a lower RH. As the cycle continues and the heat diminishes (brown circle Fig. 8a), the air contracts, leading to a higher RH inside CGD. In contrast with T, RH changes rapidly as the system operates, causing overlap phenomenon and constantly changing the RH value relatively at higher/lower conditions than the ambient. IOF-I shows this condition notably (brown rectangular Fig. 8b), while IOF-II indicates a lower cyclic effect regarding the interlapping RH value (Fig. 8c).

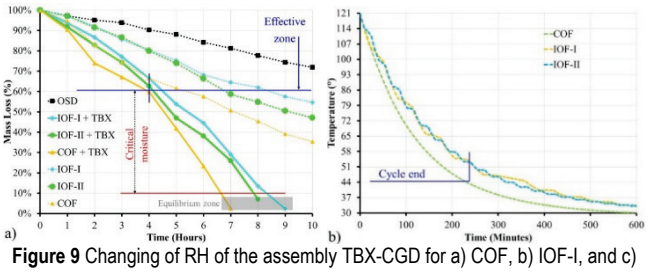


Figure 9 Changing of RH of the assembly TBX-CGD for a) COF, b) IOF-I, and c) IOF-II

The role of dehydration is presented as the decrement of mass of the product. It is plotted in Fig. 9a for the system with and without TBX for a clearer evaluation. The mass decreases higher for COF and IOF-II, indicating the key role of forced air transfer in performing dehydration at night-time. Adding thermal backup, in this case TBX, accelerates the evaporation, eventually increasing the mass loss to achieve the equilibrium moisture. However, the presence of TBX is not intended to perform direct drying to achieve equilibrium moisture due to disproportional pair-to-pair gravimetry capacity from the TBX and the load. For illustration, the MTBX was 10 kg while the load was merely 0.5 kg. It is totally imbalanced from a practical perspective since the storage has more capacity than the designed load. The key role of TBX is essential to achieve critical moisture, which occurs at mass 60-70%. One undeniable proof is that all

operations show the same duration to achieve critical moisture, indicating that the TBX is more advantageous in accelerating the initial dehydration phase. Moreover, the temperature inside TBX remains higher at capacity of 15% (according to minimum/maximum temperature deviation inside the TBX) when it comes to the end of cycle of pre-dehydration stage (Fig. 9b). It is favorable for TBX since it reduces initial charge stage which typically requires more energy under solid-sensible stage of the material inside TBX.

4 CONCLUSION

The work shows high feasibility and opportunity to utilize OSD to perform pre-dehydration stage, which probably increases the batch capacity per day and prevents severe loss during low-intensity solar season. Technical barrier such as low dehydration performance is solved by adjusting the system to operate sufficiently, reducing the net energy consumption while maintaining the reasonable total water removed. Indeed, operating the fan at higher evacuation rate with continuous active cycle results in the maximum total dehydration. Unfortunately, this act has low positive energy coefficient, which is merely 142.4%. Also, operating the system with a relatively long cyclic setting causes an extended tempering period, reducing the maximum water removed from product. At this point, shorter cyclic with suitable evacuation rate contributes to a positive achievement in terms of effective energy ratio for operating fan, which increased to 342.5%. The absence of heat input for OSD is solved by using thermal backup, allowing the system to reach the critical moisture faster. It is found that all scenarios show identic duration for the product to reach critical moisture level. Operating the backup heating energy under suitable control allows the system to achieve effectivity level of 22.7%. One critical finding shows that forcing the system to achieve equilibrium moisture (dry) is technically inefficient, reaching maximum point effectivity level at 9.3%.

Overall, the finding reveals that OSD is applicable to support the development and capacity of typical solar drying. It is only achieved by adjusting the system to operate with suitable cyclic duration and evacuation rate, even though the MC of product is considerably high (above 80%). This initial evaluation promotes OSD for optimizing the drying process, making additional work advisable to achieve a higher technology output. It is important to conduct continuous cycle from effective solar duration and assess the entire energy consumption of the process, including critical factors involving the operation and feasibility, such as techno-economy evaluation. We believe the finding from this work, which can be further improved, offers a clear breakthrough to encourage the development of clean food production, supporting responsible consumption and enhancing the affordable clean energy goal as part of community service from researcher and industrialist.

5 REFERENCES

[1] Vallejo-Díaz, A., Herrera-Moya, I., Fernández-Bonilla, A., & Pereyra-Mariñez, C. (2022). Wind energy potential assessment of selected locations at two major cities in the Dominican Republic, toward energy matrix decarbonization, with resilience approach. *Thermal Science and Engineering Progress, 32*, 101313. https://doi.org/10.1016/j.tsep.2022.101313

[2] Lahnaoui, A., Aniello, G., Barberis, S., Nwabueze Agbo, S., & Kuckshinrichs, W. (2024). Assessing the feasibility of a green hydrogen economy in selected African regions with composite indicators. *International Journal of Hydrogen Energy, 101*, 1387-1405. https://doi.org/10.1016/j.ijhydene.2024.12.233

[3] Assareh, E., Agarwal, N., Paul, M. C., Ahmadi, P., Ghodrat, M., & Lee, M. (2023). Investigation and development of a novel solar-biomass integrated energy system for clean electricity and liquid hydrogen production. *Thermal Science and Engineering Progress, 42*, 101925. https://doi.org/10.1016/j.tsep.2023.101925

[4] Wang, J., Kong, Y., Liu, Z., & Wang, H. (2023). Solar-driven interfacial evaporation: Design and application progress of structural evaporators and functional distillers. *Nano Energy, 108*, 108115. https://doi.org/10.1016/j.nanoen.2022.108115

[5] Muflikh, Y. N., Permani, R., Nuryartono, N., Pasaribu, S. H., Julianto, B. S., Sjahruddin, F., et al. (2024). Building stakeholders' mutual understanding of seaweed sustainability in Indonesia: A group model building approach. *Marine Policy, 168*, 106283. https://doi.org/10.1016/j.marpol.2024.106283

[6] Chaatouf, D., Raillani, B., Salhi, M., Amraqui, S., & Mezrhab, A. (2023). Experimental and numerical study of a natural convection indirect solar dryer with PCM tubes: Dynamic, thermal and nutritional quality analysis. *Solar Energy, 264*, 111975. https://doi.org/10.1016/j.solener.2023.111975

[7] Demissie, Y. A., Abreham, R. E., Wassie, H. M., & Getie, M. Z. (2024). Advancements in solar greenhouse dryers for crop drying. *Energy Reports, 11*, 5046-5058. https://doi.org/10.1016/j.egyr.2024.04.058

[8] Atalay, H., & Aslan, V. (2023). Advanced exergoeconomic and exergy performance assessments of a wind and solar energy powered hybrid dryer. *Renewable Energy, 209*, 218-230. https://doi.org/10.1016/j.renene.2023.03.137

[9] Atalay, H., Yavaş, N., & Turhan Çoban, M. (2022). Sustainability and performance analysis of a solar and wind energy assisted hybrid dryer. *Renewable Energy, 187*, 1173-1183. https://doi.org/10.1016/j.renene.2022.02.020

[10] Khalil, A., Khaira, A. M., Abu-Shanab, R. H., & Abdelgaied, M. (2023). A comprehensive review of advanced hybrid technologies that improve the performance of solar dryers: Photovoltaic/thermal panels, solar collectors, energy storage materials, biomass, and desalination units. *Solar Energy, 253*, 154-174. https://doi.org/10.1016/j.solener.2023.02.032

[11] Tiwari, A., & Kumar, A. (2023). Solar air heater based on double-ends open evacuated tube with and without phase change material: An experimental investigation. J*ournal of Energy Storage, 72*, 108265. https://doi.org/10.1016/j.est.2023.108265

[12] Anish, R., Mariappan, V., Joybari, M. M., & Abdulateef, A. M. (2020). Performance comparison of the thermal behavior of xylitol and erythritol in a double spiral coil latent heat storage system. *Thermal Science and Engineering Progress, 15*, 100441. https://doi.org/10.1016/j.tsep.2019.100441

[13] Rahman, R. A., Sulistyo, S., Utomo, M. S. K. T. S., & Dhewangga Putra, R. D. (2024). The optional approach in recycling plastic waste for energy storage application: A detailed evaluation of stabilized–hexadecanoic acid/plastic. *Case Studies in Chemical and Environmental Engineering, 9*, 0-5. https://doi.org/10.1016/j.cscee.2024.100751

[14] Murugesan, G. K., Murugesan, C., Sakthivel, M., & Tamilkolundu, S. (2024). Experimental investigation on a solar dryer assisted with minimum phase change material (PCM) placed on the inner walls of drying chamber. *Journal of Energy Storage, 98*. https://doi.org/10.1016/j.est.2024.113069

[15] Palacio, M., Ramírez, C., Carmona, M., & Cortés, C. (2022). Effect of phase-change materials in the performance of a solar air heater. *Solar Energy, 247*, 385-396. https://doi.org/10.1016/j.solener.2022.10.046

[16] Khanlari, A., Sözen, A., Şirin, C., Tuncer, A. D., & Gungor, A. (2020). Performance enhancement of a greenhouse dryer: Analysis of a cost-effective alternative solar air heater. *Journal of Cleaner Production, 251*. https://doi.org/10.1016/j.jclepro.2019.119672

[17] Carlos, J., Pereira, A., Pereira, W., Gomes, J. P., Jos, A., Melo, D., et al. (2020). Continuous and intermittent drying of rough rice: Effects on process effective time and effective mass diffusivity. *Agriculture (Switzerland), 10*, 0-5.

[18] Li, M., Geng, Z., Zhang, X., Yang, X., & Zhang, Q. (2025). Establishment and application of a simulation model for infrared radiation combined with hot air intermittent drying of jujube slices. *LWT, 215*, 117280. https://doi.org/10.1016/j.lwt.2024.117280

[19] Hou, Y., Wu, W., Wang, H., Zhang, H., & Guan, C. (2025). Effect of intermittent ratio on the performance of a closed-loop heat pump drying system with a novel ternary refrigerant and on the drying characteristics of corn kernels. *International Communications in Heat and Mass Transfer, 161*, 108541. https://doi.org/10.1016/j.icheatmasstransfer.2024.108541

[20] Hadibi, T., Mengjie, S., Jiangtao, R., Wang, Y., Benseddik, A., Ming, L., et al. (2025). Comprehensive analysis and experimental study on the novel intermittent solar-electric onion drying technology. *Renewable Energy, 238*, 121949. https://doi.org/10.1016/j.renene.2024.121949

[21] Van Engeland, C., Spreutels, L., Legros, R., & Haut, B. (2022). Comprehensive analysis of intermittent drying: A theoretical approach. *Food and Bioproducts Processing, 131*, 86-101. https://doi.org/10.1016/j.fbp.2021.10.009

[22] González-Pérez, J. E., Romo-Hernández, A., López-Malo, A., & Ramírez-Corona, N. (2023). Evaluation of osmodehydration and vacuum-assisted osmodehydration as pre-treatments during fruit drying process: The effect on drying rates, effective water diffusion and changes in product quality. *Journal of Engineering Research (Kuwait), 11*, 275-282. https://doi.org/10.1016/j.jer.2023.100153

[23] Bissaro, C. A., de Souza Matias, G., Defendi, R. O., & de Matos Jorge, L. M. (2022). Modeling the drying kinetics of soybeans under intermittent operation in thin layer. *Food and Bioproducts Processing, 136*, 226-235. https://doi.org/10.1016/j.fbp.2022.10.006

[24] Tsopwo Zena, C., & Jiokap Nono, Y. (2024). Investigating intermittent immersion during osmotic dehydration of mango (Mangifera indica L. Moench). Part A: Determination of optimal conditions for mango (Mangifera indica L. Moench) dehydration impregnation by immersion (D2I) and intermittent immersion. *Heliyon, 10*, e39389. https://doi.org/10.1016/j.heliyon.2024.e35808

[25] Pyrka, I., Mantzouridou, F. T., & Nenadis, N. (2023). Optimization of olive leaves' thin layer, intermittent near-infrared drying. *Innovative Food Science and Emerging Technologies, 84*, 103264. https://doi.org/10.1016/j.ifset.2022.103264

[26] Pandey, S., Kumar, A., & Sharma, A. (2024). Sustainable solar drying: Recent advances in materials, innovative designs, mathematical modeling, and energy storage solutions. *Energy, 308*, 132725. https://doi.org/10.1016/j.energy.2024.132725

[27] Chong, S., You, J., Wu, J., & Chang, I. S. (2024). Booming solar energy drives land value enhancement: Evidence from 648 photovoltaic projects in China. *Journal of Cleaner Production, 484*, 144270. https://doi.org/10.1016/j.jclepro.2024.144270

[28] Amir, N., Imron Mustajib, M., Gozan, M., & Chan, C. (2024). Development of a novel tray solar dryer for aquaculture product: Experimental study on drying kinetics and product quality in Eucheuma cottonii seaweed. *Solar Energy, 273*, 112503. https://doi.org/10.1016/j.solener.2024.112503

[29] Paing, S. T., & Anderson, T. N. (2024). Characterizing the effects of geometry on the natural convection heat transfer in closed even-span gable-roof greenhouses. *Thermal Science and Engineering Progress, 48*, 102408. https://doi.org/10.1016/j.tsep.2024.102408

[30] Srinivasan, G., & Muthukumar, P. (2021). A review on solar greenhouse dryer: Design, thermal modelling, energy, economic and environmental aspects. *Solar Energy, 229*, 3-21. https://doi.org/10.1016/j.solener.2021.04.058

[31] Ranjan Tamuli, B., & Nath, S. (2023). Analysis of micro heat pipe array based evacuated tube solar water heater integrated with an energy storage system for improved thermal performance. *Thermal Science and Engineering Progress, 41*, 101801. https://doi.org/10.1016/j.tsep.2023.101801

[32] Rahman, R. A., Sulistyo, S., Utomo, M. S. K. T. S., Ragil, D., & Suyitno, B. M. (2024). Experimental evaluation on the power characteristic of direct-photovoltaic charging for thermal storage equipment. *Mechanical Engineering for Society and Industry, 4*, 115-122.

[33] NematpourKeshteli, A., Iasiello, M., Langella, G., & Bianco, N. (2024). Using metal foam and nanoparticle additives with different fin shapes for PCM-based thermal storage in flat plate solar collectors. *Thermal Science and Engineering Progress, 52*, 102690. https://doi.org/10.1016/j.tsep.2024.102690

[34] Acero, J. A., Kestel, P., Dang, H. T., & Norford, L. K. (2024). Impact of rainfall on air temperature, humidity, and thermal comfort in tropical urban parks. *Urban Climate, 56*, 102051. https://doi.org/10.1016/j.uclim.2024.102051

[35] Lv, G., Ji, X., Yang, B., Chen, Y., & Xu, H. (2024). Operating modes study of evaporative cooling-dehumidification air conditioning under different climatic conditions. *Applied Thermal Engineering, 243*, 122597. https://doi.org/10.1016/j.applthermaleng.2024.122597

Authors' contacts:

Nursiwan, Master Student
Department of Mechanical Engineering,
Faculty of Engineering, Universitas Pancasila,
Srengseng Sawah, Jagakarsa 12640, DKI Jakarta, Indonesia
E-mail: nursiwanalif@yahoo.com

Samsul Arifin, Master Student
Department of Mechanical Engineering,
Faculty of Engineering, Universitas Pancasila,
Srengseng Sawah, Jagakarsa 12640, DKI Jakarta, Indonesia
E-mail: abngsamsul1400@gmail.com

Budhi Muliawan Suyitno, Full professor, Dr.
Department of Mechanical Engineering,
Faculty of Engineering, Universitas Pancasila,
Srengseng Sawah, Jagakarsa 12640, DKI Jakarta, Indonesia
E-mail: budhi.suyitno@univpancasila.ac.id

Reza Abdu Rahman, Assistant Professor, Dr.
(Corresponding author)
Department of Mechanical Engineering,
Faculty of Engineering, Universitas Pancasila,
Srengseng Sawah, Jagakarsa 12640, DKI Jakarta, Indonesia
E-mail: reza.a@univpancasila.ac.id

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250322010406

Original scientific paper
Received: 2025-03-22, Accepted: 2025-09-23

Resting EEG Biomarkers: Investigating Energy Asymmetry in Brain Hemispheres for Autism Diagnosis across Gender

Bashar S. Falih*, Mohannad K Sabir, Zainab H. Qadoori, Ali H. Ali, Noor H. Al-Janabi, Ahmet Aydın

Abstract: Autism spectrum disorder (ASD) is characterized by atypical brain activity, yet early and accurate diagnosis remains challenging. This study proposes a novel EEG-based method to detect ASD by analyzing hemispheric energy asymmetry during the resting state. EEG data were recorded from 44 children (22 ASD, 22 controls) at medical city, Baghdad. Signals were pre-processed using decimation, band-pass filtering, artifact subspace reconstruction, detrending, and normalization. A sliding window (8 s, 12.5% overlap) extracted maximum theta wave energy from each hemisphere. Results revealed greater right-hemispheric energy in ASD children, particularly enhanced in the Theta band. The method achieved 100% ASD detection accuracy for both genders, while control classification yielded 77% accuracy for males and 55.6% for females, resulting in an overall control accuracy of 68.1%. Total accuracy across groups was 84% (88.5% males, 77.8% females). These findings highlight right-hemispheric dominance in ASD and the need for gender-specific considerations in EEG-based diagnosis.

Keywords: Autism Diagnosis; Brain Hemispheres; Brain Hemisphere Imbalance; Diagnostic Biomarkers; Electroencephalography (EEG); Gender Differences

1 INTRODUCTION

Autism Spectrum Disorder (ASD) is a complex neurodevelopmental condition influenced by genetic, epigenetic, environmental, and immune factors. It is primarily diagnosed in children and is characterized by difficulties in communication, social interaction, and restricted or repetitive behaviors. The symptoms vary widely among individuals, making ASD a highly heterogeneous disorder [1].

Genetic susceptibility plays a significant role in ASD, with families having a history of autism showing higher occurrence rates. Genetic mutations linked to brain development increase the risk. ASD individuals often exhibit neurochemical and neuroanatomical abnormalities, including atypical brain activity and accelerated early childhood brain growth. Environmental factors like prenatal infections, air pollution, and preterm birth may contribute to ASD risk but typically interact with genetic predispositions rather than acting alone [2].

Given the multidimensionality of ASD, researchers emphasize that development of the disorder results from a combination of genetic, biological, and environmental factors. Due to the multiplicity of causative factors, it is easy to attribute no single cause to the disorder. Identification of such causative factors makes it easier to advance in diagnosis and treatment at an early stage, consequently enhancing the quality of life of people with ASD as well as integration into society [1].

The prevalence of ASD has been steadily increasing around the world. In America, the rate had risen from 1 in 68 children in 2014 to 1 in 36 children in 2018 [3]. Further, ASD is diagnosed roughly 3.9 times more frequently in boys than in females [3]. The World Health Organization (WHO) found that 1 in 100 children in the world has ASD [4]. Similarly, Autism Spectrum Australia (ASA) also noted an increase in the incidence of autism from 1 in 100 in 2014 to 1 in 70 in 2018, a rise of 40% in cases. By 2024, this figure reached 1 in 40, indicating that at least 1 in 40 individuals are on the autism spectrum, a significant increase from the previous estimate of 1 in 70 in 2018 [5].

In South Korea, rates of ASD cases registered have also increased exponentially. They numbered 1,514 as of the year 2000 but rose to a modest 42,700 and above by 2023—a 2,384% increase in just twenty years [6]. Fig.1 illustrates the steady annual rise in ASD cases in South Korea from 2000 to 2023. The graph demonstrates an acceleration pattern of acceleration with the annual growth rate increasing from approximately 1.01 thousand new cases annually in the early 2000s to approximately 2.85 thousand new cases annually in 2016.

This trend suggests an exponential increase in ASD prevalence, emphasizing the urgent need for prevention and intervention strategies. With the assumption that the current trend continues, it is projected that ASD cases in South Korea would be approximately 73.55 thousand in 2030. Based on the current annual increment rate of 2.85 thousand cases, this estimate emphasizes the necessity of understanding the reasons for this rapid increase.

Despite the increasing trend, ASD prevalence is not consistent across all countries, with some countries reporting considerably higher rates compared to others. Differences in diagnostic criteria, awareness, and accessibility of healthcare might explain these variations. It is important to understand the variations in order to formulate effective global measures to manage and alleviate the burden of ASD.

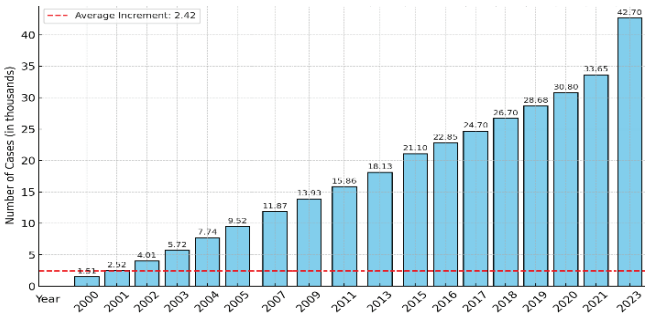


Figure 1 Number of people with ASD in South Korea from 2000 to 2023

Study on infant siblings at high risk of autism shows that some behavioral traits typical of autism start to emerge in the later stages of the 1st and 2nd year of life [7, 8]. ASD severely disables a person from performing daily activities, interacting socially, and succeeding in school and work environments. In addition, people with ASD are constantly exposed to discrimination, abuse, and infringement of their human rights all over the world, as reported by the WHO [4].

There is no cure available for ASD to date, yet early intervention is reported to produce a positive influence on brain growth, improving learning, social, and communication functions. With increased cases of ASD across the world, it is imperative to devise an accurate, effective, and highly efficient diagnosis system

The rising prevalence of autism is a big challenge to human health and the quality of education. As the number of ASD cases increases, the most potent way to improve outcomes is early detection and intervention. Therefore, high priority must be accorded to early identification to ensure timely intervention and treatment of individuals with ASD.

Early ASD detection plays a crucial role in mitigating the challenges associated with autism. The following objectives outline the importance of timely diagnosis:

- Improving Developmental Gains: Early diagnosis of ASD allows children to receive interventions during the critical period of brain development (before age 3). This helps improve social, communication, and cognitive development, enabling greater independence as they grow [9].
- Reducing Long-term Costs and Burdens: Early intervention minimizes the need for costly therapies, reducing the financial and emotional burden on families and the healthcare system. Children diagnosed early are more likely to succeed in regular educational settings, reducing the need for specialized programs.
- Facilitating Social and Communication Development: Early diagnosis enhances communication and socialization skills, helping children build essential interpersonal abilities. This promotes better academic engagement and integration into society [9].
- Personalized and Efficient Treatment Strategies: Early identification allows tailored interventions involving pediatricians, therapists, and educators. This individualized approach enhances developmental outcomes and addresses areas of difficulty effectively [9, 10].

2 RELATED WORKS

With the advancements in technology and computing, artificial intelligence (AI) has achieved numerous successes across various fields. Additionally, AI has been employed for the early detection of autism in several studies [11, 12]. However, studying and analyzing the energy distribution in both hemispheres of the brain for different genders is of great importance, as it helps highlight differences and draw valuable conclusions that should be considered in future studies.

However, in this study, a quantitative analysis of brain wave energy distribution between the right and left hemispheres, as well as gender-based differences in control and ASD children, is emphasized. This approach establishes a robust foundation for understanding neurophysiological divergences between these groups.

A study by [13] explores gender differences in brain function and structure, highlighting variations in cognitive processing, emotional regulation, and neurological susceptibility. These differences are attributed to genetic, hormonal, and environmental influences rather than inherent superiority. Understanding these distinctions aids in developing gender-specific strategies for treating neurological disorders.

Research by [14] investigates sex differences in brain-behavior correlations during early childhood using resting-state fMRI. Findings indicate that significant functional connectivity differences exist between males and females, particularly in the prefrontal cortex and temporal pole, influencing cognitive and emotional development at later stages.

A study by [15], examined how the native language influences brain connectivity. Using MRI and diffusion-weighted imaging, they compared brain scans among 94 native speakers of Arabic and German. The findings indicated more connectivity between the left and right hemispheres in Arabic speakers due to the complex semantic and phonological processing of Arabic. Conversely, the German speakers exhibited more left-hemisphere integration, possibly because of German's syntactically complex nature. The study illustrates how early language exposure affects brain structure and cognitive processing.

The study in [16] examines neural abnormalities in vocal pitch perception in children with ASD through magnetoencephalography (MEG). Results reveal atypical pitch processing in ASD, particularly in the left hemisphere, suggesting abnormal neural mechanisms that affect auditory perception.

Similarly using MEG, [17] investigates auditory processing deficits in children with specific language impairment (SLI). Findings indicate an enhanced left-hemisphere auditory response in SLI children, deviating from the typical right-hemisphere dominance, suggesting a compensatory neural mechanism for delayed language development.

In [17], the neural basis of phonological deficits in developmental dyslexia is examined through MEG. Results show atypical left-hemispheric specialization in at-risk children, affecting phonological processing and efficient word learning.

Language lateralization development is explored in [18] using fMRI, revealing strong right-hemisphere activation in young children, which diminishes with age. This suggests a compensatory mechanism for language recovery following left-hemisphere injury.

Authors in [19] examine hemispheric dominance in bimanual movements, supporting left-hemisphere control in symmetric movements. However, asymmetric movements do not show definitive hemispheric dominance, indicating task-dependent lateralization.

The classical left-brain/right-brain theory introduced by [20] posits that each hemisphere governs different cognitive styles. Though outdated, this theory provides a foundation for

understanding hemispheric specialization in analytical and creative tasks.

Hemispheric attention differences are investigated in [21], suggesting the right hemisphere processes both contralateral and ipsilateral stimuli, whereas the left hemisphere primarily processes contralateral stimuli. EEG findings support right-hemisphere dominance for attention.

The development of functional brain asymmetry in childhood is further explored in [22] through cerebral blood flow studies. Initial right-hemisphere dominance in early childhood shifts to left-hemisphere dominance by age three, aligning with the evolution of visuospatial and language functions.

The right hemisphere's role in emotional processing is analyzed in [23] through behavioral and physiological measures. Results indicate strong physiological responses to emotional stimuli in the right hemisphere, suggesting its primary contribution to emotional regulation rather than cognitive processing.

These studies collectively indicate the presence of significant hemispheric and gender differences in brain function, supporting the hypothesis that ASD involves distinct patterns of brain wave energy distribution. The current study builds upon this foundation by quantitatively comparing energy differences between hemispheres in autistic and control children, further delineating the neurophysiological distinctions between these groups.

3 DATASET DESCRIPTION

The study, conducted from March 5, 2021, to July 11, 2023, data collected from the Children's Welfare Teaching Hospital at Medical City, Baghdad, Iraq.

EEG signals were recorded using the 10/20 international electrode system [24], employing a 19-channel scheme with a sampling frequency of 0.5 kHz, ensuring channel resistance remained below 10 kΩ. All EEG data consisted of real signals to obtain accurate and practical results. The EEG machine used in this study was manufactured by Nihon Kohden Company, Japan. Fig. 2 illustrates electrode connections using 19 channels based on the 10/20 international electrode placement system.

Subjects underwent EEG recording during induced sleep, with the chloral hydrate group receiving specified doses of the drug. The study included 44 children: 22 diagnosed with ASD (mean age: 3.9 ± 1.7 years) and 22 control children (mean age: 4.8 ± 3.1 years). The recording duration ranged from 28 to 42 minutes, with an average of 32.8 minutes. Autism diagnoses were subsequently confirmed using the medical history, and clinical assessments conducted by child and adolescent psychiatrists. The DSM-5 [25] criteria at the National Autism Center, using standardized assessments by child and adolescent psychiatrists.

The control group was selected from the pediatric department of the Child Welfare Teaching Hospital, EEG unit. Anonymization techniques were applied to maintain confidentiality and adhere to ethical standards.

The dataset indicates that parents did not report any instances potentially linked to epileptic seizures among the cases.

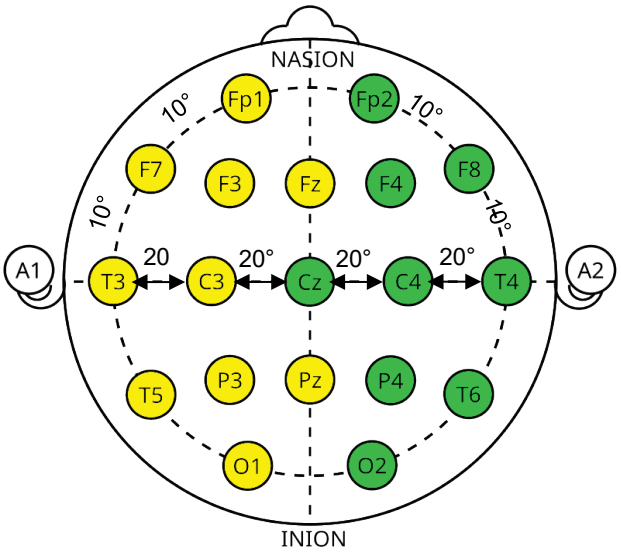


Figure 2 Illustration of electrode connections using 19 channels based on the 10/20 international electrode placement system

4 PROPOSED METHOD

The pre-processing stage included some significant steps. Data downsampling of data was initially performed in to reduce the sampling frequency from 0.5 kHz to 0.25 kHz for quicker processing. Filtering was subsequently performed utilizing linear-phase Finite Impulse Response (FIR) Band-Pass Filter (BPF) to remove frequencies less than 0.2 Hz and greater than 48 Hz. The aim of this step was to remove 50 Hz low-frequency artifacts and baseline noise. The filter was a Hamming-window design with heuristic order and transition bandwidth estimation, conducted automatically in EEGLAB [26].

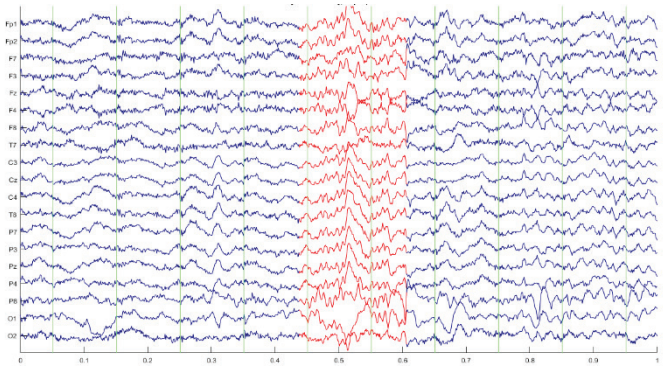


Figure 3 The work of ASR red color presented excluded segment of EEG

A key pre-processing step was Artifact Subspace Reconstruction (ASR), a technique that removes artifacts from EEG signals. ASR had three steps: 1st extracting reference data from raw EEG signals, 2nd estimating thresholds for identifying artifact components, and 3rd rejecting artifact components and reconstructing the cleaned data [27]. ASR effectively removes artifacts and reconstructs

the remaining EEGs, enhancing the quality and reliability of subsequent analysis and interpretation. Fig. 3 shows the rejected EEG signal segments identified by ASR in red and the reconstruction of the remaining segments in blue color.

Detrending is used to removes offsets or linear trends from time-domain input-output EEG data after artifact removal. It helps in detrend the data to eliminate systematic biases or drifts that may be operating it and ensures that the data is zero-centered or follows a stationary trend. While dealing with EEG signals, baseline shifts or slow linear trends are very common, which may distort the interpretation and analysis of the data. Detrending is done to eliminate these trends and focus on the dynamics of interest [28]. The technique involves estimating the trend component in the data and subtracting it to obtain detrended data. Fig. 4(B). An example of an output signal after detrending, shows a clear difference compared to Fig. 4(A), which representing the input EEG signals.

Following Amplitude Normalization (AN) was employed to increase the interpretability, removes the DC component, reliability, and effectiveness of EEG signal analysis for a broad spectrum of applications, e.g., neuroscience studies, clinical diagnosis, and brain-computer interface development [29]. The AN range was +1 to -1. Fig. 4(C). An example of an output signal after AN shows, a clear difference compared to Fig. 4 (A and B), representing the input EEG signals.

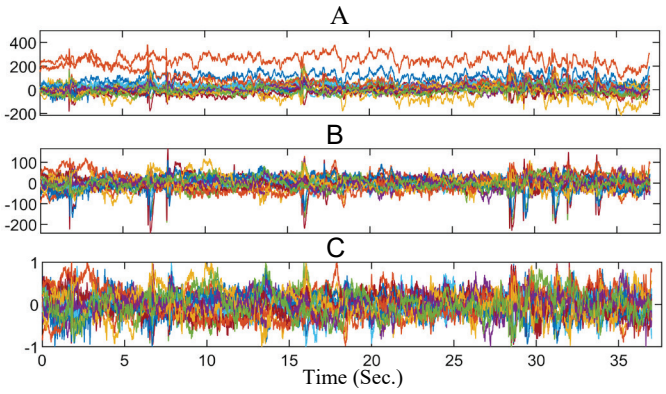


Figure 4 (A) Input EEG signals after the downsampling step, (B) output signal after detrending the input EEG signals, and (C) output signal after amplitude normalization

The EEG signal was then decomposed into sub-frequency bands using five BPFs, each implemented as a 6th-order Infinite Impulse Response (IIR) Chebyshev Type II filter and designed to extract specific EEG bands: [Delta (Δ): 0.2 – 4.0 Hz, Theta (θ): 4.0 – 8.0 Hz, Alpha (α): 8.0 – 13.0 Hz, Beta (β): 13.0 – 28.0 Hz, and Gamma (Γ): 31.0 – 48.0 Hz] as displayed in Fig. 5 a five decomposed frequency bands obtained using Fast Fourier Transform (FFT), shown in the frequency domain.

FFT is an efficient algorithm to compute the Discrete Fourier Transform (DFT), which is mathematically defined in Eq. (1):

$$X_h = \sum_{n=0}^{N-1} f(n) \cdot \mathrm{e}^{-j2\pi\frac{hn}{N}} \tag{1}$$

Where, X_h is the DFT of the sequence $f(n)$, $f(n)$ is the EEG signal in the time domain, $n = 0, 1, 2, \ldots, (N-1)$ is the time-domain sample index, $h = 0, 1, 2, \ldots, (N-1)$ represents the frequency index, and $\mathrm{e}^{-j2\pi\frac{hn}{N}}$ is the complex exponential, with j representing the imaginary unit. Fig. 5 illustrates an output of input EEG data.

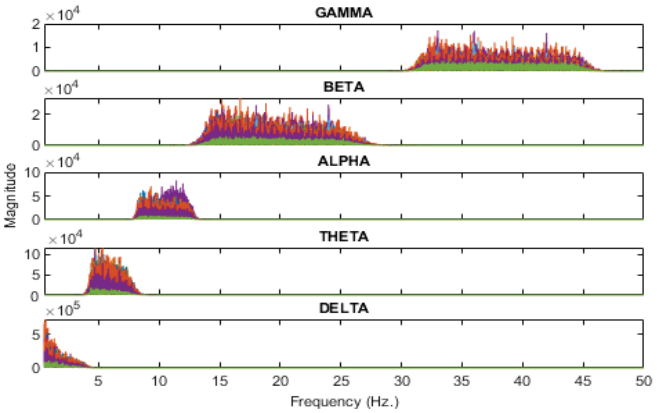


Figure 5 The FFT for the decomposed EEG bands

For analysis, EEG channels were grouped into two categories based on brain hemispheres: 1) Right hemisphere channels - Fp2, F4, C4, P4, O2, F8, T4, T6, and Cz. 2) Left hemisphere channels - Fp1, F3, C3, P3, O1, F7, T3, T5, Fz, and Pz, see Fig. 2 for the green and yellow channels, respectively.

The peak envelope method [30] was used to derive signal envelopes for all sensors. This process resulted in 10 envelope signals for the left hemisphere and 9 for the right hemisphere, as illustrated in Fig. 6. The mean of these envelopes was computed for each frequency band, generating a single envelope signal for each brain hemisphere.

The energy values were computed using a sliding window method on the mean envelope of each brain hemisphere for each decomposed band separately. The envelope signals denoted as ($\hat{H}_{\text{band, hemisphere}}[n] = \Psi$) were divided into overlapping windows. Define the index of a window as W denotes the window index and N the window length, experimentally set to 2×10^3 samples (~8 seconds), which is considered sufficient based on prior studies [11] and [29] suggests that an 8-second window provides informative data for EEG analysis.

Different overlap ratios, ranging from 12.5% to 87.5% with a step size of 250 samples, were tested to determine the most effective overlap, providing a measure of signal energy for each EEG band and hemisphere. The windowed signal envelope $E_{\text{band, hemisphere}}[n]$ is defined by the following Eq. (2):

$$E_{\text{band, hemisphere}} = \sum_{n=0}^{N-1} \left| \Psi\left[n + W(N - \Phi)\right] \right|^2 \tag{2}$$

Where Φ is the overlap in samples. The envelope energy in a window, $E_{\text{band, hemisphere}}$ is given by the sum of the squared envelope values over the window, $n = 0, 1, 2, \ldots, (N-1)$ is

the time-domain sample index within each window, and N is the total number of samples in the window.

Overlapping windows ensure continuity of the EEG signal and prevent the loss of valuable features between adjacent windows. Overlap improves feature extraction and enhances model performance. Mild overlapping allows the model to capture temporal patterns that might otherwise be overlooked without overlapping strides. Overlapping sliding windows have been shown to significantly improve classification accuracy [11]. In contrast, non-overlapping windows treat each segment as independent, potentially missing transitional features.

The proposed method is displayed in Fig. 6.

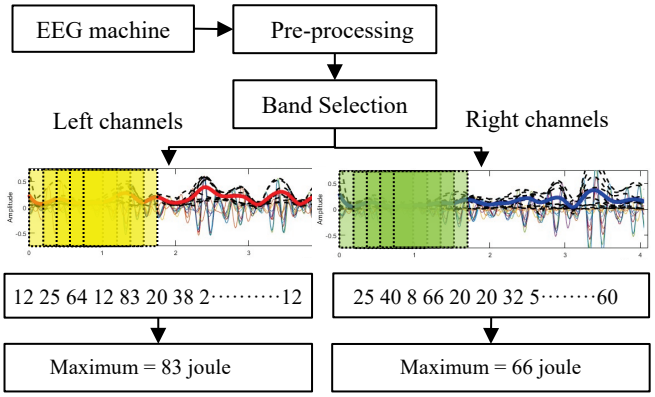


Figure 6 The proposed methodology for measuring energy in both brain hemispheres

5 RESULTS AND DISCUSSION

The experiments were conducted on a Microsoft Windows 10 Pro (22H2) operating system, with MATLAB® + plugin EEGLAB [26] on an Intel® CoreTM i7-7500U CPU @ 2.70-2.90 GHz, RAM 12.0 GB computer. The EEG signals were processed with EEGLAB V2023.1. With the help of the Automated artifact rejection tool in EEGLAB, employing the "Clean raw data" EEGLAB plugin, bad data segments were automatically excluded.

The study aimed to investigate hemispheric differences in EEG energy as potential biomarkers for ASD. The energy values were compared between the right and left hemispheres. In ASD cases, energy levels were generally higher in the right hemisphere (Left < Right), whereas in control cases, the left hemisphere exhibited higher energy (Left > Right).

Accuracy was calculated as the ratio of correctly classified samples to the total number of samples to quantify the discriminative power of these hemispheric energy differences. Specifically, for the Theta band, a stride of 250 samples was used, and the maximum energy within each window was selected as the feature for classification. Using this approach, the classification accuracy reached 100% for the ASD group and 68.18% for the control group, resulting in an overall accuracy of 84%, as presented in Tabs. 1 and 2. This demonstrates that hemispheric energy asymmetry can effectively differentiate between ASD and control subjects.

Table 1 Energy values for each hemisphere in Theta waves for the control group with a stride of 250 samples

ID	Gender	Left brain energy	Right brain energy	Left > Right
1	male	193.13	151.76	TRUE
2	male	113.85	92.52	TRUE
3	male	309.81	303.79	TRUE
4	male	157.84	147.53	TRUE
5	Female	184.92	187.41	FALSE
6	Female	192.28	176.64	TRUE
7	Female	184.70	202.92	FALSE
8	Female	182.92	177.72	TRUE
9	male	159.37	118.98	TRUE
10	male	110.27	125.66	FALSE
11	male	169.01	178.70	FALSE
12	male	139.19	145.49	FALSE
13	male	148.94	138.49	TRUE
14	male	168.17	161.15	TRUE
15	male	137.06	125.13	TRUE
16	male	215.52	179.88	TRUE
17	Female	129.69	180.37	FALSE
18	Female	158.48	178.18	FALSE
19	Female	285.15	212.42	TRUE
20	Female	193.20	174.88	TRUE
21	male	171.63	166.42	TRUE
22	Female	146.56	140.24	TRUE

Table 2 Energy values for each hemisphere in Theta waves for the ASD group with a stride of 250 samples

ID	Gender	Left brain energy	Right brain energy	Left < Right
1	male	121.54	180.96	TRUE
2	male	153.02	224.93	TRUE
3	male	151.69	178.93	TRUE
4	male	77.61	126.62	TRUE
5	male	143.91	198.98	TRUE
6	male	109.45	152.27	TRUE
7	male	125.24	162.18	TRUE
8	male	68.69	95.57	TRUE
9	male	169.67	284.28	TRUE
10	male	120.16	168.84	TRUE
11	male	72.19	127.10	TRUE
12	male	195.37	229.88	TRUE
13	Female	89.80	113.62	TRUE
14	male	110.95	170.70	TRUE
15	male	187.82	274.40	TRUE
16	male	133.56	160.57	TRUE
17	male	138.46	162.36	TRUE
18	male	139.37	188.63	TRUE
19	Female	84.84	99.54	TRUE
20	Female	97.35	213.04	TRUE
21	male	115.20	164.22	TRUE
22	Female	177.47	200.02	TRUE

The concept of comparing both brain hemispheres has been spurred by evidence showing that the right hemisphere is more active when psychotherapy occurs [31]. The right hemisphere has the capacity for implicit, integrative processing of affective, nonverbal, intuitive, and social information, which provides insight into relational aspects of the neurobiology of psychotherapy. This concurs with the relational psychotherapy model, which stipulates that right hemisphere processes that are collaboratively activated by both the therapist and the client lie at the center of change during psychotherapy.

In Tab. 3 presents the total accuracy calculated across both ASD and control groups, where accuracy is defined as the ratio of correctly classified samples for both groups to the total number of samples. Each value corresponds to a specific

EEG band and overlap ratio, and the results summarize performance in each case.

Table 3 Overall EEG band results with seven overlap ratios (total accuracy across ASD and control groups), showing the Theta band is dominant with a lower stride

Window length	Strid	Γ	β	α	θ	Δ
2000 Samples	250	52.2	72.7	75.0	84.1	65.9
	500	52.3	72.7	75.0	81.8	61.4
	750	54.5	72.7	77.2	84.0	68.1
	1000	54.5	72.7	75.0	77.3	63.6
	1250	50.0	72.7	75.0	81.8	70.4
	1500	54.6	70.5	77.3	79.5	65.9
	1750	50.0	72.7	75.0	75.0	68.2

Theta achieves the highest accuracy (84%) at stride 250 with 87.5% overlap, showing strong performance with high overlap. Its lowest value (77.3%) is at stride 1000 (50% overlap), suggesting a slight decline with decreasing overlap.

Fig. 7 shows the results of the theta band for a stride length of 250 samples, showing a summary of the differences across the 44 equally divided total samples for both groups including both genders. For ASD participants, assuming the right hemisphere had higher energy than the left, the difference was calculated as right minus left; for controls, the opposite was applied. The resulting differences are plotted to highlight group-specific patterns. These results emphasize the importance of the current study in focusing on feature extraction in future studies, considering both brain hemispheres and gender.

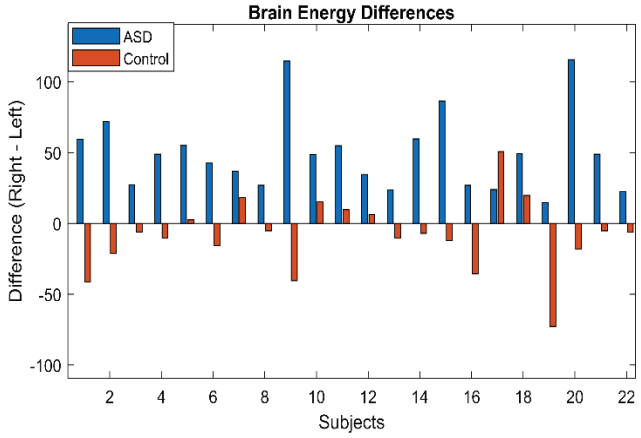


Figure 7 The results of brain energy differences, including both genders

The study examines gender differences in brain energy distributions between ASD and control groups using boxplots (Fig. 8 and 9). The boxplots illustrate the right-minus-left brain energy differences across ASD males, ASD females, control males, and control females, providing key statistical insights.

Each boxplot represents crucial statistical elements: the red line inside each box marks the median energy difference for that group, while the blue box captures the interquartile range (IQR), encompassing the middle 50% of the data. The whiskers extend to the minimum and maximum values within 1.5 times the IQR, and red crosses beyond the whiskers indicate outliers.

In the 2D visualizations (Fig. 10), male groups exhibit more apparent separation, whereas females display greater overlap, suggesting non-linear separability. The boxplots further confirm this trend, as ASD males show a positive median energy difference, indicating right-brain dominance with low variability, while ASD females exhibit stronger right-brain dominance but with higher variability and more outliers. Control males demonstrate a balanced energy distribution, whereas control females tend to have left-brain dominance with greater fluctuation.

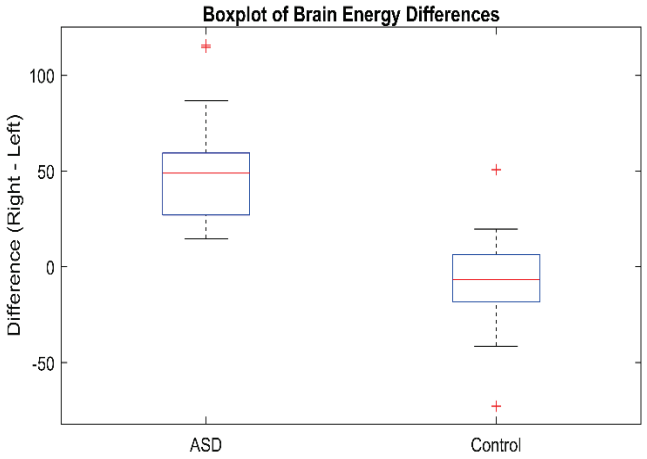


Figure 8 The results of the boxplot of brain energy differences, including both genders (male and female)

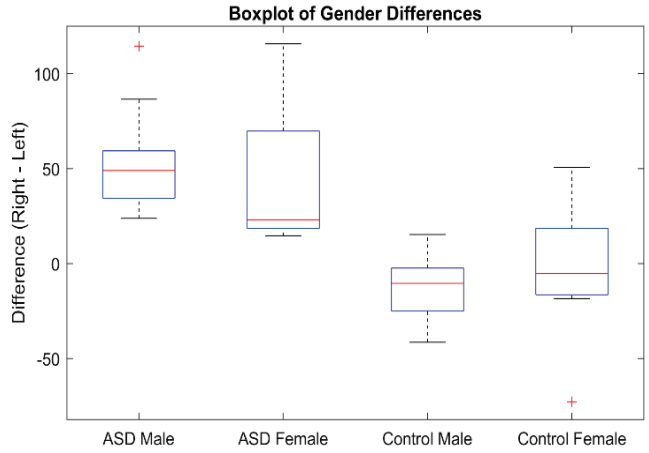


Figure 9 The results of the boxplot of brain energy differences separately for each gender (male and female)

Additionally, the energy quantities were plotted in joules, with the *X*-axis representing the left hemisphere and the *Y*-axis representing the right hemisphere (see Fig 10). This visualization helps in analyzing the energy distribution across brain hemispheres for a comprehensive assessment of ASD-related differences, shows both groups' differences and clearly highlights the error samples in the control group.

ASD participants consistently exhibit right-brain dominance, while control groups show a more balanced or left-leaning distribution. Gender differences influence variability, with females in both groups showing wider distributions and more pronounced differences.

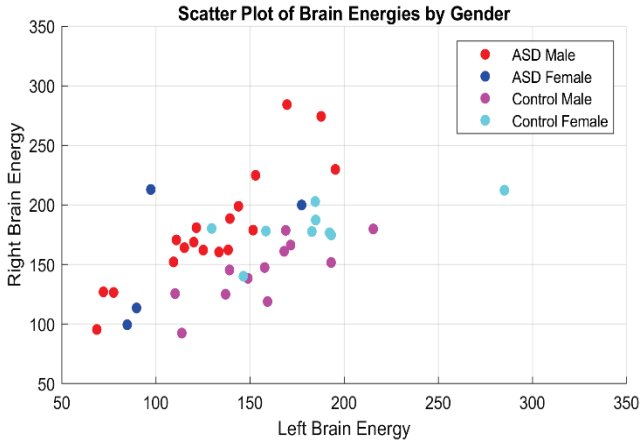


Figure 10 Two-dimensional feature representation of ASD and control groups plotted in 2D for each gender independently, the data is categorized into four groups: ASD Male, ASD Female, Control Male, and Control Female

From a diagnostic perspective, determining the best results depends on whether the focus is on consistency or group separation. If consistency is prioritized, ASD males provide the most reliable results due to lower variability and fewer outliers. However, suppose the goal is to maximize separation between the ASD and Control groups. In that case, ASD females offer the strongest distinction, as their median difference is significantly higher than that of Control females, though their greater variability may impact classification reliability.

ASD males and females achieve 100% classification accuracy, indicating a strong and reliable pattern in brain energy asymmetry. In the Control group, males reach 77% accuracy, while females have a lower accuracy of 55.56%, reflecting greater variability in their energy distributions. Overall, ASD classification remains highly accurate, whereas Control classification is more challenging, resulting in a total accuracy of 84%.

These findings highlight the potential of Theta band energy asymmetry as a biomarker for ASD diagnosis, demonstrating clear gender-specific trends. ASD individuals exhibit strong right-brain dominance, while Control participants tend to have more balanced energy distributions. Additionally, gender differences suggest that males provide more consistent data, whereas females exhibit greater variability but stronger group separation. Understanding these patterns can help refine diagnostic tools and improve the reliability of ASD classification using resting EEG data.

Theta performs best with high overlap and smaller strides (250 with 87.5% overlap), suggesting that shorter strides with larger overlaps optimize EEG feature detection (Tab. 3). The theta band is the most reliable across all conditions, consistently producing high accuracy values. While Gamma and Delta perform strongly under certain conditions, Beta exhibits the highest variability, particularly with a 62.5% overlap.

In practice, high overlap combined with small stride lengths enhances accuracy in Theta and Alpha bands, whereas Delta performs best at moderate strides with lower overlaps. These findings guide EEG data analysis, particularly in determining optimal windowing parameters for diagnostic applications like ASD detection.

The theta band demonstrated the most consistent performance across the ASD and Control groups. Energy level analysis between hemispheres revealed significant differences in functional connectivity across genders. In the Control group, most participants (15 out of 22) exhibited greater left hemisphere energy, suggesting left-hemisphere dominance (Tab. 1).

In the ASD group, right-hemisphere dominance was more prevalent (22 out of 22 participants -100%) compared to the Control group, suggesting imbalances in functional connectivity that may contribute to ASD-related challenges. ASD males and females exhibited stronger right-hemisphere dominance (Tab. 2).

The theta band emerges as the most reliable indicator of brain-behavior correlations. Male subjects, particularly those in the Control group, exhibited peak performance. These findings emphasize the role of gender-specific hemisphere energy dominance in cognitive function, particularly in ASD diagnosis.

The lower classification accuracy in the control group, especially among females, highlights the challenge of distinguishing ASD from controls. These results support the argument that ASD in females presents unique characteristics that complicate diagnosis. The higher classification accuracy in ASD and control males suggests that the method may be more effective at detecting ASD in males while underscoring the need for further research into ASD in females and improvements in diagnostic procedures.

The limitations of the current study include the relatively small sample size (44 participants), which may limit generalizability, although it is comparable to other studies in the field, with some having fewer participants and others the same sort of sample. The data set was treated rather than being split into test sets and training sets as AI based methods. While non-arbitrary in its approach, it may have a somewhat negative impact on predictive accuracy. Gender imbalance, because of the higher prevalence of ASD among males, also has an effect, as ASD in females has more subtle behavioral and neural patterns. Another limitation is that children were sedated using chloral hydrate, which contaminates EEG signals and distorts them by inducing low-frequency, high-amplitude activity that might mask or imitate actual neural features.

Future studies should focus on increasing the sample size and diversity to improve generalizability, exploring additional EEG frequency bands for a more comprehensive understanding of ASD, and developing gender-specific models, particularly for females. Longitudinal studies that track symptom progression and integrating the current features with AI technologies could further enhance diagnostic precision.

6 CONCLUSIONS

This study presents a new EEG-based method for diagnosing autism through hemispheric energy asymmetry in resting states. The approach showed excellent potential for segregating ASD from controls, particularly when accounting for gender-patterned trends. The uniform right-hemispheric energy dominance among ASD subjects, as opposed to the more random energy distribution in controls,

highlights the significance of lateralized neural activity in neurodevelopmental disorders.

The findings validate the use of hemispheric energy analysis as an input in diagnostic processes, with particular attention being placed on the Theta band as a reliable biomarker. Furthermore, the evidence of gendered differences underscores the importance of differential diagnostic models. Future work should prioritize the improvement of feature extraction algorithms and larger datasets to maximize diagnostic accuracy and generalizability.

7 REFERENCES

[1] Masini, E., Loi, E., Vega-Benedetti, A. F., Carta, M., Doneddu, G., Fadda, R., & Zavattari, P. (2020). An overview of the main genetic, epigenetic and environmental factors involved in autism spectrum disorder focusing on synaptic activity. *International journal of molecular sciences*, *21*(21), 8290. https://doi.org/10.3390/ijms21218290

[2] Hernawan, A. D., Diningrum, A., Jati, S. N., & Nasip, M. (2018). Risk factors of autism spectrum disorder (ASD). *Unnes Journal of Public Health*, *7*(2), 104-112. Retrieved from https://www.verywellmind.com/autism-causes-and-risk-factors-5195311

[3] Maenner, M. J. (2023). Prevalence and characteristics of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2020. *MMWR. Surveillance Summaries*, *72*. https://doi.org/10.15585/mmwr.ss7202a1

[4] World Health Organisation (WHO). (15 Nov.2023). autism. Retrieved from https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders

[5] Autism Spectrum Australia (ASA). (2 April 2024). At least 1 in 40 Australians are Autistic: New estimate by Aspect on World Autism. *Autism Spectrum Australi*a. Understanding Day**.** Retrieved from https://www.aspect.org.au/news/at-least-1-in-40-australians-are-autistic

[6] Statistics. Retrieved from https://www.statista.com/statistics/1250352/south-korea-number-of-registered-autism-cases/

[7] Ozonoff, S., Iosif, A. M., Baguio, F., Cook, I. C., Hill, M. M., Hutman, T., ... & Young, G. S. (2010). A prospective study of the emergence of early behavioral signs of autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, *49*(3), 256-266. Retrieved from https://pmc.ncbi.nlm.nih.gov/articles/PMC2923050/

[8] Estes, A., Zwaigenbaum, L., Gu, H., St. John, T., Paterson, S., Elison, J. T., ... & IBIS network. (2015). Behavioral, cognitive, and adaptive development in infants with autism spectrum disorder in the first 2 years of life. *Journal of neurodevelopmental disorders*, *7*, 1-10. https://doi.org/10.1186/s11689-015-9117-6

[9] Faja, S. U. S. A. N., & Dawson, G. E. R. A. L. D. I. N. E. (2006). Early intervention for autism. *Handbook of preschool mental health: Development, disorders, and treatment*, 388-416.

[10] Zwaigenbaum, L., Bauman, M. L., Stone, W. L., Yirmiya, N., Estes, A., Hansen, R. L., ... & Wetherby, A. (2015). Early identification of autism spectrum disorder: recommendations for practice and research. *Pediatrics*, *136*(Supplement_1), S10-S40. https://doi.org/10.1542/peds.2014-3667C

[11] Falih, B. S., Sabir, M. K., & Aydın, A. (2024). Impact of Sliding Window Overlap Ratio on EEG-Based ASD Diagnosis Using Brain Hemisphere Energy and Machine Learning. *Applied Sciences*, *14*(24), 11702. https://doi.org/10.3390/app142411702

[12] Mróz, K., Jonak, K., Plechawska-Wójcik, M., Siejka, A., Zimenkovsky, A., & Rejdak, R. (2025). Analysis of visual perception in children using an eye tracker–A pilot study. *Advances in Science and Technology Research Journal*, *19*(2), 255-270. https://doi.org/10.12913/22998624/195719

[13] Zhang, K., Du, X., Liu, X., Su, W., Sun, Z., Wang, M., & Du, X. (2022). Gender differences in brain response to infant emotional faces. *BMC neuroscience*, *23*(1), 79. https://doi.org/10.1186/s12868-022-00761-5

[14] Fenske, S. J., Liu, J., Chen, H., Diniz, M. A., Stephens, R. L., Cornea, E., ... & Gao, W. (2024). Sex differences in brain-behavior relationships in the first two years of life. *bioRxiv*. https://doi.org/10.1101/2024.01.31.578147

[15] Wei, X., Adamson, H., Schwendemann, M., Goucha, T., Friederici, A. D., & Anwander, A. (2023). Native language differences in the structural connectome of the human brain. *Neuroimage*, *270*, 119955. https://doi.org/10.1016/j.neuroimage.2023.119955

[16] Stroganova, T. A., Komarov, K. S., Sysoeva, O. V., Goiaeva, D. E., Obukhova, T. S., Ovsiannikova, T. M., ... & Orekhova, E. V. (2020). Left hemispheric deficit in the sustained neuromagnetic response to periodic click trains in children with ASD. *Molecular Autism*, *11*, 1-22. https://doi.org/10.1186/s13229-020-00408-4

[17] Nora, A., Renvall, H., Ronimus, M., Kere, J., Lyytinen, H., & Salmelin, R. (2021). Children at risk for dyslexia show deficient left-hemispheric memory representations for new spoken word forms. *NeuroImage*, *229*, 117739. https://doi.org/10.1016/j.neuroimage.2021.117739

[18] Olulade, O. A., Seydell-Greenwald, A., Chambers, C. E., Turkeltaub, P. E., Dromerick, A. W., Berl, M. M., ... & Newport, E. L. (2020). The neural basis of language development: Changes in lateralization over age. *Proceedings of the National Academy of Sciences*, *117*(38), 23477-23483. https://doi.org/10.1073/pnas.1905590117

[19] Blinch, J., Flindall, J. W., Smaga, Ł., Jung, K., & Gonzalez, C. L. (2019). The left cerebral hemisphere may be dominant for the control of bimanual symmetric reach-to-grasp movements. *Experimental Brain Research*, *237*, 3297-3311. https://doi.org/10.1007/s00221-019-05672-2

[20] Murrell, D. M. (2018). Left brain vs right brain: Fact and fiction, *Med. News Today*. Retrieved from https://www.medicalnewstoday.com/articles/321037.php

[21] Heilman, K. M., & Abell, T. V. D. (1980). Right hemisphere dominance for attention: the mechanism underlying hemispheric asymmetries of inattention (neglect). *Neurology*, *30*(3), 327-327. https://doi.org/10.1212/wnl.30.3.327

[22] Chiron, C., Jambaque, I., Nabbout, R., Lounes, R., Syrota, A., & Dulac, O. (1997). The right brain hemisphere is dominant in human infants. *Brain: a journal of neurology*, *120*(6), 1057-1065. https://doi.org/10.1093/brain/120.6.1057

[23] Spence, S., Shapiro, D., & Zaidel, E. (1996). The role of the right hemisphere in the physiological and cognitive components of emotional processing. *Psychophysiology*, *33*(2), 112-122. https://doi.org/10.1111/j.1469-8986.1996.tb02115.x

[24] Morley, A., Hill, L., & Kaditis, A. (2016). 10-20 system EEG Placement. *European Respiratory Society, European Respiratory Society*. Retrieved from http://en.wikipedia.org/wiki/10-20_system_%28EEG%29

[25] Volkmar, F. R., & Reichow, B. (2013). Autism in DSM-5: progress and challenges. *Molecular autism*, *4*, 1-6. https://doi.org/10.1186/2040-2392-4-13

[26] Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods*, *134*(1), 9-21. https://doi.org/10.1016/j.jneumeth.2003.10.009

[27] Chang, C. Y., Hsu, S. H., Pion-Tonachini, L., & Jung, T. P. (2019). Evaluation of artifact subspace reconstruction for automatic artifact components removal in multi-channel EEG recordings. *IEEE transactions on biomedical engineering*, *67*(4), 1114-1121. https://doi.org/10.1109/TBME.2019.2930186

[28] De Cheveigné, A., & Arzounian, D. (2018). Robust detrending, rereferencing, outlier detection, and inpainting for multichannel data. *NeuroImage*, *172*, 903-912. https://doi.org/10.1016/j.neuroimage.2018.01.035

[29] Urbach, T. P., & Kutas, M. (2006). Interpreting event-related brain potential (ERP) distributions: implications of baseline potentials and variability with application to amplitude normalization by vector scaling. *Biological psychology, 72*(3), 333-343. https://doi.org/10.1016/j.biopsycho.2005.11.012

[30] Yang, Y. (2017). A signal theoretic approach for envelope analysis of real-valued signals. *IEEE Access*, *5*, 5623-5630. https://doi.org/10.1109/ACCESS.2017.2688467

[31] Schore, A. N. (2020). The right brain is dominant in psychotherapy. In The Divided Therapist (pp. 70-92). https://doi.org/10.4324/9781003049876-1

Authors' contacts:

Bashar S. Falih, Ph.D.
(Corresponding author)
(1) Department of Computer Techniques Engineering,
Al Salam University College, Baghdad, Iraq
(2) Department of Electrical and Electronics Engineering,
Çukurova University, Adana, Türkiye
bashar.s.falih@alsalam.edu.iq

Mohannad K Sabir, Ph.D.
Department of Biomedical Engineering,
Al-Khwarizmi College of Engineering,
University of Baghdad, Baghdad, Iraq
dr.mohaned@kecbu.uobaghdad.edu.iq

Zainab H. Qadoori, Ph.D.
Department of Educational and Psychological Sciences,
College of Education and Human Science Diyala University, Diyala, Iraq
zaynab.psv.hum@uodiyala.edu.iq

Ali H. Ali, MD Board
National Autism Centre, Child and Adolescent Psychiatrist,
Children's Welfare Teaching Hospital, Medical City, Baghdad, Iraq
ali_idris95@yahoo.com

Noor H. Al-Janabi, MD Board
EEG Unit, Clinical Neurophysiologist, Children's Welfare Teaching Hospital,
Medical City, Baghdad, Iraq
hamza64@gmail.com

Ahmet Aydın, Ph.D.
Department of Biomedical Engineering,
Çukurova University, Adana, Türkiye
aaydin@cu.edu.tr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250325234233

Original scientific paper
Received: 2025-03-25, Accepted: 2025-04-15

Enhancing Gas Turbine Cooling with Upstream Fogging and Nanofluids: Evaluating Evaporative Cooling Effects

Alaa Jasim Abdulah*, Hashim A. Hussein, Abdul Jabbar Owaid

Abstract: In recent years, Nano fluids have gained significant interest in thermal applications due to their enhanced heat transfer properties. Among these, Al_2O_3-water Nano fluids stand out for their high thermal conductivity and effectiveness in cooling systems. A previous study investigated the effect of cylindrical obstacles on convection within an inclined enclosure using Al_2O_3-water Nano fluids, showing that geometric configurations strongly influence thermal performance. These findings underscore the importance of optimizing design and working fluids to improve heat transfer in advanced cooling systems. Building on this, the current study examines the integration of Fogging Upstream technology with Al_2O_3-water Nano fluids to improve gas turbine cooling performance. Fogging Upstream involves injecting a fine mist of water or Nano fluid before the compression stage to lower the intake air temperature, thereby enhancing thermal efficiency. The study also investigates the role of pipe diameter (¼ inch vs. ½ inch) on cooling effectiveness. Experimental results show that Fogging Upstream alone improves cooling efficiency by 12–18%. When combined with Nano fluids, heat transfer improves by 15–25%. Additionally, using a ¼ inch pipe boosts cooling performance by 10–15% due to greater surface area and induced turbulence. Overall, the combined system reduces intake air temperature by 5–7°C and increases gas turbine efficiency by 8–12%.

Keyword: cooling performance; fogging upstream; gas turbines; heat exchange; nano fluid

1 INTRODUCTION

Gas turbines are an integral part of modern energy production in aviation and various industries, thanks to their high efficiency and reliable operation. But one of the most prominent challenges to improving the performance of gas turbines is the control of high inlet air temperatures, which leads to reduced operational efficiency and increased fuel consumption, in addition to the possibility of turbine components being damaged. One of the main strategies to mitigate these problems is to cool the inlet air before it enters the turbine, as many cooling techniques have been studied to improve turbine performance. Among these technologies, "cooling by spray into the air duct" is an effective method of introducing water or spray into the air stream before it reaches the compressor, cooling the inlet air through a large evaporation process, which improves the thermal efficiency of the turbine. Because of their great efficiency and dependable operation, gas turbines are essential to modern power production in aviation and industrial applications. But controlling high intake air temperatures, which can result in decreased operational efficiency, increased fuel consumption, and possible damage to turbine components, is one of the major obstacles to gas turbine performance optimization [1, 2]. A key tactic to lessen these problems is to cool the intake air before it enters the turbine, and several cooling technologies have been investigated to improve turbine performance. An efficient method called "fogging upstream" is introducing water or mist into the air stream prior to it reaching the compressor [3]. The injected mist undergoes an evaporative cooling process which significantly lowers the temperature of the intake air improving the thermal efficiency of the turbine [4]. In recent years, Nano fluids have received increasing attention in thermal applications due to their enhanced heat transfer capabilities. Among these, the Al_2O_3-water Nano fluid is widely used for its improved thermal properties. Authors examined the impact of cylindrical obstacles on convection within an inclined enclosure containing an Al_2O_3-water Nano fluid. The study revealed that the geometric configuration of the obstacles significantly affects heat transfer performance. Such findings open the door to optimizing cooling systems in industrial and high-performance thermal applications. Building on this, the present research aims to investigate the effects of laser energy on the mechanical and thermal properties of aluminium alloys coated with Nano silicon particles, contributing to the development of more efficient and durable thermal systems [5]. In order to improve gas turbine performance, especially in areas with high ambient temperatures, this cooling technique has been extensively researched. Recent research has demonstrated the benefits of fogging systems in raising gas turbine power production and lowering the amount of fuel needed to maintain ideal operating conditions [6]. Nano fluids, which are fluids created by dispersing nanoparticles (such SiO_2, CuO, or AlO_3) in a base fluid like water, have drawn a lot of attention lately as a cutting-edge way to enhance heat transmission. Nano fluids are excellent options for enhancing evaporative cooling performance because of their higher thermal conductivity as compared to traditional fluids [7]. Recent research has demonstrated that the use of Nano fluids in fogging systems can further enhance cooling efficiency and increase heat transfer rates especially under high-temperature conditions [8]. Several recent studies have focused on the integration of fogging and Nano fluid technology. In another study by Mukesh M. et al 2020 [9] the effectiveness of a hybrid air conditioning system using Nano fluids was evaluated. The results showed that the system's coefficient of performance (COP) was significantly improved with Nano fluids which are known for their high thermal conductivity due to the presence of solid particles. Similarly R. Agbadede and B. Kainga 2021 [10] reported that introducing inlet fogging into a gas turbine reduces intake air temperature by 10 °C. In their simulation of industrial gas turbine inlet fogging a 0.4% water-to-air ratio was used. Zhou Sun et al. 2021 [11] - This research monitors gas turbine bearing

temperatures using regression modeling. It addresses high compressor bearing temperatures at Geregu Power Plant, Nigeria, using Multi-Linear Regression (MLR) models to predict bearing temperature based on various lube oil parameters, improving gas turbine health and performance. Ghassan Fadhil Smaisim et al. 2022 [12] - This paper reviews the properties and applications of nanofluids, focusing on their enhanced thermal conductivity and heat transfer properties. It highlights potential health and environmental risks associated with nanoparticles, stressing the need for careful handling. Mohammad Zamen et al. 2022 [13] - The study introduces a new photovoltaic/thermal (PVT) system using Al2O3/water nanofluid as the coolant. The use of this nanofluid significantly improves the thermal and electrical performance of the PVT system.Shailendra Gajbhiye et al. 2023 [14] - This paper focuses on optimizing the performance of hybrid air conditioning systems with Al2O3 nanofluid. It finds that a 0.8% by weight concentration of the nanofluid improves the Coefficient of Performance (COP) by 34% and the refrigeration effect by 32%. The study of Atazaz Hassan et al. 2023 [15] examines the effectiveness of various parameters in improving the heat transfer performance of a hybrid photovoltaic thermal system (PV/T). It found that a water-based PV/T with a 16 mm tube diameter and 1.02 l/minute has the highest efficiency at 44.5% 14.8% and 59.3%. The study also found that increasing Reynolds numbers increases the system's efficiency and a maximum decrease in cell temperature can be achieved by increasing tube diameter and back plate thickness. The study of Balakrishnan S. et al. 2025 [16] investigated the effectiveness of shell and tube heat transfer on Nano fluids made of copper and alumina nanoparticles. The Nano fluids increased thermal conductivity by 29% and 39% when 0.12 wt % of these nanoparticles were added. The Nano fluids also increased friction factor and heat transfer coefficient by convection. Al_2O_3 Nano fluid outperformed Cu Nano fluid in pressure drop. A study examining the impact of pipe diameter on fogging system efficiency found that smaller diameters enhance turbulence, improving heat transfer but increasing pressure drop. The research aims to evaluate the effectiveness of Fogging Upstream combined with Nano fluids in improving gas turbine cooling efficiency. The study compares the impact of different pipe diameters (1/4 inch and 1/2 inch) on the overall performance of the fogging system. The goal is to determine the optimal combination of fogging technique and pipe diameter for optimal cooling performance. The results will provide insights into integrating advanced cooling technologies to enhance gas turbine efficiency, and sustainability in power generation and other applications.

2 THE REYNOLDS NUMBER (*Re*)

The Reynolds number a dimensionless quantity is useful for forecasting flow patterns in a range of fluid flow situations. Usually laminar or layer-like flow prevails when low Reynolds numbers whereas in situations with higher Reynolds numbers turbulent flow usually wins out. Variations in the speed and direction of the fluid give rise to turbulent flow which can sometimes cause collisions or even flow directions that are opposed to the main stream (called eddy currents). These eddy currents use energy and raise the possibility of cavitation in a liquid when combined with flow variations. Reynolds numbers are important because they are dimensionless quantities in fluid mechanics. They show how the viscous and inertia forces in a fluid are proportionate [17], [18].

$$Re = \frac{\rho V D_h}{\mu} = \frac{V D_h}{\nu} \quad (1)$$

Where: *Re* - Reynolds number, a dimensionless quantity used to predict the flow regime (laminar or turbulent) in fluid dynamics; ρ - Density of the fluid (kg/m³); V - Average flow velocity (m/s); D_h - Hydraulic diameter (m), defined depending on the geometry of the flow channel; μ - Dynamic viscosity of the fluid (Pa·s or N·s/m²); ν - Kinematic viscosity of the fluid (m²/s).

Where the impact of different pipe diameters (1/4 inch and 1/2 inch) on cooling efficiency using an Al_2O_3-water Nano fluid mixture is analysed the Reynolds number (*Re*) plays a crucial role in determining the flow characteristics inside these pipes and their effect on heat transfer.

The Effect of Reynolds Number on Flow inside the Pipes:

1) Smaller diameter pipes (1/4 inch):
- Result in higher flow velocities leading to an increase in Reynolds number making the flow more turbulent.
- Turbulent flow enhances thermal mixing improving heat transfer efficiency and boosting the cooling performance.
- However this comes at the cost of higher pressure drop which may affect the system's efficiency.

2) Larger diameter pipes (1/2 inch):

a) Lead to lower flow velocities resulting in a lower Reynolds number making the flow closer to laminar conditions.

b) In laminar flow heat transfer is less efficient compared to turbulent flow but pressure losses are reduced improving hydraulic efficiency.

3 EXPERIMENTAL SETUP

3.1 Fogging Cooling System (Upstream)

The experiment Nanomaterial's were prepared at the University of Technology, Department of nanotechnology. A test rig was constructed to the intake cooling process of a gas turbine. The system consists of a variable-speed water pump (BF-3800 model), a fluid atomization nozzle for fogging, pipelines with different diameters (¼ inch and ½ inch), and a controlled air intake duct. The nozzle was positioned upstream of the compressor simulation chamber to allow uniform distribution of the mist into the incoming airflow. Two working fluids were used:

1) Distilled Water for the baseline fogging experiment.
2) Al_2O_3–water Nano fluid with a nanoparticle concentration of 2% by volume.

The Nano fluid was prepared using ultrasonic agitation to ensure uniform dispersion of the nanoparticles and

stability of the mixture. The fogging cooling system installed in the upstream section of a gas turbine is designed to enhance turbine performance and efficiency by cooling the intake air before it reaches the compressor. This process reduces intake air temperature increasing air density and allowing a greater mass of air to enter the compressor thereby improving power output and efficiency. The system operates through water injection where a fine mist of water is introduced into the intake air stream. As the water droplets evaporate they absorb heat from the air effectively lowering its temperature through evaporative cooling by utilizing the latent heat of vaporization. Increased mass flow as a result boosts power production and improves combustion efficiency. High-pressure pumps provide steady water flow and pressure [19], high-pressure nozzles atomize water into fine droplets for efficient cooling, and sophisticated control systems adjust water flow according to turbine operations and ambient conditions to optimize cooling efficiency. One efficient way to improve turbine performance is to install a fogging cooling system in the upstream area of a gas turbine. This method raises the density of the intake air by chilling it, increasing the turbine's overall efficiency and power output. The cooling system, which is positioned ahead of the filters, keeps the compressor and turbine blades safe from excessive humidity. Additionally, it does not require demineralized water and shields the system from dust and water spray entering the compressor.

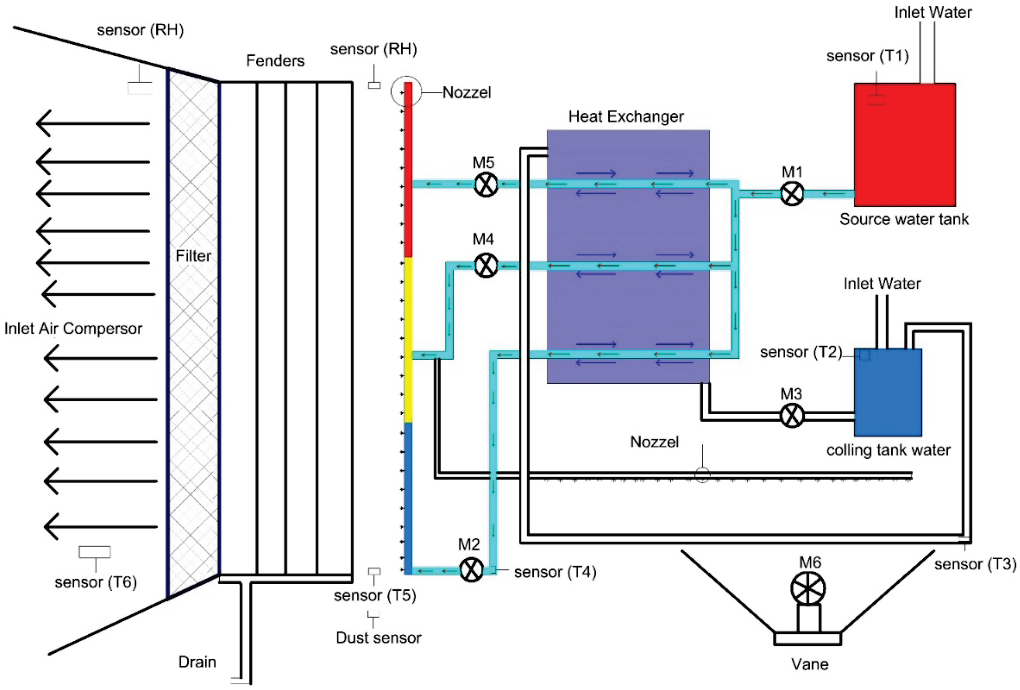


Figure 1 A Psychometric Diagram and Schematic Representation of the Fogging (Upstream) Cooling System

As seen in Fig. 1, the fogging (upstream) cooling system is mounted externally in front of the weather hoods. This configuration is perfect for retrofit applications since it permits installation and maintenance without interfering with turbine performance. This technology is very helpful when there is not enough room in the current air intake. It also eliminates the prolonged downtime brought on by expensive filter house modifications or warranty restrictions. Additionally, the system has a mesh that removes big particles, prolonging the life of air filters and making installation and maintenance easier [20]. A more recent technique for cooling gas turbine intake air (GTIC) is an indirect evaporative cooling system (IECS). With this method, water is directly evaporated into the secondary air stream, which is a component of the primary air flow. The IECS seeks to improve gas turbine performance and efficiency by effectively cooling the incoming air stream. Using the MS9001E turbine model, the fogging (upstream) system was put in place at the Al-Quds gas power station to cool the air inlets for the third and fourth generating units. By considerably lowering the temperatures inside the generating units, this initiative enhanced production capacity and performance throughout the hottest summer months [8].

The addition of a water tank with a cooling combination of nanomaterial's improved the system. Cooled water is pumped through pipes to a heat exchanger which simultaneously receives hot water from the main tank. This setup effectively reduces the water temperature before it reaches the nozzles resulting in a lower temperature of the air entering the compressor.

3.2 Fundamentals of Heat Exchanger Design

Heat is transferred from a hot fluid to a cold fluid typically with the help of a separator or barrier. However heat transfer can also occur directly between fluids without physical barriers such as when hot air interacts with a cold fluid. This type of heat transfer known as direct heat transfer is common in heat exchangers like industrial water coolers which can efficiently dissipate large amounts of heat. This

study explores the use of high thermal conductivity heat exchangers in industrial applications. Specifically it examines heat exchangers designed with pure copper tubes which are valued for their excellent thermal conductivity corrosion resistance and ease of joining through soldering [21, 22]. In addition a comparison will be made between two types of pipe diameters 1/4 inch and 1/2 inch to evaluate their impact on cooling efficiency heat transfer rates and system performance. The aim is to determine which pipe diameter offers better overall efficiency in heat exchange processes.

3.3 Methodology

The experimental process involved the development and testing of a practical cooling system. The system was constructed with a focus on empirical testing and real-world application mirroring the theoretical setup. The core of the experiment was the heat exchanger. This heat exchanger consisted of a large tank which stored hot water connected to a 0.5-0.25 inch copper tube.

3.3.1 Pipe Diameter Comparison

The experiment compared 0.5-inch and 0.25-inch copper tubes to examine how pipe diameter affects cooling performance:

- 0.25-inch Tube:
- Provided higher cooling efficiency due to increased turbulence enhancing heat transfer.
- Had a greater heat transfer surface area per unit volume leading to more effective cooling.
- Caused a slightly higher pressure drop which was a trade-off for improved thermal performance.
- 0.5-inch Tube:
- It was more suited to preserving flow stability since it had a reduced pressure drop.
- Showed lower turbulence which slightly reduced its heat transfer efficiency compared to the smaller tube.
- Was better for maintaining consistent flow rates but less effective in enhancing cooling performance.

Because of its superior thermal conductivity, copper was chosen to ensure efficient heat transmission from the hot water. Based on earlier studies, a Nano fluid with a 2% concentration of aluminum oxide (Al_2O_3) was placed in a secondary tiny tank to maximize cooling efficiency. Using a 1-inch tube, the Nano fluid tank and heat exchanger were integrated as part of the experimental setup. This arrangement made it possible for the Nano fluid to interact with the hot water, promoting cooling and heat absorption. While the Nano fluid circulated through the connecting tube, absorbing heat from the water, hot water from the main tank flowed through the heat exchanger's copper tube as the system was operationalized. Temperature sensors were positioned at key locations throughout the cooling system to track its performance. These sensors provided real-time information on the cooling performance by recording the temperatures of the hot water entering and leaving the heat exchanger. The primary goal was to achieve a substantial reduction in water temperature validating the cooling system's effectiveness. The data collected from the temperature sensors were analyzed to assess the performance of the cooling system. This analysis compared the observed temperature reductions with the theoretical predictions confirming whether the experimental outcomes aligned with the expected results. Based on the performance data adjustments were made to optimize the system. These adjustments included modifying flow rates and Nano fluid concentrations to enhance cooling efficiency. Throughout the process detailed documentation was maintained capturing the experimental setup procedures results and any modifications made. In essence the experiment involved constructing and testing a physical cooling system to empirically validate theoretical designs. This hands-on approach provided practical insights into the system's performance ensuring that theoretical predictions were accurately reflected in real-world conditions.

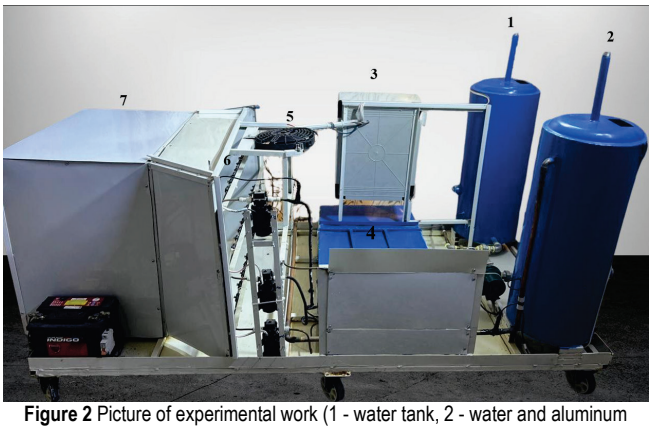


Figure 2 Picture of experimental work (1 - water tank, 2 - water and aluminum oxide mixture tank (Al2O3), 3 - electrical board, 4 - heat exchanger, 5 - cooling fan, 6 - water sprayers, 7 - filter)



Figure 3 Top view Explain the impeller and heat exchanger

3.4 System Components and Their Functions

- **Arduino Mega 2560:** The central control unit of the system, operating at 16 MHz. It has 256 KB of flash memory and 8 KB of SRAM. It processes sensor data,

applies control algorithms, and controls the relays for motors and the fan.

- **Arduino Nano:** A secondary controller with 16 MHz clock speed. It converts digital signals to analog, ensuring compatibility with the Arduino Mega and reducing the computational load on the main controller.
- **Sensors:** The system includes DS18B20 temperature sensors for accurate temperature readings, DHT11 humidity sensors, and the GP2Y1014AU0F dust sensor to monitor environmental conditions.
- **Actuators:** The system has six DC motors (12V) and one 220V AC fan, controlled by relays that turn the devices on or off based on commands from the Arduino Mega.
- **Relays:** These act as switches to control high-power actuators while isolating low-power control circuitry. They ensure safety and reliable operation by handling 5V control signals and 12V/220V loads.

3.5 Calibration and Signal Processing

- **Calibration:** The system compares sensor data to known reference values during initialization to correct any inaccuracies. For example, temperature sensors are calibrated against a precise thermometer. Calibration ensures consistent sensor accuracy.
- **Signal Filtering:** A low-pass filter (with transfer function 1/s+1) is implemented in the Arduino Mega to smooth out high-frequency noise and stabilize incoming signals. A hardware filter further reduces electrical noise.
- **Simulink Simulation:** Calibration is simulated using Simulink to analyze sensor outputs and generate correction equations, which are then applied in the control program of the Arduino Mega. Periodic recalibration ensures long-term accuracy.

3.6 System Operation

- **Data Collection and Processing:** The system continuously collects data from sensors, which is processed in real-time by the Arduino Mega. The control algorithm uses the calibrated data to determine the necessary actions for actuators.
- **Actuator Control:** Relays activate or deactivate motors or the fan based on specific environmental thresholds. For example, high temperatures may turn off certain motors, or high humidity may activate the fan.
- **Dust Level Control:** If dust levels exceed a set threshold, the system shuts down all motors and the fan as a protective measure. This prevents damage to the system due to poor air quality.

The system integrates precise sensors, efficient actuators, and robust control algorithms to achieve accurate and responsive environmental regulation, ensuring high reliability and adaptability for various applications.

4 RESULTS AND DISCUSSION

4.1 Case 1: Comparison of Cooling Efficiency Between 1/4 Inch and 1/2 Inch Systems Pipes in Heat Exchange is used

By analyzing the results shown in the three figures in Fig. 4 some important conclusions can be drawn about the effect of the addition of nanoparticles on the temperature distribution using tubes of different sizes (1/4 inch and 1/2 inch).

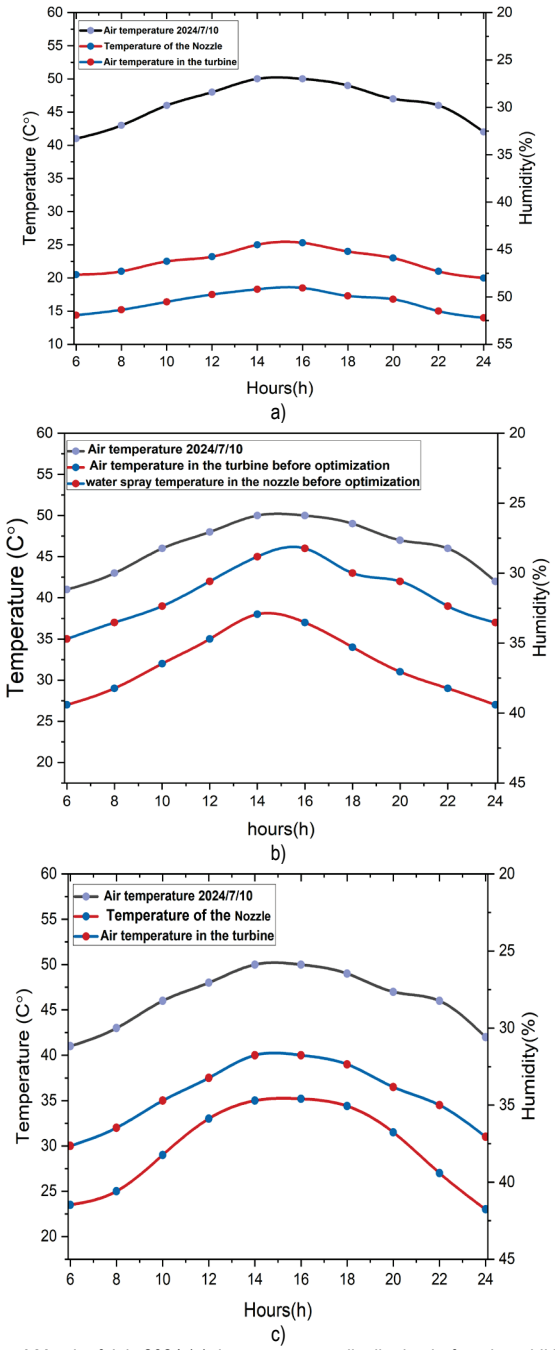


Figure 4 Month of July 2024 (a) the temperature distribution before the addition of the nanoparticles, (b) the temperature distribution after the addition of the nanoparticles in (1/4 inch) tube, (c) The temperature distribution after the addition of the nanoparticles in (1/2 inch) tube

1) Fig. 4a: thermal distribution before the addition of nanoparticles this figure shows the air temperature inside the turbine the nozzle temperature and the air temperature inside the turbine before optimization. We note that the air temperature inside the turbine gradually rises during the day reaching a peak at about 14:00 - 16:00 where it is about 50 °C. The nozzle temperature is lower than the turbine temperature which indicates that cooling is taking place but is not as efficient as required.

2) Fig. 4b: the effect of adding nanoparticles in a 1/4 inch tube when using a 1/4 inch diameter tube with nanoparticles we observe a noticeable decrease in temperatures inside the turbine and nozzle compared to Fig. 4a. The air temperature in the turbine after optimization is about 3-5 °C lower in the peak period than in the case before the addition of nanoparticles. This suggests that the use of nanoparticles has improved the efficiency of heat exchange within the system but there is still a clear rise in temperatures during the afternoon.

3) Fig. 4c: the effect of adding nanoparticles in a 1/2 inch tube when using a pipe with a diameter of 1/2 inch we notice a greater improvement in turbine cooling compared to a smaller pipe. The air temperature in the turbine decreases even more since the thermal difference compared to the Fig. 4a reaches 5-8 °C during the peak. The temperature distribution shows a smoother gradient during daylight hours suggesting that this system is more efficient at heat dissipation.

This analysis shows that increasing the diameter of the pipe to 1/2 inch enhances the efficiency of heat transfer and cooling which reduces the temperatures inside the turbine even more compared to the 1/4 inch pipe. The addition of nanoparticles significantly improved cooling in both cases but the effect was more pronounced when using a tube with a diameter of 1/2 inch. Pipes with a larger diameter (1/2 inch) achieve a better improvement in heat exchange compared to smaller pipes (1/4 inch) which helps reduce the temperatures in the turbine and nozzle. A system using 1/2 inch with nanoparticles is the most efficient in terms of reducing temperatures and therefore will be most useful in preventing overheating of the turbine and improving its performance in hot operating environments. Based on these results the use of 1/2 inch tubes with nanoparticles can be recommended to improve the cooling efficiency of turbine systems especially in high-temperature environments.

4.2 Case 2: Effective Cooling in Pipe with a Diameter of 1/4 Inch is used

Fig. 5a before adding nanoparticles the ambient air temperature reaches 50 °C at the peak (14:00).The nozzle temperature is about 45°C.The air temperature inside the turbine reaches 35-40 °C which indicates insufficient cooling. After the addition of nanoparticles Fig. 5b the ambient air temperature has not changed (50 °C).The nozzle temperature dropped to 35-40 °C which indicates an improved heat transfer. The air temperature inside the turbine dropped to 25-30 °C which is about 10 °C lower than the state before the improvement

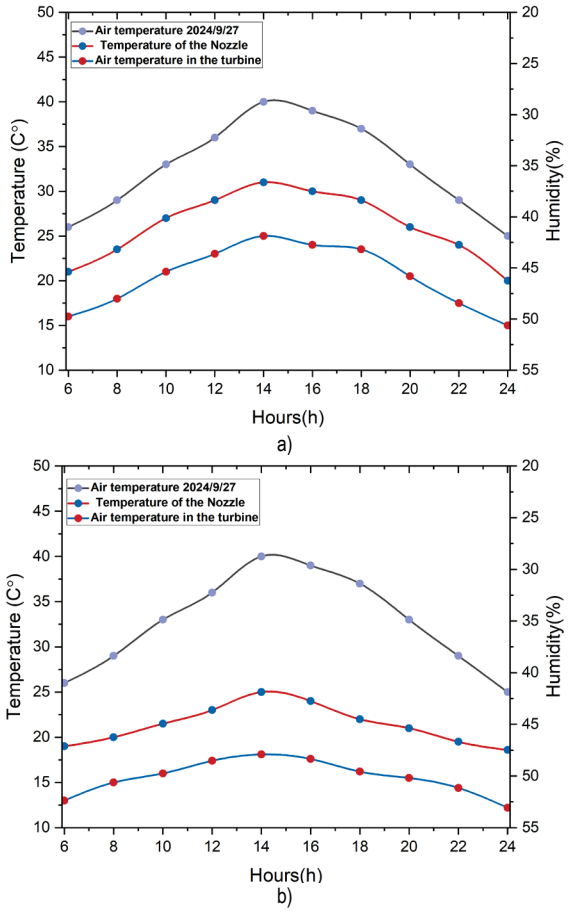


Figure 5 Month of September 2024 (a) the temperature distribution before the addition of the nanoparticles, (b) the temperature distribution after the addition of the nanoparticles

The distribution of convection before optimization Fig. 6a shows a similar pattern to September when the temperature reaches a peak of about 48-50 °C in the middle of the day. After optimization with nanoparticles Fig. 6b we observe a clear decrease in temperatures but at a lower rate compared to September 2024 when the decrease is only 2-4 °C. This shows that cooling with a 1/4 inch pipe becomes less efficient in months with a colder climate (such as October 2024) when the difference between the ambient temperature and the operating temperature is less. In general the results confirm that the addition of nanoparticles in a 1/4 inch tube leads to a noticeable improvement in cooling during extreme thermal conditions but its effectiveness gradually decreases with decreasing outdoor temperatures. This suggests that cooling strategies need to be adjusted according to the season and environmental conditions to achieve optimal performance.

In November 2024 the cooling system is shut down due to the naturally low ambient temperatures which provide adequate cooling without the need for mechanical systems. This decision is made to enhance operational efficiency as the cold weather achieves the required cooling levels naturally reducing energy consumption and optimizing overall system performance. As shown in the figure this seasonal shutdown demonstrates the system's integration with environmental conditions reflecting effective energy

management and operation aligned with the surrounding climate.

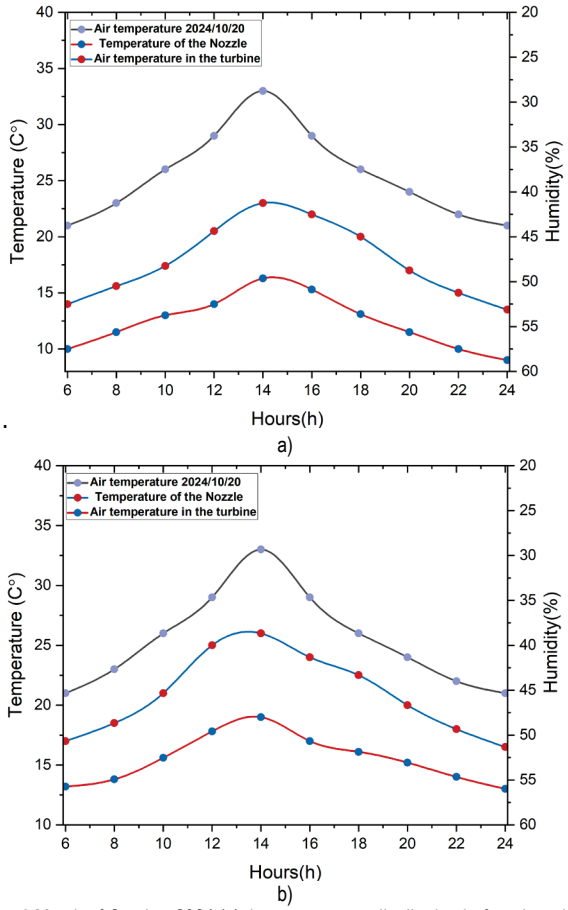


Figure 6 Month of October 2024 (a) the temperature distribution before the addition of the nanoparticles, (b) the temperature distribution after the addition of the nanoparticles

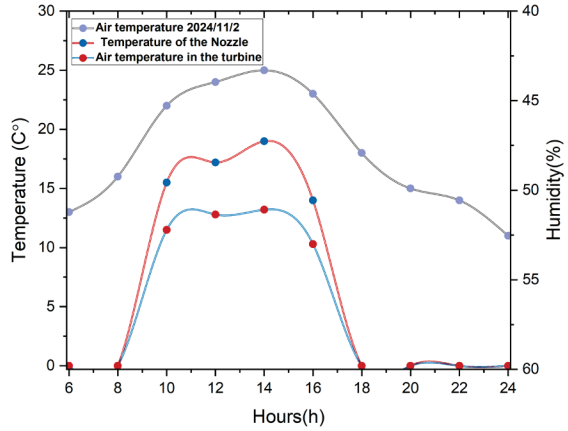


Figure 7 Month of November 2024 the temperature distribution before the addition of the nanoparticles

5 CONCLUSION

1) Enhancement of Heat Transfer Efficiency:
- The air temperature inside the turbine decreased by 5-7 °C which represents a 10-14% improvement in cooling efficiency compared to the pre-nanoparticle condition.
- This improvement is due to the increased thermal conductivity of the coolant allowing 15-20% better heat absorption from the hot air.

2) Effect of Pipe Diameter (1/4 Inch vs. 1/2 Inch):
- Using a 1/4-inch pipe resulted in better temperature control with a cooling efficiency increase of 12-18% compared to a 1/2-inch diameter pipe.
- The heat transfer rate improved by approximately 15% when using a smaller pipe with nanoparticles due to increased turbulence and surface area interaction.

3) Effect on Relative Humidity After Nanoparticle Addition:
- The relative humidity decreased by 5-8% indicating better cooling efficiency and improved moisture control within the system.

4) Impact of Lower Temperatures Inside the Turbine:
- The reduction in thermal stresses on internal components resulted in an estimated 20-25% extension in equipment lifespan and a 10-15% reduction in maintenance frequency. When equipment such as engines or turbines are operating under high temperatures, they are subjected to thermal stress caused by the frequent expansion and contraction of components due to thermal changes. Such stress can lead to deterioration of the material, which reduces the service life of the equipment
- Improved cooling efficiency also led to a 5-10% reduction in fuel consumption optimizing turbine performance.

5) Seasonal Influence and System Performance in September:
- In September the system was able to reduce peak air temperatures by 8-12% ensuring efficient turbine operation even during high daytime temperatures.
- The cooling effectiveness in transitional months (e.g. September) improved by approximately 10-15% demonstrating the system's adaptability to various climatic conditions.

6 REFERENCE

[1] Stefan, E., Talic, B., Larring, Y., Gruber, A., & Peters, T. A. (2022). Materials challenges in hydrogen-fuelled gas turbines. *International Materials Reviews, 67*(5), 461–486. https://doi.org/10.1080/09506608.2021.1981706

[2] Yang, Y., Nikolaidis, T., Jafari, S., & Pilidis, P. (2024). Gas turbine engine transient performance and heat transfer effect modelling: A comprehensive review, research challenges, and exploring the future. *Applied Thermal Engineering, 236*, 121523. https://doi.org/10.1016/j.applthermaleng.2023.121523

[3] Sahu, S., Mohapatra, A. K., Thatoi, D., & Panigrahi, N. (2021). Improving gas turbine performance using mechanical chiller: A thermodynamic approach. In *1st Odisha International Conference on Electrical Power Engineering, Communication and Computing Technology (ODICON 2021)*. https://doi.org/10.1109/ODICON50556.2021.9428940

[4] Cecere, D., Giacomazzi, E., Di Nardo, A., & Calchetti, G. (2023). Gas turbine combustion technologies for hydrogen blends. *Energies, 16*(19). https://doi.org/10.3390/en16196829

[5] Kaddouri, N., Kherris, S., Mostefa, K., Mekroussi, S., Saleh, M. S. M., & Zebbar, D. (2024). Cylindrical obstacle effect on

convection inside an inclined enclosure filled with a nanofluid. *Applied Engineering Letters, 9*(1), 22–36. https://doi.org/10.46793/aeletters.2024.9.1.3

[6] Zheng, K., Tian, W., Qin, J., & Hu, H. (2021). An experimental study on the improvements in the film cooling performance by an upstream micro-vortex generator. *Experimental Thermal and Fluid Science, 127*, 110410. https://doi.org/10.1016/j.expthermflusci.2021.110410

[7] Pourhedayat, S., Hu, E., & Chen, L. (2024). Yearly hour-by-hour performance evaluation of a SIEMENSE gas turbine power plant for Australia's climate condition with and without a novel intake air pre-cooling strategy. *Energy Conversion and Management: X, 21*, 100501. https://doi.org/10.1016/j.ecmx.2023.100501

[8] Williams, J. D., & Peterson, G. P. (2021). A review of thermal property enhancements of low-temperature nano-enhanced phase change materials. *Nanomaterials, 11*(10). https://doi.org/10.3390/nano11102578

[9] [Autor nepoznat], Gupta, M., Katyal, P., & Kumar, R. (2020). A comparative study of nanofluid (Al_2O_3) and distilled water in terms of thermal conductivity. *International Journal of Chemical Studies, 8*(2), 1148–1154. https://doi.org/10.22271/chemi.2020.v8.i2r.8923

[10] Agbadede, R., & Kainga, B. (2021). Performance and techno-economic analysis of inlet fogging system implementation in heavy duty industrial gas turbines. *International Journal of Thermodynamics, 24*(2), 75–81. https://doi.org/10.5541/ijot.782485

[11] Sun, Z., Chen, S., Hu, Z., & Lei, D. (2022). Vibration response analysis of a gear-rotor-bearing system considering steady-state temperature. *Nonlinear Dynamics, 107*(1), 477–493. https://doi.org/10.1007/s11071-021-07024-8

[12] Smaisim, G. F., et al. (2022). Nanofluids: Properties and applications. *Journal of Sol-Gel Science and Technology, 104*(1), 1–35. https://doi.org/10.1007/s10971-022-05859-0

[13] Zamen, M., Kahani, M., Rostami, B., & Bargahi, M. (2022). Application of Al_2O_3/water nanofluid as the coolant in a new design of photovoltaic/thermal system: An experimental study. *Energy Science & Engineering, 10*(11), 4273–4285. https://doi.org/10.1002/ese3.1067

[14] Gajbhiye, S., & Dakre, V. (2023). Experimental analysis of hybrid air-conditioning system using Al_2O_3 nanofluid via heat & mass transfer characteristics. *P. Technology, 23*(2), 4711–4730. https://doi.org/10.5281/zenodo.7778371

[15] Hassan, A., et al. (2023). An experimental and numerical study on the impact of various parameters in improving the heat transfer performance characteristics of a water-based photovoltaic thermal system. *Renewable Energy, 202*, 499–512. https://doi.org/10.1016/j.renene.2022.11.087

[16] Balakrishnan, S., Suriya, M., Tamilarasan, P., & Thanigaivel, S. (2025). Impact of nanoparticle concentration on thermal properties of nanofluids in heat exchangers. *ICSGET 2025 - E3S Web of Conferences 619*, 05008. https://doi.org/10.1051/e3sconf/202561905008

[17] Mousmoulis, G. (2020). *Development of cavitation detection methodology for hydraulic turbomachines using experimental and numerical tools* (Doctoral dissertation).

[18] Abass, W. A. (2023). *Numerical investigation on effects of plain, perforated, and dimpled twisted tape inserts in a tube to flow and heat transfer characteristics* (Master's thesis, Karabuk University).

[19] Ajmi, R. E., Hussein, H. A., & Numan, A. H. (2025). Experimental investigation of inception cavitation phenomena effect in centrifugal pumps using hydrophone sensor. *Jordan Journal of Mechanical and Industrial Engineering, 19*(1), 147–155. https://doi.org/10.59038/jjmie/190110

[20] Peband, P. (2023). *B 11245*.

[21] Barakat, S., Ramzy, A., Hamed, A. M., & El Emam, S. H. (2019). Augmentation of gas turbine performance using integrated EAHE and fogging inlet air cooling system. *Energy, 189*, 116133. https://doi.org/10.1016/j.energy.2019.116133

[22] Barik, D., Chandran, S. S. R., Dennison, M. S., Raj, T. G. A., & Roy, K. E. R. (2023). Investigation on fluid flow heat transfer and frictional properties of Al_2O_3 nanofluids used in shell and tube heat exchanger. *International Journal of Photoenergy, 2023*. https://doi.org/10.1155/2023/6838533

Authors' contacts:

Alaa Jasim Abdulah
(Corresponding author)
University of Technology, Elctromechanical Eng. Dep.,
Al-sinaa Street, Area 72, Baghdad 19006, Iraq
eme.19.18@grad.uotechnology.edu.iq

Hashim A. Hussein
University of Technology, Elctromechanical Eng. Dep.,
Al-sinaa Street, Area 72, Baghdad 19006, Iraq

Abdul Jabbar Owaid
University of Technology, Elctromechanical Eng. Dep.,
Al-sinaa Street, Area 72, Baghdad 19006, Iraq

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250421095705

Original scientific paper
Received: 2025-04-21, Accepted: 2025-05-09

Perception Management of Product Photography through Digital Channel Modifications

Ratko Knežević*, Miroslav Mikota, Krunoslav Hajdek, Ivana Žiljak Stanimirović

Abstract: Product photography conveys information and influences consumer perception, and today, most product photographs are created through the reproduction of digital image data. The digital file—i.e., the product photograph—as a medium for transmitting information based on visual assessment, is something consumers tend to trust. Each reproduction of a digital file consists of blue, green, and red channels. This research paper explores the possibility of managing the perception of product photography by altering the blue, green, and red channels of the digital file, as well as the impact of these changes on the perceived acceptability of the product image. It also aims to determine the boundaries within which the realistic appearance is preserved and how these boundaries can be interpreted. The study is conducted on two groups of participants: expert evaluators and regular consumers, both of whom assess the acceptability of the images through visual evaluation. In addition to participant-based assessments, characteristic color tables are defined for each type of promotional photographic image, and CIE Lab color values and overall color differences (ΔE00) are determined. The research demonstrates that by intervening in the blue, green, and red channels of the digital file—i.e., by modifying them—it is possible to influence the perception of product photography. The results of the evaluations show varying thresholds of acceptability for changes in individual color channels of the digital image.

Keywords: digital image file; product photography; RGB channels; visual evaluations

1 INTRODUCTION

A photographic image as a medium can also be defined as a representational image that expresses a certain aspect of reality, i.e., it represents a sign that denotes a specific object. The fundamental classification used today regarding signs and their meanings is the one that divides them into iconic, indexical, and symbolic signs [1].

A type of photography that distinctly uses semantic properties and is considered a specific form of illustrative photography is advertising or promotional photography. Its primary role is to attract the attention of the consumer and, along with conveying information, to encourage the consumption of a particular service or product. As the most effective method of capturing attention, promotional photography often employs close-up shots, direct gaze, distortion of reality, and shock.

Types of promotional photographic images include: mood photography, persuasive photography, explanatory photography, and product photography [2]. As previously stated, product photography is one form of promotional photography that conveys information to the consumer and influences their perception. Naturally, the quality of visual information transmission is crucial, as influencing perception directly affects the consumer's decision-making process regarding the promoted product. Based on the promotional message, the core concept of integrated marketing communication is aimed at encouraging consumers to engage with a specific product. In order to present product photography of desired items in a more acceptable and informative manner through various promotional channels, image processing and reproduction standards emerge as key elements. The visual expression captures the viewer's attention and emotionally engages them, while product photography itself is essential for presenting different products that attract consumer attention and guide them in making decisions. The goal of reproducing a photograph—in this case, specifically product photography—is to achieve sales, considering that consumers spend 80% of their time focusing on the image and only 20% on the textual content [3]. The objective of product photography is to evoke specific emotions in the consumer that would lead to consumption, making clear communication between the reproduced image and the consumer highly important. We can agree that photography is the ideal medium to evoke a reaction in the viewer—in this case, the consumer—because it offers excellent possibilities for expression and, therefore, for conveying information and eliciting emotions. Product photography captures attention, establishes communication, and transmits information. Properly reproduced product photographs can encourage consumption and create a better user experience.

Research shows that by intervening in the red, green, and blue channels of the digital image file—or by manipulating them—it is possible to influence the perception of product photography [4]. The results of evaluations of product image files show varying thresholds of acceptability for changes in individual channels in relation to their impact on perception. The obtained results confirm the hypothesis regarding the importance of product photography and its role in attracting the consumer's attention [5]. The authors in this study analyse product photography and explore the limits of acceptable channel manipulation within a digital image file while maintaining a realistic representation of the product, as well as the possibilities for perception management.

2 EXPERIMENTAL PART

The aim of the experimental research is to explore the possibility of managing the perception of product photography through changes in the red (R), green (G), and blue (B) channels of the digital file of the promotional photographic image. The authors also investigate the impact of changes in individual channels on the acceptability assessments of the product photograph and establish the limits for maintaining a realistic representation, i.e., the iconic character. The iconic character of the promotional

product photograph is defined in relation to the product photograph defined as the original [6].



Figure 1 Product photograph original

For the purposes of the research paper, a product photograph was created and captured. The photograph was taken with a Canon EOS 5D Mark III photographic system, and its digital recording is defined in the finest JPEG format in the Adobe RGB color space. The digital file of the captured product photograph is then transferred to the reproducible sRGB color space for perceptual rendering, thus defining the original product photograph [7]. The digital file of the original product photograph is processed by changing the values of the RGB channels in the range of ±12%, with a step of change or degradation of 3% over the entire image, using Adobe Photoshop image editing software (Fig. 2).

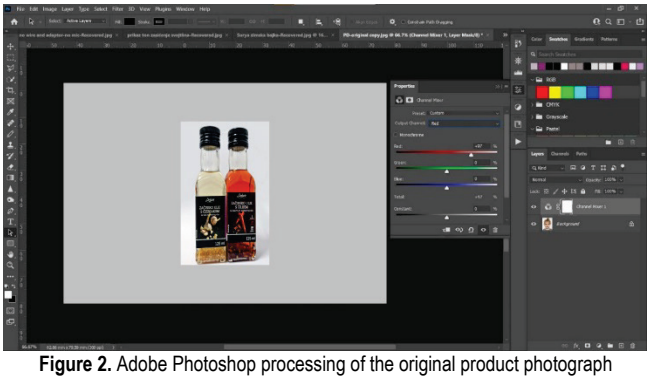


Figure 2. Adobe Photoshop processing of the original product photograph



Figure 3 Degraded product photograph of the original within a range of ±12% with a change step of 3%.

After degrading the original product photograph for each individual RGB channel within a range of ±12% with a change step of 3%, we obtained 24 degraded photographs (Fig. 3).

All photographs were prepared on a neutral gray background adapted to the screen format on which the promotional product photograph assessment took place. To ensure the best possible reproduction quality, the dimensions of the photographic images were adjusted to fit the format of the monitors used in the study, TCL 55"P615 4K Android TV. Each degraded product photograph was assigned a code consisting of a combination of letters and numbers, so that when comparing, we could accurately determine which photograph met the given criteria [7]. After defining the original product photograph and creating the set of degraded images, the experimental research was conducted on two groups of participants. The first group consisted of experts with experience in creating and selecting promotional photographs, specifically product photography, while the second group consisted of consumers of promotional photographs who had less professional knowledge, namely students. The expert group consisted of 5 participants, while the consumer group consisted of 44 participants with an average age of 21, with 22 males and 22 females. All participants underwent the Ishihara color recognition test [8]. The interpretation of the degraded product photograph samples was carried out using the method of simultaneous comparison for experts and memory comparison for consumers with less professional knowledge. Participants assessed the product photographs on a scale from 1 to 5, indicating which of the degraded photographs was considered least acceptable (1) to most acceptable (5), where a highly acceptable product photograph was rated as 5 [5]. The online questionnaire was created using Google Forms, and participants accessed it via a link sent to their mobile devices. During the evaluation of the product photographs, data were recorded on a server and were subsequently used for further processing in the research.

In the course of the study, a characteristic color table for product photography was created, which identified three specific colors marked with red dots and one white dot, indicating the specific locations for measuring and taking the specific colors, as shown in Fig. 4.



Figure 4 Display of the measurement locations for specific colors in product photography.

In addition to the three specific colors of the product photograph obtained by measurement at specific locations, named "Oil 1," "Oil 2," and "Glass," the color table also contained standard primary and achromatic colors from the standard color table (Fig. 5).

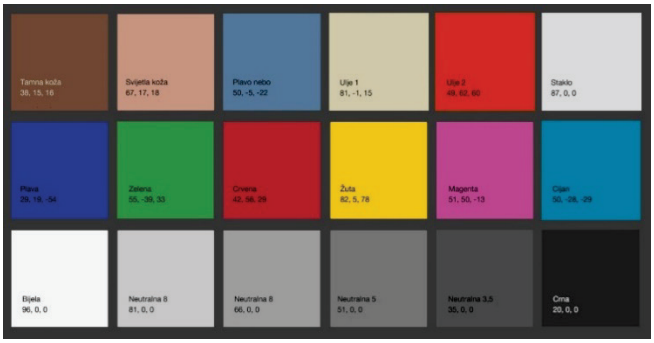


Figure 5 Characteristic color table with the three specific colors from the product photograph and other standard primary and achromatic colors.

The characteristic color table for product photography was then degraded in the RGB channels by ±3 to 12% for each channel individually, meaning the R channel from ±3% to 12%, followed by the G channel from ±3% to 12%, and the B channel from ±3% to 12% [9, 10]. In the next phase of the research, color changes in the digital files of the characteristic color tables previously defined for product photography are determined. Based on the determination of CIE Lab* values in Photoshop 2020, the total color differences ΔE00 for each field of the characteristic color table are calculated [11, 12].

3 RESULTS AND DISCUSSION

The results of the research conducted by the authors on expert participants, using the method of simultaneous comparison of the original product photograph and the degraded product photographs, were rated by the participants on a scale from 1 to 5, where 1 represents a completely unacceptable reproduction and 5 represents a highly acceptable reproduction. These results are presented in Tab. 1. The evaluation results of the research conducted on consumers with less professional knowledge, who used memory comparison [13] to compare the original product photograph with the degraded product photographs, are shown in Tab. 2.

As shown in Tab. 1, the assessments from five experts under standard observation conditions using memory comparison reveal varying acceptance thresholds for changes in individual channels of the digital file for the observed photographs. These results strongly correlate with the findings from the consumer study, presented in Tab. 2, conducted on consumers with less professional knowledge.

For the characteristic color table created for the tested product photograph (Fig. 5), Lab values were taken, and the average ΔE00 was calculated. In this research, the tables showing the Lab values of specific colors, as well as the calculations of their average ΔE00, were retained. Below are the average ΔE00 (Lab) values specifically for each color channel (R, G, and B), the absolute size of the change (3%,

6%, 9%, 12%), and the direction of the change (positive or negative).

Table 1 Expert participants' ratings for the product photograph with code P on a scale from 1 to 5, where 1 is completely unacceptable and 5 is highly acceptable.

Image code	Expert 1's Response	Expert 2's Response	Expert 3's Response	Expert 4's Response	Expert 5's Response	Degradation
P1	5	3	2	4	5	–3% R
P2	4	2	3	1	5	–6% R
P3	4	1	4	1	4	–9% R
P4	3	1	5	1	3	–12% R
P5	4	5	2	4	1	+3% R
P6	4	3	3	1	4	+6% R
P7	4	1	4	1	2	+9% R
P8	4	1	4	1	1	+12% R
P9	4	1	2	4	3	–3% G
P10	3	3	3	1	2	–6% G
P11	3	2	4	1	1	–9% G
P12	3	1	5	1	1	–12% G
P13	4	2	2	4	5	+ 3% G
P14	3	2	3	1	4	+ 6% G
P15	3	1	4	1	3	+9% G
P16	3	1	5	1	2	+ 12% G
P17	4	5	2	1	3	–3% B
P18	3	4	2	1	5	–6% B
P19	3	2	4	1	4	–9% B
P20	3	2	5	1	1	–12% B
P21	4	5	2	4	5	+ 3% B
P22	4	5	2	1	5	+6% B
P23	4	2	3	1	4	+ 9% B
P24	4	1	4	1	4	+ 12% B

Table 2 Consumer participants' ratings with less professional knowledge for the product photograph with code P, where 1 is completely unacceptable and 5 is highly acceptable.

1	Image code																																				Degradation
2	P1	3	3	4	5	5	5	3	2	5	4	5	2	5	5	5	5	3	5	4	5	5	4	5	4	5	5	5	5	4	4	5	3	3	4	5	5	4	5	4	4	3	- 3% R
3	P2	4	2	3	5	5	4	3	3	3	2	4	3	4	4	3	3	3	4	2	5	3	3	3	3	5	4	5	3	3	3	4	3	2	4	4	5	3	3	3	3	2	- 6% R
4	P3	4	2	3	4	5	3	2	3	2	2	3	3	2	1	3	2	2	3	1	3	1	2	4	2	2	3	4	1	1	2	2	2	3	3	3	4	2	3	1	1	1	- 9% R
5	P4	3	1	2	2	4	2	2	3	2	2	3	3	1	1	3	2	2	2	1	2	1	2	4	2	1	2	3	1	1	1	1	2	5	1	2	4	2	2	1	1	1	- 12% R
6	P5	5	1	5	4	3	4	5	2	5	5	5	5	5	5	5	5	4	5	5	5	3	3	5	5	5	5	5	5	3	3	5	4	5	4	5	5	4	5	5	4	5	+3% R
7	P6	3	1	4	4	3	4	4	2	4	5	4	5	3	3	3	5	2	4	3	4	3	4	4	4	4	4	5	5	3	3	5	5	4	4	4	5	3	5	5	3	4	+6% R
8	P7	3	1	4	2	3	1	4	2	2	3	4	4	2	1	2	5	2	4	2	4	3	5	4	3	3	3	5	4	3	3	4	5	3	3	4	5	3	3	3	1	3	+9% R
9	P8	3	2	3	2	3	1	4	3	2	2	4	4	1	1	3	5	3	3	1	4	3	3	4	1	2	5	4	3	2	2	3	4	3	2	3	4	2	1	1	1	2	+12% R
10	P9	4	3	5	5	3	5	5	3	5	5	5	3	5	5	5	5	4	5	4	4	2	5	5	4	4	5	5	5	3	2	5	3	5	4	5	5	5	5	5	4	5	- 3% G
11	P10	4	1	4	2	4	2	3	2	2	4	4	3	4	2	5	5	2	4	2	3	4	3	4	3	3	3	5	3	4	4	5	3	4	3	4	4	2	4	4	2	1	- 6% G
12	P11	2	1	3	1	2	1	3	2	2	3	3	3	3	1	4	5	3	3	1	2	3	1	3	2	2	3	5	1	2	2	3	2	1	2	2	3	2	2	1	1	1	- 9% G
13	P12	4	1	2	1	3	2	2	2	2	1	2	3	1	1	3	3	2	2	1	1	2	1	3	2	1	3	4	1	2	1	2	1	1	1	2	2	2	1	1	1	1	- 12% G
14	P13	4	2	5	4	3	3	3	4	4	3	5	3	5	4	3	5	3	5	4	1	3	3	5	5	5	5	4	4	5	2	4	3	5	4	5	5	3	4	3	5	4	+ 3% G
15	P14	3	1	4	4	4	2	2	4	2	3	5	3	4	3	3	3	3	4	2	5	2	2	5	4	4	3	5	3	5	2	3	3	2	3	4	4	2	2	2	4	3	+ 6% G
16	P15	3	1	3	1	2	1	2	3	2	2	3	3	3	1	2	3	3	4	1	4	1	2	4	3	4	3	4	2	4	1	3	3	1	2	3	3	2	1	1	1	2	+9% G
17	P16	4	1	2	1	3	1	1	3	1	2	3	4	1	1	2	2	3	3	1	1	1	2	4	2	3	3	3	1	2	1	2	2	1	1	3	3	2	1	1	1	2	+ 12% G
18	P17	4	4	4	5	3	4	4	4	5	5	5	4	5	5	4	5	5	5	5	2	5	4	5	5	5	5	5	5	5	4	5	4	5	4	5	5	4	5	5	5	5	- 3% B
19	P18	4	4	4	5	4	4	3	3	5	5	5	4	5	4	3	5	5	5	4	1	5	5	5	5	5	4	5	5	5	4	5	4	4	3	5	5	4	5	5	5	5	- 6% B
20	P19	5	1	2	2	3	1	4	4	2	3	3	4	2	3	2	2	5	3	3	5	2	2	4	2	4	3	4	2	4	2	3	2	3	2	4	3	3	4	1	5	2	- 9% B
21	P20	3	2	4	1	3	5	4	5	3	4	4	3	4	4	3	5	4	4	4	3	2	4	2	5	2	4	2	3	1	2	1	2	1	3	2	3	3	1	3	2	- 12% B
22	P21	4	3	5	5	2	5	5	5	4	4	5	3	5	5	3	5	5	4	4	1	5	3	5	5	5	5	5	3	4	4	4	3	5	5	5	5	4	5	5	5	4	+ 3% B
23	P22	4	2	4	5	3	5	5	5	3	5	5	3	3	5	4	5	5	4	3	5	5	2	4	5	5	5	5	2	4	4	4	3	5	5	4	5	3	5	5	2	2	+6% B
24	P23	4	2	4	4	3	5	5	5	2	5	5	3	3	5	3	5	5	5	3	2	4	4	3	4	4	4	4	5	3	3	2	4	3	5	5	3	5	2	4	4	2	+ 9% B
25	P24	4	1	4	4	4	4	5	5	2	5	5	3	3	5	4	5	5	5	3	1	3	3	3	4	4	3	4	4	2	2	1	4	3	5	4	3	5	3	5	3	2	+ 12% B

Table 3 Product photograph R channel -3%

	Original			Measured Sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	81.00	–3.00	15.00	2.4012
Oil2	49.00	62.00	60.00	47.00	60.00	59.00	2.0567
Glass	87.00	0.00	0.00	87.00	–2.00	–1.00	2.9522

Table 4 Product photograph R channel -6%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	80,00	–6.00	13.00	6.1869
Oil2	49.00	62.00	60.00	46.00	58.00	57.00	3.1432
Glass	87.00	0.00	0.00	86.00	–4.00	–1.00	5.3893

Table 5 Product photograph R channel -9%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	80.00	–8.00	12.00	8.4329
Oil2	49.00	62.00	60.00	45.00	56.00	5600	4.2444
Glass	87.00	0.00	0.00	86.00	–7.00	–2.00	8.6382

Table 6 Product photograph R channel -12%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	79.00	–11.00	12.00	11.0688
Oil2	49.00	62.00	60.00	44.00	53.00	55.00	5.5115
Glass	87.00	0.00	0.00	85.00	–9.00	–3.00	10.6199

Table 7 Product photograph R channel +3%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	82.00	2.00	16.00	3.7807
Oil2	49.00	62.00	60.00	50.00	64.00	61.00	1.1439
Glass	87.00	0.00	0.00	88.00	2.00	1.00	3.0211

Table 8 Product photograph R channel +6%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	82.00	3.00	18.00	5.0400
Oil2	49.00	62.00	60.00	51.00	66.00	63.00	2.2559
Glass	87.00	0.00	0.00	89.00	5.00	2.00	6.7296

Table 9 Product photograph R channel +9%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	83.00	6.00	18.00	8.2786
Oil2	49.00	62.00	60.00	53.00	68.00	64.00	4.2779
Glass	87.00	0.00	0.00	89.00	7.00	2.00	8.7080

Table 10 Product photograph R channel +12%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	83.00	8.00	19.00	10.1672
Oil2	49.00	62.00	60.00	53.00	69.00	65.00	4.3714
Glass	87.00	0.00	0.00	90.00	9.00	3.00	10.7108

Table 11 Product photograph G channel -3%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	80.00	1.00	14.00	2.6757
Oil2	49.00	62.00	60.00	48.00	62.00	60.00	0.9929
Glass	87.00	0.00	0.00	86.00	3.00	–3.00	4.8648

Table 12 Product photograph G channel -6%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	78.00	5.00	11.00	8.6540
Oil2	49.00	62.00	60.00	48.00	63.00	60.00	1.0686
Glass	87.00	0.00	0.00	84.00	7.00	–5.00	9.4203

Table 13 Product photograph G channel -9%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	77.00	7.00	9.00	11.8633
Oil2	49.00	62.00	60.00	48.00	63.00	60.00	1.0686
Glass	87.00	0.00	0.00	82.00	10.00	–7.00	12.4887

Table 14 Product photograph G channel -12%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	75.00	11.00	7.00	16.9921
Oil2	49.00	62.00	60.00	48.00	64.00	60.00	1.2658
Glass	87.00	0.00	0.00	81.00	13.00	–10.00	15.1582

Table 15 Product photograph G channel +3%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	83.00	–4.00	18.00	3.8604
Oil2	49.00	62.00	60.00	49.00	61.00	60.00	0.3997
Glass	87.00	0.00	0.00	89.00	–3.00	2.00	4.6134

Table 16 Product photograph G channel +6%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	84.00	–7.00	20.00	6.7774
Oil2	49.00	62.00	60.00	49.00	61.00	60.00	0.3997
Glass	87.00	0.00	0.00	91.00	–6.00	4.00	8.4478

Table 17 Product photograph G channel +9%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	86.00	–10.00	22.00	9.4555
Oil2	49.00	62.00	60.00	49.00	60.00	60.00	0.8042
Glass	87.00	0.00	0.00	93.00	–9.00	7.00	11.9255

Table 18 Product photograph G channel +12%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	88.00	–13.00	25.00	11.7740
Oil2	49.00	62.00	60.00	49.00	60.00	60.00	0.8042
Glass	87.00	0.00	0.00	94.00	–12.00	10.00	14.6932

Table 19 Product photograph B channel –3%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	81.00	–2.00	18.00	1.9940
Oil2	49.00	62.00	60.00	49.00	62.00	60.00	0.0000
Glass	87.00	0.00	0.00	87.00	0.00	3.00	2.8103

Table 20 Product photograph B channel –6%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	81.00	–2.00	21.00	3.4155
Oil2	49.00	62.00	60.00	49.00	62.00	60.00	0.0000
Glass	87.00	0.00	0.00	87.00	–2.00	6.00	5.8282

Table 21 Product photograph B channel –9%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	81.00	–3.00	24.00	5.0525
Oil2	49.00	62.00	60.00	49.00	61.00	60.00	0.3997
Glass	87.00	0.00	0.00	87.00	–2.00	10.00	8.4536

Table 22 Product photograph B channel –12%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	81.00	–3.00	26.00	5.9025
Oil2	49.00	62.00	60.00	49.00	61.00	60.00	0.3997
Glass	87.00	0.00	0.00	87.00	–3.00	13.00	10.5024

Table 23 Product photograph B channel +3%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	81.00	0.00	13.00	1.7292
Oil2	49.00	62.00	60.00	49.00	62.00	60.00	0.0000
Glass	87.00	0.00	0.00	88.00	1.00	–3.00	3.1840

Table 24 Product photograph B channel +6%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	82.00	0.00	10.00	3.4781
Oil2	49.00	62.00	60.00	49.00	62.00	60.00	0.0000
Glass	87.00	0.00	0.00	88.00	2.00	–7.00	6.5329

Table 25 Product photograph B channel +9%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	82.00	1.00	8.00	5.3846
Oil2	49.00	62.00	60.00	49.00	62.00	60.00	0.0000
Glass	87.00	0.00	0.00	88.00	2.00	–10.00	8.4779

Table 26 Product photograph B channel +12%

	Original			Measured sample			
	L1	a1	b1	L2	a2	b2	ΔE00
Oil1	81.00	–1.00	15.00	82.00	1.00	5.00	7.4845
Oil2	49.00	62.00	60.00	49.00	62.00	60.00	0.0000
Glass	87.00	0.00	0.00	88.00	4.00	–13.00	10.8432

After conducting the research and analyzing the results related to the experiment, observing the main intervention, i.e., the degradation of the RGB channels by 3% within the range of ±12%, it was concluded that the acceptability ratings for the product photographs by the consumer participants with less professional knowledge mostly fall within the range when ΔE00 is equal to or less than 6.

4 CONCLUSION

The conducted research shows that in relation to a photographic image defined as the original, interventions and manipulations in digital channels lead to noise in communication, i.e., degradation of the information that the photographic image transmits. Such interventions and manipulations can be performed throughout all stages of the digital system, with maximum control over the effect being achieved during the processing phase of the digital image file, where each channel can be adjusted individually while maintaining the values of the other channels. The research also indicates that by changing the values of the blue, green, or red channels, the realism or iconicity of the photographic image can be preserved. Most of the research conducted so far has focused on changes in the values of the channels in the digital recording of portrait images, which, in the context of promotional photography, can also be viewed as potential "persuasive" images, landscape photography, which in promotional photography can be viewed as potential "mood" images, and food photography, which in promotional photography can be viewed as potential "catalog" images. The most challenging photographic subject by this criterion has proven to be the "persuasive" photographic image. Maintaining the iconic character, or realism, is associated with the boundaries of the overall change in standard skin colors. Studies conducted on the digital file of promotional product photography show that the optimal regulator of perception and interpretation of promotional photographic images during the digital image processing phase is the regulation of enhancing and weakening channels within the boundaries of ±6%.

For other types of photographic images studied, the primary link between the perception of dominant or isolated colors and changes in the values of the digital channels within the range of ±9% was demonstrated. This cannot be tracked using standard color tables used for evaluating the technical reproduction of colors in all phases of the digital

photographic system. It was also determined that small changes in colors, within acceptable limits, based on the criteria of determining the total color difference (ΔE00), affect the perception of the photographic image.

5 REFERENCES

[1] Mikota, M. (2000). *Kreacija fotografijom*. ISBN 953-6838-01-X. (in Croatian)

[2] Mikota, M. (2016). Fotografija – Tehnički i semantičko-sintaktički izazovi medija u novom okruženju. *Printing & Design*, 175–178. (in Croatian)

[3] Pavlović, I. (2021). Impact of the channels changes on the interpretation of photographic image. *Doctoral dissertation*, Faculty of Graphic Arts. (in Croatian)

[4] Milković, M., Bolanča, S., Mrvac, N., & Zjakić, I. (2006). The influence of standard rendering methods on the manifested intensity of the chromatic induction effect. *Technical Gazette*, 13(1–2), 5–13.

[5] Periša, M., Mrvac, N., & Mikota, M. (2015). The "Visual Grammar" of photographic images produced by media convergence. In *Communication Management Forum Proceedings* (pp. 647–667). Zagreb.

[6] Hajdek, K., Matijević, M., & Mikota, M. (2017). Effect of ICC profile on reproduction. In *21st International Conference on Printing, Design and Graphic Communications* (pp. 21–34). Zagreb.

[7] Dragčević, K., Mikota, M., Pavlović, I., & Maričević, M. (2024). Dependence of the emotional experience of newspaper photographic illustration on the medium of realization. *Technical Journal*, 18(4), 560–567. https://doi.org/10.31803/tg-20231018125023

[8] Dragčević, K., Miroslav, M., & Dasović, A. (2019). Istraživanje doživljaja fotografske ilustracije ovisno o mediju prikaza. In *Proceedings PDC conference* (pp. 98–112). Zagreb: Hrvatsko društvo grafičara. (in Croatian)

[9] Mikota, M., Brozović, M., & Pavlović, I. (2008). Kvaliteta fotografske prezentacije modnih noviteta u medijima vanjskog oglašavanja. *Textile*, 57(9), 457–464. (in Croatian)

[10] Ahtik, J., & Starešinič, M. (2017). Eye movement analysis of image quality parameters compared to subjective image quality assessment. *Technical Gazette*, 24(6), 1833–1839. https://doi.org/10.17559/TV-20161213185321

[11] Mohammadi, P., Ebrahimi-Moghadam, A., & Shirani, S. (2014). Subjective and objective quality assessment of image: A survey. *arXiv*. https://arxiv.org/abs/1406.7799

[12] Mikota, M., Matijević, M., & Pavlović, I. (2016). Colour reproduction analysis of portrait photography in cross-media system: Image on the computer monitor–electrophotographic printing. *The Imaging Science Journal*, 64(6), 299–304. https://doi.org/10.1080/13682199.2016.1184875

[13] Pilarczyk, J., & Kuniecki, M. (2014). Emotional content of an image attracts attention more than visually salient features in various signal-to-noise ratio conditions. *Journal of Vision*, 14(12), 4–24. https://doi.org/10.1167/14.12.4

Authors' contacts:

Ratko Knežević,
(Corresponding author)
University of Zagreb Faculty of Graphic Arts
Getaldićeva 2, 10000 Zagreb, Croatia
E-mail: rknezevic@media-graphics.hr

Miroslav Mikota, PhD, Associate professor
University of Zagreb Faculty of Graphic Arts,
Getaldićeva 2, 10000 Zagreb
miroslav.mikota@grf.unizg.hr

Krunoslav Hajdek, PhD, Associate professor
University North,
Trg dr Žarka Domjana 1, 48000 Koprivnica, Croatia
khajdek@unin.hr

Ivana Žiljak Stanimirović, PhD, Full professor
University of Zagreb Faculty of Graphic Arts,
Getaldićeva 2, 10000 Zagreb, Croatia
ivanaziljakstanimirovic@grf.unizg.hr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20251021120858

Original scientific paper
Received: 2025-10-21, Accepted: 2026-08-07

Multi-Scenario Optimization of Fuel Distribution Transportation Using Goal Programming

David Billy Martin Salangka, Hwi-Chie Ho*

Abstract: Inefficient distribution routes of oil could lead to supply deficits and increased costs. Several goal programming models were developed in this study to determine an optimal distribution model for fuel oil transportation using goal programming that meets target demand while optimizing capacity, time, and cost. Data were obtained through sources directly involved in operational performance. A total of four scenarios reflecting variations in demand and route use were evaluated. The optimization results showed increased fuel fulfilment by 1,033 KL; 2,533 KL; 1,378 KL; 3,973 KL for Scenario 1 to 4, respectively. Also, the average travel time and distribution cost per litre were significantly reduced across all scenarios. The findings demonstrate that multi objective goal programming models enable in providing strategic allocation of transport resources in the networks of fuel-distribution. Moreover, implications for logistics managers include prioritizing route efficiency and modelling trade-offs between fulfilment, time, and cost in dynamic supply-chain nature.

Keywords: energy allocation; oil distribution; operational efficiency; optimization; supply chain

1 INTRODUCTION

The need for fuel oil has grown significantly over the last few decades due to several factors, including ongoing economic expansion, expanded access to personal vehicles, and global integration, which has raised transportation demands. Despite the recent rise in electric vehicles, conventional fuels are predicted to continue to be used at least until 2050 [1], which will have an impact on the environment. Even though the industrial sector uses a substantial quantity of energy globally, there are very few economic research on industrial energy demand [2]. Industries operating in hard-to-reach areas, such as Industrial Plantation Forests, face unique challenges in addressing this issue. Common problems include long distribution processes, high costs, limited access, and the socio-economic conditions of the surrounding communities.

Harvesting [3] is a production process that consists of a series of activities including Felling, Skidding, Transporting, and Loading/Unloading. The expense of fuel is the most expensive aspect of these activities. In Fig. 1, the fuel costs are as follows: 34% for felling, 36.9% for skidding, 39.3% for transporting (Pompong), 48.6% for transporting (Tronton), 54.3% for loading/unloading at TPn or Forest Landing Site, 54.3% for loading/unloading at TPK or Log Yard, and 45.3% for loading/unloading at the logpond. This suggests that the expense of fuel throughout the harvesting process needs to be given careful consideration.

All operational activities, including nursery, planting, and road construction, would be impacted by the up to 31% increase in the harvesting target, according to the company's report data. This will also affect the company's operations, which would require more gasoline. The goal planting numbers, the standard age of the plants during harvest, and the age of the standing trees in the field are used to calculate the expected increase in the Wood Delivery goal.

There would be a sharp rise in fuel consumption, especially in a few key operational and supporting activities including road construction and plantation operations. According to the Standard Calculation of Estimated Fuel Consumption in harvesting, the company's fuel consumption for harvesting activities is 8 Liters for every ton of wood harvested. For planting activities, fuel consumption is assumed to be 74% of the harvesting consumption, as shown in Tab. 1. Furthermore, weekly and annual fuel use will undoubtedly be impacted by road construction fuel consumption.

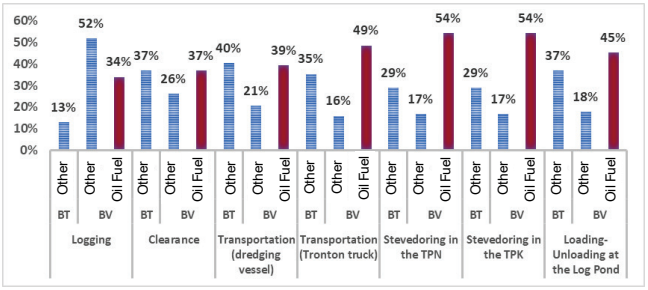


Figure 1 The percentage of fuel costs compared to other costs in each operational activity

Table 1 Monthly fuel needs for the years 2021-2025 (Tons)

No.	Business Unit	2021	2022	2023	2024	2025
1	Harvesting	3,733,333	3,993,246	4,052,060	4,463,665	5,167,424
2	Plantation	2,758,750	2,954,325	2,998,524	3,302,554	3,822,860
3	Road construction	222,917	396,250	449,110	476,770	504,430
	Total	6,715,000	7,343,821	7,499,694	8,242,989	9,494,714

Table 2 Number of non-productive heavy equipment and potential loss

No.	Date	Total unproductive heavy equipment (units)	Potential losses (tons)
1	09/22/2022	11	440
2	09/23/2022	22	880
3	09/28/2022	3	120
4	10/02/2022	10	400
5	10/13/2022	5	200
	Total	51	2,040

Based on data and reports collected over a two-month period, from September to October 2022, some operational harvesting regions have had substantial challenges hurting Wood Shipping production. According to statistics from two districts, a total of 28 heavy equipment units had their production impeded while on standby for fuel delivery. The company will undoubtedly lose opportunities because of

heavy equipment's inefficiency. Based on the results of the time study conducted by the Harvest Management team, the productivity of one heavy equipment unit is approximately 40 tons per day, thus the company's expected losses, as summarized in Tab. 2.

The data on unproductive heavy machinery is merely a sample of operational data collected from two districts over a two-month period. Given the comparatively lower fuel production and consumption targets in 2022 compared to subsequent years, the corporation is likely to face more potential challenges if changes are not implemented quickly. These enhancements are designed to optimize fuel delivery, both in terms of timing and prices. The many distribution channels and hard-to-reach places will undoubtedly affect access time and distribution costs, therefore the option of using multimodal transportation in its distribution is critical.

Furthermore, in the plantation operations, fuel oil distribution is a crucial factor in the supply chain, where inefficiencies in routing, scheduling and resource allocation could lead to significant deficits and cost inefficiency. Although optimization techniques in the logistics and supply chain area have advanced considerably, many existing studies focus on single objective models, such as minimizing cost or delivery time alone. For example, [4] used a single objective Mixed-Integer Linear Programming (MILP) model to minimize total costs in solving closed-loop supply chain, location-allocation, transportation, and supplier selection problems.

Although many studies investigate logistics and supply chain optimization, a gap remains in models that integrate multiple fuel distribution scenarios such as varying demand, route cycles with simultaneous objectives of fulfilment, time, and cost. The objective of this study is to find optimal scenarios for petroleum distribution routes using goal programming mathematical models to meet target demands, as well as to generate distribution route scenarios using existing resources to optimize time, costs, and available transport capacity. This paper contributes by 1) developing a goal programming model for multi-scenario fuel distribution, 2) analyzing trade-offs between three key objectives, and 3) applying the model to real plantation data to provide practical decision-support for route and resource allocation.

2 LITERATURE REVIEW

2.1 Concept of Optimization

An action or activity that aims to produce the best or most ideal results is called optimization. The goal of optimization is to minimize the amount of work or resources needed, while they can also be used to maximize the desired outcomes [5]. When something is accomplished as efficiently and maximally as possible, it is said to be optimized. Optimization can also be interpreted as a standard by which to measure whether the actions taken satisfy the criteria and needs [6]. There must be a declaration about the purpose of the choice variables used in the optimization procedure. This is used to define or explain the procedure for determining the function's maximum or minimum value. Investigating operations is one of the steps or procedures that must be followed when performing optimization activities to solve an optimization problem [7].

2.2 Distribution

Distribution is the process of delivering goods or products to customers after the sales process [8]. Distribution is the transfer of commodities or products from the manufacturer to consumers or users via distribution channels, as well as the reverse flow of payments between buyers and sellers or back to the original provider [9]. Distribution is defined as marketing operations that facilitate the distribution of goods or products from sellers to customers [10]. Thus, the distribution procedure refers to the steps or series of marketing activities for products to customers that are repeated.

The implementation method of distribution contains numerous alternatives or types of channels that can be used, and these possibilities can be classified as consumer goods and industrial goods. First, consumer goods are items or objects purchased by consumers for use or consumption. Customer products are acquired based on the preferences or needs of the user or end customer, not the manufacturer, because the purchased commodities are no longer goods to be produced. Second, industrial goods are those that, once purchased, will be further processed with the intention of fulfilling the interests and demands of the relevant industry. Buyers of industrial commodities include businesses, institutions, and organizations.

2.3 Oil Fuel

Fuel is a type of material that can be transformed into energy. Fuel contains thermal energy that can be converted into energy for specialized applications, such as those used in automobiles or engines. Fuel types are classified according to their content or octane rating, which is divided into solid fuel and liquid fuel. Solid fuels are employed during combustion operations. Examples of solid fuels include wood, coal, and uranium. Liquid fuel is a mixture of hydrocarbon compounds acquired naturally or through artificial means, typically from petroleum. Oil fuel is one type of fuel, specifically a liquid fuel with a looser structure than solid fuel. Its molecules move freely. Liquid fuels include gasoline, kerosene, and diesel. Liquid fuels are commonly used in industrial sector [11].

2.4 Transportation

In industrial activities, Operations and Supply Chain Management (OSCM) continues to evolve by adopting various strategies, frameworks, and technologies to address increasingly complex challenges across industries [12–14]. It plays a vital role for all companies and influences many areas of their activities [15]. As one of the important aspects in OSCM, transportation positioned as an important role in ensuring the smooth flow of goods, services, and information throughout the supply chain.

Transportation is the movement of commodities or persons from one location to another. Transportation begins at the origin (where the initial activity occurs) and ends at the destination. Transportation is critical in locations that supply raw materials, market, and produce goods [16]. The transportation system can be defined as a system that

provides facilities and infrastructure, as well as services that enable activities throughout the region, allowing for effective movement and mobility from one location to another [17].

Transportation services are one of the systems that make up transportation. Transportation companies produce transportation services such as shipping, airline, and bus transportation, among other things. Transportation services are one of the input variables for operations such as manufacturing, trading, and others. Transportation services are closely related to the transportation process [18]. Transportation promotes and stimulates development while also contributing to a country's economic growth. As a result, transportation infrastructure is created before any other development projects. Transportation works best when journey time is kept to a minimum, accidents are minimized, service frequency is acceptable, and passengers are comfortable.

Existing study in transportation and logistics have adopted quantitative modelling approaches to optimise cost, time and fulfilment objectives as [19] for an example in short-haul airfare markets. Our study extends the optimisation framework to fuel distribution in the plantation sector.

2.5 Goal Programming Method

Goal programming extends the linear programming technique [20]. It is one of many mathematical methods used in decision-making to analyze and develop optimal solutions or resolutions incorporating numerous objectives, with the goal of obtaining the most optimal solution [21, 22]. The basic approach used in the goal programming process is to determine and set goals expressed numerically or, in other words, mathematically, to compile and formulate a function that becomes the objective function of the optimization problem, and to seek a solution or solutions by minimizing or optimizing the amount of deviation from the objective function [23].

2.6 Related Works

A number of scholars, particularly those working with optimization models, have expressed serious concerns about the distribution problem. Aside from the type of goods, this literature review identifies several distribution objectives in optimizing the distribution process, including maximizing profits [24, 25], supplier selection and allocation [26], fairness and effectiveness in fulfilling demand [27], multi-objective green logistics optimization [28], and maximizing distance and modes of transportation [29]. In addition, a study of [30] applied a multi-criteria approach to optimise last-mile distribution networks by considering cost, time, and resource utility. However, because distribution optimization has numerous objectives, goal programming optimization models are widely utilized.

The study [31] compared the optimization methodologies of Analytical Hierarchy Process (AHP) and goal programming, with a particular emphasis on supplier selection and evaluation processes aimed at attaining an optimal order quantity. The findings showed that using the interactive model method, namely the goal programming method, yields the best outcomes. Similar results were found in a study [32] comparing goal programming and AHP approaches for identifying optimal solutions in bank asset liability management. The study's findings revealed that the goal programming model is an effective and simple way for setting targets and limits to produce optimal solutions.

According to this literature review, the following trends could be found. First, several researches have proven that goal programming optimization models are effective in achieving various objectives of the problem. This is supported by the improved results gained as well as the overview provided by this strategy. Second, no researchers (to the best of our knowledge) have developed goal programming models for the problem of distributing fuel via intermodal transportation. As a result, the research described in this paper represents a novel contribution in a topic with important economic contributions and repercussions for the Industrial Plantation Forest Company.

3 METHODS

The purpose of this study was to develop an optimization model for the distribution of fuel oil using goal programming to attain multiple objectives. The research methodology is designed to systematically determine the most optimal distribution scenarios by considering the scientific principles and addressing the previously defined problem. Data collected from the Industrial Plantation Forest Company were processed and analyzed based on established theoretical foundations.

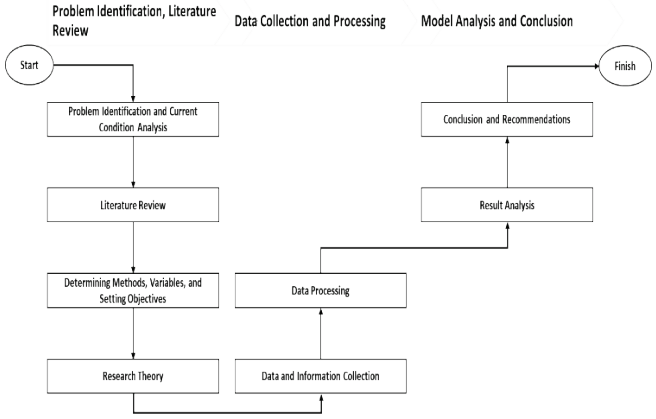


Figure 2 Illustration of the research flowchart

3.1 Research Flowchart

The research flowchart of this study consisted of three main phases as illustrated in Fig. 2. First, we performed problem identification and literature review. This stage provides a foundation for the research direction, where the research methods and variables were determined based on a reliable and compatible theoretical basis. Consequently, in the subsequent stages, the collected data must support the chosen research methods. Second, data collection was carried out. The collected data was ensured to support the research methods. Once the data was obtained, it was processed according to the mathematical model of the chosen research method. Lastly, the model resulting from data processing was

then analyzed to determine whether it surpassed the current conditions. If the model proved to be superior, the research has achieved its objectives.

3.2 Optimization Model using Goal Programming

The constraints and parameters used in the problem analysis process were expressed into mathematical models. The objective of this data modeling was to define the decision variables, the objective function, and the constraints of the identified problem. In this study, the initial model was Linear programming model, which was later converted into a goal programming model. There are some fundamental concepts of goal programming including [33]:

1) In a goal programming model, different objectives are expressed as goals. Thus, there will be multiple goals.
2) All goal constraints are expressed as equations that contain deviation variables: d^+ and d^-.
3) Positive deviation variable: d^+, represents the surplus value over the target level of the goal.
4) Negative deviation variable: d^-, represents the deficit value against the target level of the goal.
5) At least one (or both) of the deviation variables of the goal constraint must have a value of zero.
6) Deviation variables cannot be negative.
7) The objective function of all goal programming models is to minimize the deviations from all goal constraints levels.

In this study, several key assumptions underpin the goal-programming model. The decision variables were variables with unknown values. The decision variables were the West distribution route, which requires intermodal transportation, and the East distribution route, which uses a single mode of transportation. The objective function defined the goals to be achieved, which include fulfilling the monthly fuel needs, optimizing the distribution time, minimizing distribution costs, and maximizing carrying capacity of the fuel distribution transportation. Meanwhile, constraints were variables that serve as limits in achieving the objective function. The constraints in this study were the types of transportation used for fuel distribution. In addition, to ensure validity of the model, the model structure, parameter values, and decision rules were discussed by domain experts of the company who operate the fuel distribution system. In addition, the results from the goal programming model were also compared against historical distribution data.

The general steps of goal programming used in this study [34]:

1) Determination of the decision variables

There are seven parameters, including:

X_1 = West Route (Intermodal); X_2 = East Route (Direct/sea route only); P_1 = First Priority; P_2 = Second Priority; P_3 = Third Priority; d^- = Negative Deviation; d^+ = Positive Deviation.

2) Definition of the objectives and deviation variables for each goal

There were two objectives for fulfilling fuel needs, with two scenarios each being presented. However, in general, this scenario was used to compare two distribution routes in four different conditions. This was conducted to prove whether we must choose one distribution route or must use both routes on several trips. Scenario 1 was the fulfillment of the east demand target in 2023. Scenario 2 was the fulfillment of the highest east demand target in 2025. Furthermore, Scenarios 3 and 4 were the fulfillment of the west and east demand targets in 2023 and 2025, respectively.

a) Goals and priorities of Scenario 1:

1) Allocating 2 units of 500 KL Pontoons to pass through 2 routes (East and West) with priority to meet the minimum fuel needs of the East Area by at least 500,000 liters.
2) The company aims to distribute as quickly as possible given the absence of safety stock in each area. The maximum distribution time within 1 month is 28 days.
3) The company strives to minimize Distribution costs as much as possible.

b) Objective and priority of Scenario 2:

1) Allocating 2 units of 250 KL Pontoons to pass through 2 Routes (East and West), and 1 pontoon of 500 KL through the East route, with the priority to fulfill the minimum fuel needs of the East Area by at least 750,000 liters.
2) The company aims to distribute goods as quickly as possible considering the absence of safety stock in each area. The maximum distribution time within 1 month is 28 days.
3) The company aims to minimize Distribution costs as much as possible

c) Objective and priority of Scenario 3:

1) Allocating all available units to compare 2 Routes (East and West), with priority for fulfilling the fuel needs of the entire Area (West and East) minimally by 7,499,694 liters.
2) The company aims to distribute goods as quickly as possible given the absence of safety stock in each area. The maximum distribution time within 1 month is 28 days.
3) The company strives to minimize distribution costs as much as possible

d) Objective and priority of Scenario 4:

1) Allocating all available units to compare 2 Routes (East and West), with priority for fulfilling the fuel needs of the entire Area (West and East) minimally by 9,494,714 liters.
2) The company aims to distribute goods as quickly as possible given the absence of safety stock in each area. The maximum distribution time within 1 month is 28 days.
3) The company strives to minimize distribution costs as much as possible.

3) Formulation of the goal programming models

$$Z = X_1 + X_2 + d^- - d^+ \tag{1}$$

Minimize: $P_1 d_1^-$, $P_2 d_2^+$, and $P_3 d_3^+$

a) Scenario 1

1) Carrying Capacity

$$500{,}000X_1 + 500{,}000X_2 + d_1^- - d_1^+ = 1{,}286{,}609 \tag{2}$$

2) Time

$$7X_1 + 15X_2 + d_2^- - d_2^+ = 28 \tag{3}$$

3) Cost

$$103{,}900{,}000X_1 + 231{,}000{,}000X_2 + d_3^- - d_3^+ = 0 \tag{4}$$

b) Scenario 2

1) Carrying Capacity

$$750{,}000X_1 + 750{,}000X_2 + d_1^- - d_1^+ = 2{,}292{,}828 \tag{5}$$

2) Time

$$7X_1 + 15X_2 + d_2^- - d_2^+ = 28 \quad (6)$$

3) Cost

$$153,740,000X_1 + 346,500,000X_2 + d_3^- - d_3^+ = 0 \quad (7)$$

c) Scenario 3

1) Carrying Capacity

$$2,595,000X_1 + 2,595,000X_2 + d_1^- - d_1^+ = 7,499,604 \quad (8)$$

2) Time

$$7X_1 + 15X_2 + d_2^- - d_2^+ = 28 \quad (9)$$

3) Cost

$$501,950,000X_1 + 1,198,890X_2 + d_3^- - d_3^+ = 0 \quad (10)$$

d) Scenario 4

1) Carrying Capacity

$$2,595,000X_1 + 2,595,000X_2 + d_1^- - d_1^+ = 9,494,714 \quad (11)$$

2) Time

$$7X_1 + 15X_2 + d_2^- - d_2^+ = 28 \quad (12)$$

3) Cost

$$504,290,000X_1 + 1,198,890X_2 + d_3^- - d_3^+ = 0 \quad (13)$$

3.3 Data Collection

The data collection process was conducted to obtain information regarding the distribution process of fuel at PT EFG in South Sumatra Province. As a result, this research utilized secondary data. The collected data included fuel distribution from January to February 2023. Data collection began by gathering information about the background information through field observations, case studies, and data related to the fuel distribution process. The data to be processed in this study were those related to transport capacity (Tab. 4), distribution time (Tab. 5), and distribution costs (Tabs. 6-8).

Table 3 The indicator of research variable used in this study

Variables	Definition	Indicator
Transportation capacity (X_1)	The capacity of a facility or medium to transport people or products from origin to destination	1) Load quantity 2) Duration of travel 3) Trips 4) Target within one period
Delivery time (X_2)	The series of processes from the ongoing and future conditions, in this case, the process referred to is the distribution process of fuel	1) Distribution location 2) Quantity of shipments 3) Total cycle time
Distribution costs (X_3)	Sacrifices that are economic in nature measured in monetary units	1) Distance travel 2) Weighted of the load 3) Price calibration

Table 4 Transport capacity data

No.	Capacity (KL)	Quantity (units)
1	500	4
2	250	2
3	95	1
4	10	1
5	20	1

Table 5 Travel time data in one trip cycle

No.	Track	Time (days)	Description
1	West	7	Total cycle time
2	East	15	

Table 6 Transportation cost

No.	Track	Pontoon cost
1	West	IDR 190/Liter (Fuel Range 10.001-15.000)
2	East	IDR 462/Liter (Fuel Range 10.001-15.000)

Table 7 Calculation of freight costs from West to East (1/2)

Truck capacity	Fuel consumption		Rental truck (IDR)	Freight			Distance (km)
	Liter/km	Fuel price (IDR)		KL/month	Trip/week	KL/week	
	(a)	(b)	(c)	(d)	(e) = (d) : 4	(f) = (d) : (e)	(g)
10 KL	6.25	13,000	18,000,000	150	4	40	80
20 KL	10	13,000	20,000,000	288	4	77	80

Table 8 Calculation of freight costs from West to East (2/2)

Truck capacity	Round-trip fuel costs (IDR)	Total units per Target		Total weekly cost (IDR)	
		322	573	250	573(h)
	(h) = (a) × (b) × (g) × 2 × (e)	(i) = 322 : (f)	(j) = 573 : (f)	(k) = [(h) × (i)] + [(c) : 4]	(l) = [(h) × (j)] + [(c) : 4]
10 KL	1,248,000	9	15	15,732,000	23,220,000
20 KL	780,000	5	8	8,900,000	11,240,000

4 RESULTS AND DISCUSSION

4.1 Data Processing and Analysis

The data obtained in this study were analyzed using LINGO software. The data processing steps were divided into two steps: 1) formulating mathematical equations and entering the mathematical model into the LINGO, and 2) interpreting the output generated by LINGO.

In our study, four scenario models were examined to achieve optimal objectives by considering three key criteria (fuel fulfillment, travel time, and transportation costs) that minimize deviations from each criterion. The results of these models were expected to provide an optimal solution based on the priority scale stated in the minimization objectives.

The instructions @GIN(X1) and @GIN(X2) in LINGO output indicate that the variables X1 and X2 must be treated as integer variables. By defining X1 and X2 as integer variables, the model guarantee that the generated solutions were applicable in real-world scenarios, where distribution trips must be in whole numbers. Goal programming serves as a tool for making optimal decisions in situations with various priorities and complex constraints.

It should be noted that the finding from the four scenarios was the absence of a company-designated budget for distribution costs (D3_PLUS). Therefore, we decided to set the distribution cost target at 0 Rupiah indicating that any additional costs incurred were considered insignificant in affecting the final decision. This nominal value represents the optimal and reasonable budget recommended for expenditure which is aligned with the priorities established in the goal programming model.

a) Scenario 1. In Scenario 1, we analyzed the fuel fulfillment targets in the East region in 2023 using goal programming analysis in LINGO. The main objectives were to minimize the negative deviation from the fuel fulfillment target (D1_MIN), minimize the positive deviation from the

number of travel days (D2_PLUS), and minimize the positive deviation from the required costs (D3_PLUS) as shown in Fig. 3.

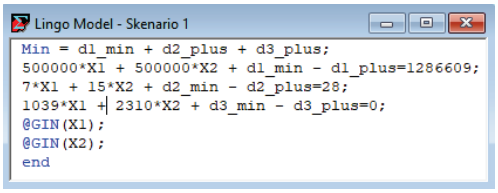


Figure 3 Mathematical equations for Scenario 1

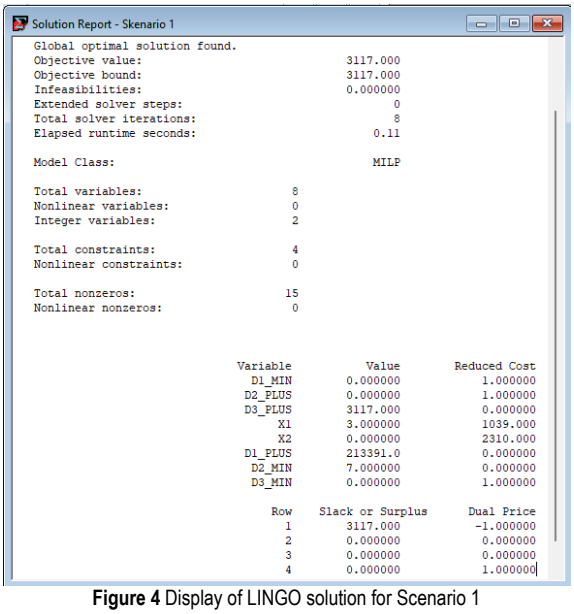


Figure 4 Display of LINGO solution for Scenario 1

The optimization results as shown in Fig. 4 revealed that:

1) Fuel Fulfillment Target. The fuel fulfillment target for the East region in 2023 is 1,286,609 liters. To achieve this target, three trips were planned in the West area (X1 = 3) and no trips were scheduled in the East area (X2 = 0).

2) Deviation from the Target. D1_MIN (negative deviation from the fuel fulfillment target) was 0, which means that the fuel fulfillment target was fully met. Next, D2_PLUS (positive deviation from the number of travel days) was 0, which indicates that there was no excess in the planned travel days. Lastly, D3_PLUS (positive deviation from the required costs) was 3117, which shows that the planned distribution cost is optimal and reasonable.

3) Distribution Costs. Although there was no budget that specifically allocated for distribution costs by the company, this study set the target cost at 0 to minimize D3_PLUS. This means that the calculated cost of Rp. 311,700,000 is considered optimal for achieving the best solution in the optimization scenario. Thus, the results of the goal programming model using LINGO showed an optimal distribution strategy that successfully met the fuel fulfillment target with high efficiency and minimize the deviations from the priority targets.

b) Scenario 2. In Scenario 2, we analyzed the fuel fulfillment targets in the East region in 2025. The main objectives of this model were to minimize the negative deviation from fuel fulfillment (D1_MIN), minimize the positive deviation from travel time (D2_PLUS), and minimize the positive deviation from transportation costs (D3_PLUS) as shown in Fig. 5.

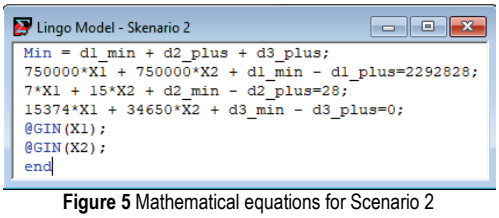


Figure 5 Mathematical equations for Scenario 2

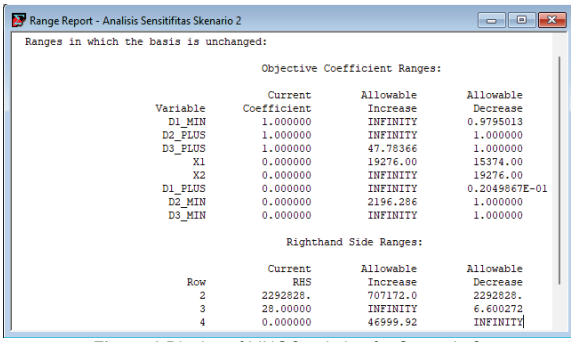


Figure 6 Display of LINGO solution for Scenario 2

Based on the LINGO output (Fig. 6), the optimal distribution strategy was to conduct four trips using the West area (X1 = 4) and no trips via the East area (X2 = 0). This strategy showed minimal deviation from the fuel fulfillment target (D1_MIN = 0) that indicates fuel needs were fully met. Minimal deviation in travel time (D2_PLUS = 0), which means the planned trips align well with the target schedule. Lastly, a positive deviation in transportation costs of Rp. 614,960,000 (D3_PLUS = 61,496), which represents the estimated optimal cost.

In addition, setting the distribution cost budget at 0 Rupiah as a target to minimize D3_PLUS guarantees that the results are both optimal and reasonable. This indicates that the estimated cost in the model already covers all relevant needs without requiring additional allocation from the company. Thus, the company could make efficient fuel distribution decisions based on this goal programming analysis to achieve the 2025 target with minimal costs and maximum fulfillment of fuel demand.

c) Scenario 3. In Scenario 3, we analyzed the fuel fulfillment targets in the East and West region in 2023. The main objectives of this model were to minimize the negative deviation from fuel fulfillment needs (D1_MIN), minimize the positive deviation from travel time (D2_PLUS), and minimize the positive deviation from transportation costs (D3_PLUS) as shown in Fig. 7.

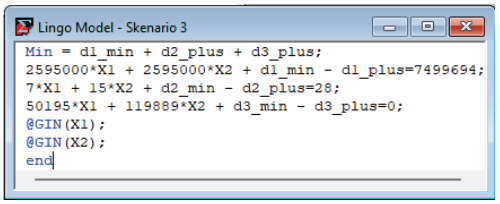


Figure 7 Mathematical equations for Scenario 3

The results showed that the optimal distribution strategy was to carry out three trips through the West area (X1 = 3) and no trips through the East area (X2 = 0). Deviation from the fuel fulfillment target (D1_MIN) was zero which means that fuel needs were met precisely as planned. Travel time was also optimal (D2_PLUS = 0) which indicates there was no excess travel time. In addition, there was a positive deviation in transportation costs of 150,585 or Rp. 1,505,850,000 (D3_PLUS) which indicates the cost exceeded the initial target (Fig. 8).

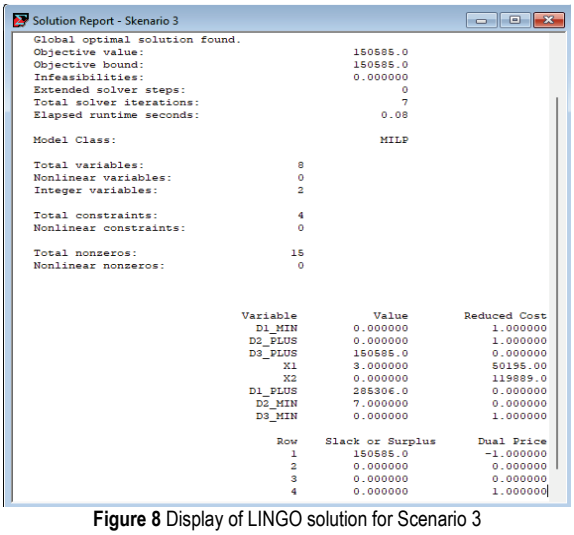


Figure 8 Display of LINGO solution for Scenario 3

With the distribution budget set at 0 Rupiah to minimize D3_PLUS, the costs incurred in this model were considered both optimal and reasonable. The company could use these findings to formulate efficient fuel distribution decisions that allows fuel needs to be met without excessive travel time, despite a slight deviation in transportation costs. This model served as a valuable guide for more effective resource allocation in 2023.

d) Scenario 4. In Scenario 4, we analyzed the fuel fulfillment targets in the East and West region in 2025. The main objectives of this model were to minimize the negative deviation from the fuel fulfillment target (D1_MIN), minimize the positive deviation in travel time (D2_PLUS), and minimize the positive deviation in transportation costs (D3_PLUS) as shown in Fig. 9.

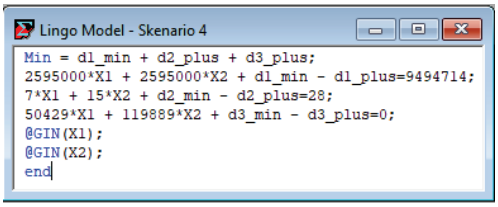


Figure 9 Mathematical equations for Scenario 4

The LINGO output (Fig. 10) displayed two decision variables, X1 and X2 represent the number of distribution trips through the West and East areas, respectively. The optimal results indicated that to efficiently meet the fuel distribution target, the best strategy was to conduct four trips through the West area and no trips via the East area were optimal.

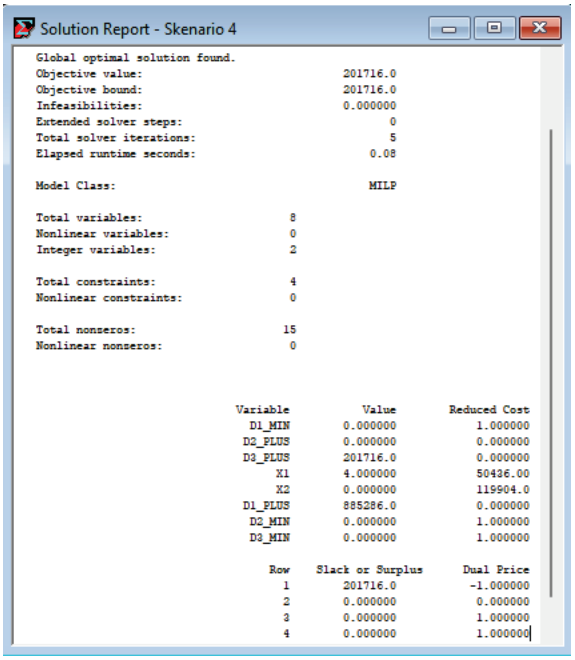


Figure 10 Display of LINGO solution for Scenario 4

A positive deviation in transportation costs (D3_PLUS) of Rp. 2,017,160,000 was identified, which indicates an increase in costs that should be considered in optimizing distribution trips. Meanwhile, the negative deviation from the fuel fulfillment target (D1_MIN) and the positive deviation in travel time (D2_PLUS) have been successfully minimized based on the established priorities. Thus, the goal programming results provide insights into how distribution trips should be allocated to minimize deviations from the set targets, with a focus on cost efficiency and optimal travel time.

4.2 Comparison of Proposed Models with Current Conditions

In this study, a comparison was also made between proposed models and current fuel distribution conditions. The planning and formulation of distribution routes were performed manually based on demand and utilize only one type of transportation. Fuel distribution was divided according to the same area within a single route; thus, time and cost aspects were not considered. The current distribution conditions could be observed in Tab. 9, which illustrates the fuel distribution allocation based on the area.

Tab. 9 summarized the fuel distribution using several pontoons in the West and East areas. In the West area, there were seven pontoons with various capacities and different departure schedules. For example, Pontoon 2 starts the first trip, followed by the other pontoons with a one-day gap from the previous one. Each pontoon has a total travel time of 7 days (including loading-unloading and round trips). The total fuel transported per month in the West area was 5,940 KL. Meanwhile, in the East area, there was only one pontoon (Pontoon 1) with a capacity of 250 KL and a total travel time of 15 days that allows it to transport 467 KL of fuel per

month. In total, the combined fuel transported from both areas per month was 6,407 KL.

Table 9 Current Condition of Fuel Distribution

Area	Poonton	Capacity (KL)	Travel duration*	Departure schedule	Trip	Date	Total transported per month
West	Poonton 2	250	7	0	3	3	750
	Poonton 3	500	7	+2 from Poonton 2	3	5	1,500
	Poonton 4	500	7	+2 from Poonton 3	3	7	1,500
	Poonton 5	500	7	+2 from Poonton 4	2	9	1,000
	Poonton 6	500	7	+2 from Poonton 5	2	11	1,000
	Poonton 7	95	7	+2 from Poonton 6	2	13	190
						Total:	**5,940**
East	Poonton 1	250	15	0	1.87	3	467
						Grand total:	**6,407**

*Total loading-unloading + Round trip travel

There was a distribution issue on the East Area route, which takes 15 days (round trip), with a time constraint that must not exceed 28 days per period. The pontoon used has a capacity of 250 KL. Meanwhile, the fuel demand in the East Area in 2023 on average was 1,287 KL per month out of a total demand of 7,500 KL per month across all areas. Due to the high demand in the West area, most of the pontoons were directed to support the West region more extensively. This concern was further provoked by the unproductivity of some heavy equipment since the fuel distribution delays. The problem worsened with the projected increase in monthly demand, which was expected to reach 9,495 KL per month by 2025.

Given these challenges, a proposal has emerged to add more distribution pontoon units. However, this solution must be supported by data to convince management. In addition, the company encounters difficulties in finding vendors capable of providing pontoons in the future. Meanwhile, the addition of new units requires a significant increase in storage capacity in each area. This process was not only time-consuming but also requires a substantial financial investment.

To solve this problem, the company intended to evaluate the optimization of available resources. The goal was to improve the efficiency of fuel distribution to the West and East Areas by considering target fulfillment, optimal distribution time to reduce the need for large safety stock, and minimizing additional operational costs. The company was limited to two main distribution routes: West and East. However, if the West route is used to support the needs of the East, the company must consider the use of intermodal transportation, which could potentially increase additional costs.

Tab. 10 summarized the fuel distribution solution based on the scenario models that have been developed by referring to the current target and future highest projections. The objective was to optimize the needs of both the East and West Areas simultaneously. The comparison highlights differences before and after the implementation of the proposed model. Each scenario has a specific fuel target and considers various factors such as the number of trips, fuel consumption, time, and cost.

Table 10 Comparison between current conditions and proposed model

Scenario	Target	Route	Current condition				Proposed model				Comparison of current conditions vs proposed model	
			Trip	Fuel (KL)	Time (days)	Cost (Rp/ltr)	Trip	Fuel (KL)	Time (days)	Cost (Rp/ltr)	Fuel (KL)	Time (days)
1	1,286,609	West	-	-	-	-	3	1,500	7	207.80	–1,033	8
		East	1.87	467	15	462	-	-	-	-		
2	2,292,828	West	-	-	-	-	4	3,000	7	204,99	–2,533	8
		East	1.87	467	15	462	-	-	-	-		
3	7,499,694	West	2-3	6,407	$\bar{x} = 7.58$	209,83		7,785	7	193,43	–1,378	0.58
		East	1.87					-	-	-		
4	9,494,714	West	2-3	6,407	$\bar{x} = 7.58$	209,83		10,380	7	193,33	–3,973	0.58
		East	1.87					-	-	-		

1) Scenario 1: With a target of 1,286,609 liters, the current conditions for the East route included 1.87 trips, 467 KL of fuel, 15 days, and a cost of Rp 462 per liter. In the proposed model, for the West route, 3 trips were required, carrying 1,500 KL of fuel, with a travel duration of 7 days and a cost of Rp 207.8 per liter. The comparison showed an increase in time by 8 days and a cost adjustment to Rp 254.2 per liter, with an achievement percentage of 221% for fuel, 53% for time, and 55% for cost.

2) Scenario 2: With a target of 2,292,828 liters, the current conditions for the East route remained the same as in the first scenario. The proposed model for the West route required 4 trips, carrying 3,000 KL of fuel, with a travel duration of 7 days and a cost of Rp 204.99 per liter. The comparison showed an increase in time by 8 days and a cost adjustment to Rp 257.01 per liter, with an achievement percentage of 542% for fuel, 53% for time, and 56% for cost.

3) Scenario 3: With a target of 7,499,694 liters, the current conditions for the West route included 3-2 trips, 6,407 KL of fuel, 7.58 days, and a cost of Rp 209.83 per liter. In the proposed model, 3 trips were required, carrying 7,785 KL of fuel, with a travel duration of 7 days and a cost of Rp 193.43 per liter. The comparison showed an increase in time by 0.58 days and an adjusted cost of Rp 16.4 per liter, with an achievement percentage of 22% for fuel, 8% for time, and 8% for cost.

4) Scenario 4: With a target of 9,494,714 liters, the current conditions for the West route remained the same as in the third scenario. The proposed model required 4 trips, carrying 10,380 KL of fuel, with a travel duration of 7 days and a cost of Rp 193.33 per liter. The comparison showed an increase in time by 0.58 days and an adjusted cost of Rp 16.5 per liter, with an achievement percentage of 62% for fuel, 8% for time, and 8% for cost.

Fig. 11 listed a comparison of fuel demand fulfillment capability between the current conditions and the proposed model across four different scenarios.

1) Scenario 1: With a fuel fulfillment target of 1,286,609 liters, the current conditions showed a deficit of 820 KL. Meanwhile, the proposed model achieved an improvement of 213 KL. This indicated that the proposed model could reduce the existing deficit and improved fuel fulfillment capability.

2) Scenario 2: At a target of 2,292,828 liters, the current conditions showed a larger deficit of 1,826 KL; however, the proposed model demonstrated an increase of 707 KL in fulfillment capability. This showed that the proposed model not only reduced the deficit but also enhanced the fuel demand fulfillment.

3) Scenario 3: For a target of 7,499,694 liters, the current conditions showed a deficit of 1,093 KL, while the proposed model showed an increase of 285 KL in fulfillment capability. This indicated that the proposed model contributed to improving fuel fulfillment efficiency, even for high targets.

4) Scenario 4: At a target of 9,494,714 liters, the current conditions showed a deficit of 3,088 KL. The proposed model achieved an increase of 885 KL in fulfillment capability. This suggested that the proposed model reduced the deficit in fuel fulfillment for scenarios with very high targets.

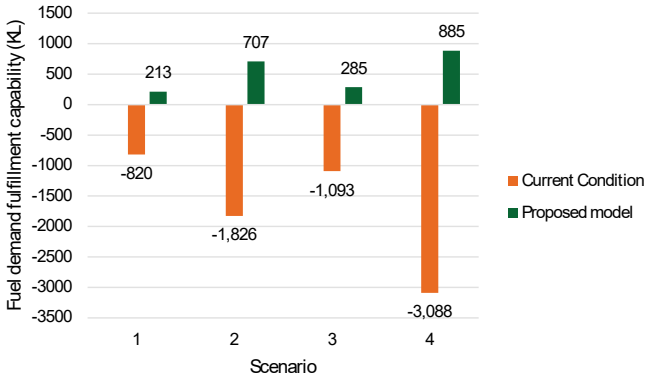


Figure 11 Comparison between current conditions and proposed model

The proposed model in this study demonstrated its effectiveness in improving fuel demand fulfillment capability compared to current conditions. These findings indicated that applying goal programming to optimize fuel distribution provided valuable benefits to the industry by reducing deficits and improving overall operational efficiency.

4.3 Analysis of Constraints in Implementing the Proposed Model

This study determined an optimal scenario for fuel oil distribution routes using goal programming mathematical models to fulfill target demands. The proposed goal programming model offers several noteworthy benefits. It enables simultaneous optimization of multiple objectives such as fuel demand fulfilment, transport time, and cost, rather than focusing on a single criterion. Thus, it reflects real-world trade-offs more accurately. Existing study of [35] pointed out that goal programming models were highly effective for transportation problems involving conflicting objectives, but their performance depends sensitively on weights and scenario definitions. Similarly, [36] demonstrated a goal programming strategy for fuel distribution routing in Indonesia and aligned with our model's focus. In addition, [37] highlighted that while multi objective frameworks provide rich trade-off insights, they require significant data, computational effort, and careful interpretation of results.

The implementation of the new model to optimize fuel distribution might face several challenges that need to be solved as summarized in Tab. 11. For example, the complexity of scheduling and route coordination requires an integrated scheduling system to optimize travel time. In addition, extra costs and additional workforce must be well-prepared to secure sustainability in operation [35–37].

Table 11 Challenges and proactive plans in implementing the proposed model

No.	Challenges	Proactive plans to solve challenges
1	Scheduling and route coordination complexity	Implementing an integrated scheduling system to maximize time efficiency and vehicle route coordination
2	Cost and workforce enlargement	Estimating additional costs and preparing an sufficient budget plan as well as a recruitment strategy to solve increased workload
3	Provide appropriate infrastructure	Developing a plan for the maintenance and development of crucial infrastructure to support optimal fuel distribution operations
4	Fuel loss risk management	Developing methods and procedures to minimize fuel loss risks during transportation, such as strict management and supervision
5	Development of training and human resource competency	Conduction training on the use of the new system and developing human resource competencies in understanding and managing change
6	System integration and data security	Providing smooth and secure system integration for managing sensitive data related to fuel distribution

Furthermore, the company must also consider sufficient infrastructure to support optimal distribution operations, including necessary maintenance and infrastructure development. The other key focus is to manage the risk of fuel loss during transportation that required the development of methods to minimize risks and strict supervision. Proper training activities to enhance human resource competencies are also crucial to provide that the team could understand and manage changes that come from the implementation of this new model. Moreover, effortless system integration and tolerable data security must be prioritized to protect information regarding the fuel distribution. By implementing these proactive plans into action, the company could guarantee that the necessary changes in fuel distribution operations will be carried out effectively and sustainably in the future. Thus, it was expected that these plans would aid in obtaining the goals of optimizing time, cost, and transport capacity.

Although the scope of this study is on fuel distribution in plantation operations, the proposed model framework has the potential to be adapted to other industries that involve

bulk transport and distribution. For example, chemical product delivery, cold chain logistics, mining supply bases, or remote infrastructure support services that face similar challenges of fulfilling demand, optimizing time, and controlling cost under constrained networks. The future studies need to reformulate the scenario sets, goal programming targets. It is expected the same approach could support organization management to allocate transport resources more strategically, simulate future scenarios, and balance operational trade-offs effectively.

5 CONCLUSIONS

The present study aimed to determine an optimal scenario for fuel oil distribution routes by utilizing goal programming mathematical models to fulfill target demands. This study investigated four different scenarios. Based on a comparison of present conditions and the proposed model, the results revealed that the suggested model greatly reduced the gap in fuel demand fulfillment across all scenarios. The proposed model indicated a greater ability to achieve fuel demand requirements, with increases of 1,033 KL (Scenario 1), 2,533 KL (Scenario 2), 1,378 KL (Scenario 3), and 3,973 KL (Scenario 4). These findings showed the goal programming model's efficiency in optimizing distribution routes to meet fuel demand. The data showed that the proposed strategy not only enhanced fuel demand fulfillment, but it also increased time and cost efficiency. In each scenario, the proposed model's travel time and fuel cost per liter were more efficient than the current conditions. In Scenario 1, travel time and cost were optimized, with a 221% increase in fuel, 53% in time, and 55% in cost. Scenarios 2, 3, and 4 also result in considerable time and cost savings. The proposed approach improved resource use by optimizing time, cost, and transport capacity for fuel distribution routes.

For the practical implications of this study, manager of logistics could use the models in prioritizing route options, allocating tanker, and pontoon resources more effectively, and simulating future demand scenarios for strategic decision-making. The model serves as a tool to support decision making in balancing fulfilment, time and cost in fuel-distribution operations. However, several limitations also need to be addressed. The model assumes constant route capacities, deterministic demand forecasts, and without any disruption events (e.g., breakdowns or weather delays) that may not fully capture the real-world process. In addition, the study case application of this study is specific to a fuel distribution for plantation that may need to be adjusted for other industries. Several recommendations were also made for future investigation. First, building upon this work, future research should focus on developing models with additional extensions, such as scheduling and vehicle route integration, to improve loading and unloading process efficiency. Second, it is proposed to take into account other objective functions, such as limiting handling loss or depreciation during transportation. Finally, using data mining or machine learning approaches to discover patterns, predict future demand, and improve decision-making in fuel distribution could result in additional cost, time, and resource savings.

6 REFERENCES

[1] British Petroleum Company. (2017). *BP energy outlook 2017*. https://www.bp.com/

[2] Bernstein, R., & Madlener, R. (2015). Short- and long-run electricity demand elasticities at the subsectoral level: A cointegration analysis for German manufacturing industries. *Energy Economics, 48*, 178–187. https://doi.org/10.1016/j.eneco.2014.12.019

[3] Julaikah, J., Hardiansyah, G., & Roslinda, E. (2022). Biaya pemanenan tanaman akasia krasikarpa (Acacia crasicarpa A. Cunn. ex Benth) di PT Kalimantan Subur Permai Kabupaten Kubu Raya Kalimantan Barat. *Jurnal Lingkungan Hutan Tropis, 1*(2). (in Indonesian)

[4] Xu, W., Luis, M., & Yuce, B. (2023). A hybrid method for the closed-loop supply chain to minimize total logistics costs. *International Journal of Technology, 14*(7), 1449. https://doi.org/10.14716/ijtech.v14i7.6710

[5] Lukman, I., & Hanafi, R. H. (2019). Optimasi biaya distribusi pada HFVRP menggunakan algoritma particle swarm optimization. *Jurnal Optimasi Sistem Industri, 18*(2), 164–175. (in Indonesian) https://doi.org/10.25077/josi.v18.n2.p164-175.2019

[6] Mashuri, C., Mujianto, A. H., Sucipto, H., & Arsam, R. Y. (2020). Sistem optimasi penjadwalan mesin produksi menggunakan metode GUPTA berbasis Android. *Jurnal Sistem Informasi Bisnis, 10*(1), 20–27. (in Indonesian) https://doi.org/10.21456/vol10iss1pp20-27

[7] Imansuri, F., Hadiguna, R. A., & Afrinaldi, F. (2019). Model optimasi perancangan jaringan rantai pasok biomassa dari tandan kosong kelapa sawit di Sumatera Barat. *Jurnal Optimasi Sistem Industri, 18*(1), 1–13. (in Indonesian) https://doi.org/10.25077/josi.v18.n1.p1-13.2019

[8] Viarani, S. O. V. M., Henmaidi, & Bronto Adi, A. H. (2024). Model jaringan distribusi produk dengan pendekatan fuzzy multi objective programming. *Jurnal Optimasi Sistem Industri, 17*(1), 1–15. (in Indonesian) https://doi.org/10.25077/josi.v17.n1.p1-15.2018

[9] Meilani, D., & Iswara, A. (2018). Aplikasi penentuan rute distribusi LPG 3 kg. *Jurnal Optimasi Sistem Industri, 17*(2), 208–219. (in Indonesian) https://doi.org/10.25077/josi.v17.n2.p208-219.2018

[10] Hidayat, S., Nurhasanah, N., & Zulkifli, A. S. (2016). Perencanaan dan penjadwalan distribusi pakaian jadi dengan metode Distribution Resource Planning. *Jurnal Optimasi Sistem Industri, 12*(2), 343–351. (in Indonesian) https://doi.org/10.25077/josi.v12.n2.p343-351.2013

[11] Maridjo, I. Y. A. R. (2019). Pengaruh pemakaian bahan bakar premium, pertalite dan pertamax terhadap kinerja motor 4 tak. *Jurnal Teknik Energi, 9*(1), 73–78. (in Indonesian) https://doi.org/10.35313/energi.v9i1.1648

[12] Petropoulos, F., Akkermans, H., Aksin, O. Z., Ali, I., Babai, M. Z., Barbosa-Povoa, A., … Zheng, T. (2025). Operations and supply chain management: Principles and practice. *International Journal of Production Research*, 1–184. https://doi.org/10.1080/00207543.2025.2555531

[13] Prathapage, H., Ivanov, D., & Ivanova, M. (2025). Methodological approaches and managerial insights in supply chain stress testing. *International Journal of Production Research*, 1–40. https://doi.org/10.1080/00207543.2025.2561779

[14] Sawik, T., & Sawik, B. (2024). Risk-averse decision-making to maintain supply chain viability under propagated disruptions. *International Journal of Production Research, 62*(8), 2853–2867. https://doi.org/10.1080/00207543.2023.2236726

[15] Sawik, B. (2020). Selected multiple criteria supply chain optimization problems (pp. 31–58). https://doi.org/10.1108/S0276-897620200000020003
[16] Fatma, E., & Kartika, W. (2017). Penjadwalan dan penentuan rute distribusi komoditas ke wilayah Timur Indonesia. *Jurnal Optimasi Sistem Industri, 16*(1), 40–49. (in Indonesian) https://doi.org/10.25077/josi.v16.n1.p40-49.2017
[17] Ahmed, S., & Dey, K. (2020). Resilience modeling concepts in transportation systems: A comprehensive review based on mode and modeling techniques. *Journal of Infrastructure Preservation and Resilience, 1*(1), 8. https://doi.org/10.1186/s43065-020-00008-9
[18] Bekiaris, E. D. (2019). Optimisation and personalization technologies and algorithms for future transportation systems. *European Transport Research Review, 11*(1), 39. https://doi.org/10.1186/s12544-019-0376-2
[19] Otręba, P., & Sawik, B. (2025). Determinants of airfares in the short-haul European market: An empirical study. *Expert Systems with Applications*, 130266. https://doi.org/10.1016/j.eswa.2025.130266
[20] Hashemizadeh, A., & Ju, Y. (2021). Optimizing renewable energy portfolios with a human development approach by fuzzy interval goal programming. *Sustainable Cities and Society, 75*, 103396. https://doi.org/10.1016/j.scs.2021.103396
[21] Kouaissah, N., & Hocine, A. (2020). Optimizing sustainable and renewable energy portfolios using a fuzzy interval goal programming approach. *Computers & Industrial Engineering, 144*, 106448. https://doi.org/10.1016/j.cie.2020.106448
[22] Canelas, E., Pinto-Varela, T., & Sawik, B. (2020). Electricity portfolio optimization for large consumers: Iberian electricity market case study. *Energies, 13*(9), 2249. https://doi.org/10.3390/en13092249
[23] Sirodj, D. A. N., & Gunawan, I. (2020). Recycling solid waste of coconut oil industry: A response surface-goal programming approach. *Jurnal Optimasi Sistem Industri, 19*(2), 111–121. https://doi.org/10.25077/josi.v19.n2.p111-121.2020
[24] Gupta, S., Ali, I., & Ahmed, A. (2018). Efficient fuzzy goal programming model for multi-objective production distribution problem. *International Journal of Applied and Computational Mathematics, 4*(2), 76. https://doi.org/10.1007/s40819-018-0511-0
[25] Sofiyanto, S., Yulianto, T., & Faisol, F. (2021). Penerapan metode goal programming dalam mengoptimalkan pendistribusian gas LPG di SPPBE Tlanakan. *Zeta - Math Journal, 6*(2), 48–53. https://doi.org/10.31102/zeta.2021.6.2.48-53
[26] Jia, R., Liu, Y., & Bai, X. (2020). Sustainable supplier selection and order allocation: Distributionally robust goal programming model and tractable approximation. *Computers & Industrial Engineering, 140*, 106267. https://doi.org/10.1016/j.cie.2020.106267
[27] Cheng, J., Feng, X., & Bai, X. (2021). Modeling equitable and effective distribution problem in humanitarian relief logistics by robust goal programming. *Computers & Industrial Engineering, 155*, 107183. https://doi.org/10.1016/j.cie.2021.107183
[28] Sawik, B., Faulin, J., & Pérez-Bernabeu, E. (2017). A multicriteria analysis for the green VRP: A case discussion for the distribution problem of a Spanish retailer. *Transportation Research Procedia, 22*, 305–313. https://doi.org/10.1016/j.trpro.2017.03.037
[29] Gliwan, S. E., & Crowe, K. (2022). A goal programming model for dispatching trucks in an underground gold mine. *European Scientific Journal, 18*(36), 1. https://doi.org/10.19044/esj.2022.v18n36p1
[30] Sawik, B. (2024). Optimizing last-mile delivery: A multi-criteria approach with automated smart lockers, capillary distribution and crowdshipping. *Logistics, 8*(2), 52. https://doi.org/10.3390/logistics8020052
[31] Sadatian, S. A., Sadatian, S. D., Asiaei, M., & Davoodi, S. A. R. (2022). Providing a model for evaluating and selecting suppliers of three-phase self-sustaining cables using the interactive approach of analytical hierarchical process and goal programming. *OPSEARCH, 59*(4), 1649–1666. https://doi.org/10.1007/s12597-021-00570-0
[32] Tanwar, J., Scholar, R., Kumar, A., & Rao, M. (2020). Mathematical modeling of asset liability management in banks using goal programming and AHP. *Indian Journal of Finance and Banking, 4*(4). https://www.cribfb.com/journal/index.php/ijfb
[33] Faisal, D. M. N., Bagus, H., & Sunarya, S. (2021). Perhitungan metode goal programming untuk optimasi perencanaan produk keripik singkong pada PT Cassava Chips. *Bulletin of Applied Industrial Engineering Theory, 12*(1). (in Indonesian)
[34] Huang, F. (2022). Low-carbon supplier selection using fuzzy AHP and goal programming approach. *Discrete Dynamics in Nature and Society, 2022*(1). https://doi.org/10.1155/2022/3115490
[35] Gütmen, S., Roy, S. K., & Weber, G.-W. (2024). An overview of weighted goal programming: A multi-objective transportation problem with some fresh viewpoints. *Central European Journal of Operations Research, 32*(3), 557–568. https://doi.org/10.1007/s10100-023-00861-5
[36] Melani, T., Abusini, S., & Marjono. (2025). Fuel distribution optimization in Malang: A hybrid routing approach. *CAUCHY: Jurnal Matematika Murni dan Aplikasi, 10*(1).
[37] Abdelati, M. H., Abd-El-Tawwab, A. M., Ellimony, E. E. M., & Rabie, M. (2023). Solving a multi-objective solid transportation problem: A comparative study of alternative methods for decision-making. *Journal of Engineering and Applied Science, 70*(1), 82. https://doi.org/10.1186/s44147-023-00247-z

Authors' contacts:

David Billy Martin Salangka
Industrial Engineering Department,
BINUS Graduate Program – Master of Industrial Engineering,
Bina Nusantara University,
Jakarta 11480, Indonesia
E-mail: david.salangka@binus.ac.id

Hwi-Chie Ho, Associate Professor
(Corresponding author)
Industrial Engineering Department,
Faculty of Engineering, Bina Nusantara University,
Jakarta 11480, Indonesia
E-mail: hhchie@binus.edu

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20260320081742

Original scientific paper
Received: 2026-03-20, Accepted: 2026-06-02

A Study on the Stabilization Methods of Structural Systems and the Form-active Characteristics of Buildings (Focusing on Early Modern Buildings)

Sung Hyun Jo

Abstract: Integrated designs involving structure and design are closely intertwined with the initial stages of architectural design. Although the integration of structural systems and architectural aesthetics is a core element of contemporary design, existing studies have often treated structural efficiency and spatial form as independent factors. In most cases, the focus has been on areas where structural systems play a primary role, such as the relationship between structure and building materials, the development of structural analysis software, and the design of high-rise buildings. Consequently, there is a lack of empirical analysis on how structural systems directly influence the composition of architectural space. To address this, the MIDAS program was used to visualize areas of concentrated bending moments and load flow, and analyze the interrelationship between structural stability and spatial form. The analysis results indicate that structural systems classified as Skeleton and Truss types perform distinct spatial roles. Structures that emphasize architectural forms (e.g., lattice trusses, geodesic domes, and tapered skeleton structures) are primarily used for roof construction. Conversely, structures defining the architectural space (rectangular skeleton structures and Vierendeel trusses) are primarily employed for horizontal floor slab construction through stress redistribution.

Keywords: Skeleton structure; Structural stability factors; Structural system; Truss structures

1 INTRODUCTION

1.1 Background and Purpose of the Study

A building forms a space through the physical combination of floor slabs and walls, and its form consists of a roof that responds to stress. Buildings pursue different stabilization methods based on the structural system that responds to stress, and each has its own unique formal characteristics. Therefore, a building is characterized by its floor plan, which defines the space and form of the roof with its characteristic envelope.

Vitruvius, in his book De Architectura around 25 BC, defined the essence of architecture as commoditas, firmitas, and venustas. This is explained by Italian architect and engineer Pier Luigi Nervi in his book New Structures (1963), in which he describes function, structure, and form as essential elements of a building, stating that architecture is completed when these three elements are harmoniously combined. The structure, as the most fundamental element of a building, is the structural system. The structural system that influences the architectural form usually has a naturally curved shape that follows the flow of loads.

Historians are interested in buildings as a way of evaluating an era because architecture is the result of the complex interaction of history, culture, and scientific technology of that time. Knowledge of structural behavior and mechanical properties is necessary to understand buildings and set design criteria. The structural system has an absolute influence on determining the physical size of the floor plan and constituting the external form. Modern architecture can be considered a case in which the new materials of that era—plate glass, reinforced concrete, and steel—and the technology of the structural system were actively reflected in the form of buildings.

Examples of prior research are as follows.

Sev and Tugrul [1] classified the structural systems used in high-rise building designs and specifically investigated structural analysis and digital design methods for high-rise buildings with nonstandard forms. They also emphasized the importance of integrating architectural and structural designs.

Uihlein [2] conducted interviews with practitioners to examine structural integration. The interviews were organized into areas such as the collaborative relationship with structural engineers during the design phase, the role of the structure, the criteria for evaluating architectural design, and the characteristics of structural integration. In other words, expert interviews were used as the research methodology.

Li et al. [3] regarded structure as the primary agent of morphogenesis by conducting experiments on tension, compression, curved surfaces, and load deformation. They observed that structural rationality profoundly influenced spatial composition, whereas the performance of construction materials also served as a key factor in generating form. Through these four experiments, they advocated the integration of structure and design.

Jin-ho et al. [4] proposed an integrated structural design system and a model for buildings. Based on structural analysis, they developed a design model by modeling the main structure, substructure, and connection structure. In other words, this research focused on developing a model in which the structural analysis results took precedence over the integration of the architectural design and structure.

Janakieska [5] defined structural systems as factors influencing building design, setting parameters such as room size, circulation paths, furniture, and functionality, and conducted a case study limited to hotel architecture.

Ahmed El-Khadrawy [6] examined the influence and correlation between geometric patterns found in nature (triangles, hexagons, and lattice structures) and the structural systems and spatial configurations of modern architecture. He analyzed buildings with truss structures and focused his research on the geometric composition of exterior forms rather than on interior spaces.

Kookalani et al. [7] focused on developing programs to enhance productivity and achieve optimized designs through automation of architectural and structural designs. The

authors presented methods for automating the structural design using generative design (GD), deep generative models (DGM), and topology optimization (TO). They also proposed a hybrid approach that integrates these design methods according to the project characteristics. However, the program primarily focuses on structural systems, with insufficient coverage of integrated designs and spatial considerations.

Firdevs Kozan [8] used structural analysis software (SAP2000) to model various exterior structural forms, including tubular frames, tubular diagrids, and lattice diagrids, and conducted a quantitative study of seismic loads and load behavior. The study focused on high-rise buildings, where the use of structural systems is essential. This research also primarily focuses on external structural forms, and it can be considered to lack sufficient discussion of interior floor plans.

Acican and Luyten [9] conducted a project involving architecture students, in which each student's project progressed through structural systems and structural analysis within the architectural design curriculum. The structural simulation results were applied to architectural project design. Students received feedback from experts to develop their designs and approached the integration process of architectural structure and design within an educational context.

Mario Rinke and Robbe Pacquée [10] proposed the concept of "structural porosity", which refers to the degree to which a structural system is open to space, as a factor determining the adaptability and flexibility of buildings. Their research primarily focused on typical reinforced concrete column-beam frames and rib-slab grids, demonstrating that structures with more gaps (porosity) between the framework—such as typical column-beam structures or rib slabs—are better suited for future building renovations or flexible spatial use. However, the study does not extend to truss structures or long-span buildings, and the selection of structural systems applicable to new construction or the early design stages is excluded from the research. Instead, it can be viewed as a case study examining the role of structural systems primarily in the context of renovating existing buildings or changing their use.

Previous research has focused on complex buildings such as high-rise structures and hotels, where the structural system often exerts a decisive influence on the architectural form. Hotel architecture typically requires specialized structural systems, such as transfer girders, because it involves placing large-span banquet halls in the lower zones and repetitive small-scale unit rooms in the upper zones. The verification methods in existing research include expert interviews, experimental studies on structures and materials, and the development of structural design models.

Previous studies have focused on cases where structural systems inevitably play a crucial role in architectural design, primarily examining the relationship between exterior form and structural systems.

However, there are very few research cases that identify structural stability factors from the early design stages, select appropriate structural systems, and integrate these with the building’s exterior form and interior spaces.

The testing method used in this study involved identifying stress concentration points using structural analysis software (MIDAS) and examining structural stabilization methods as a solution. Each stabilization method is represented as a structural system.

This study categorizes structural systems into architectural forms and spatial organizations. It aims to identify the stabilizing elements within each structural system and clarify their respective impacts on form and spatial organization.

The significance of this study lies in the fact that the structural requirements for the building envelope and those for the floor plan necessitate different structural systems, and that by considering how to address these structural requirements from the initial design stage, a foundation can be established for integrating the structural system with the architectural design.

1.2 Scope and Methods of the Study

A building maintains its form and position because the structural system that responds to internal stress creates the form, and the load from above and the bearing capacity of the ground are balanced, thereby maintaining its position.

In early modern buildings, technological elements were directly reflected in the buildings due to the development of new materials and connection methods. Kenneth Frampton defined modern architecture in his book Modern Architecture as "a break with historicist styles and a search for a new formal logic" within the technological and social changes following the Industrial Revolution at the end of the 19th century. In History of Modern Architecture, Leonardo Benevolo defines it as the totality of newly emerging architectural thought and practice within the context of changes in technology, social structure, and urban conditions following the Industrial Revolution. He stated that ‘modern architecture’ signifies not only new forms, but also a fundamental shift in thinking, viewing it as the representative concept of architectural movements formed from the late 19th century to the mid-20th century.

In this study, modern architecture is defined as emerging after the Industrial Revolution (late 19th century) and focuses on buildings that actively influenced spatial and formal qualities through the advent of new materials and construction methods, such as iron, glass, and reinforced concrete (mid-20th century).

Buildings that use the skeleton structure method are small-scale buildings, many of which are residential buildings; whereas those that use the truss structure method consist of long-span buildings.

This research method classifies buildings into skeleton and truss structure types, and represents them as free-body diagrams with support points and reaction forces. Free-body diagrams simplify buildings by separating them from their envelopes, serving as a straightforward model for structural analysis, and effectively expressing the effects of reaction forces and loads. Using Midas Gen, this free-body diagram approach examined the stability methods, corresponding forms, and spatial effects at nodal points where the bending moments are concentrated from a system integration perspective.

This program was initially developed within an internal venture of the POSCO Group, and is currently commercialized and distributed by MIDAS Information Technology as a commercial finite element structural analysis program. Both dead and live loads were applied, and the self-weight of the frame was automatically calculated using the analysis program (1.2D + 1.6L).

For the boundary conditions, the reinforced concrete (RC) frame structure employed rigid joints (fixed) to accommodate bending moments at the points of occurrence by utilizing the member thickness or shape. The steel frame structure places the steel members (bars) at the points of axial force occurrence and connects those using pinned joints (fins).

This research is not intended for structural design purposes, but rather aims at system integration between the structural system, architectural form, and architectural design; therefore, numerical analysis was excluded. Such research is expected to yield positive effects by approaching the design process, which is traditionally conducted separately within each architectural discipline, from an integrated design perspective that encompasses both architectural structure and architectural design.

An efficient structure is one in which forces move in the form of structural elements. A structurally efficient form minimizes the generation of bending moments or responds to the generated bending moments through the thickness or shape of the members.

The architectural examples of skeletal structures in this study focused on residential buildings by Le Corbusier and Mies van der Rohe. Primarily, small-scale buildings were selected in which load-bearing walls functioned as rigid diaphragms to ensure stability.

Skeleton structures primarily transfer loads in the vertical direction through the combination of columns and beams, but shear walls are used to resist lateral loads generated on the exterior of the building. The residential building case study in this research was selected based on a structure in which shear walls are installed parallel to the lateral loads, centered around the core. Additionally, examples featuring tapered structural frames were selected, in which the thickness or shape of the members served as stabilizing elements.

While tapered members are primarily used in PEB systems for steel structures, we selected buildings featuring them in skeleton structures—specifically, monolithic structures where the spectator seating and roof are structurally integrated—focusing on examples where the depth of the members is intentionally thickened at points where structural forces are concentrated, such as in cultural and assembly facilities or the interior columns of factories.

For the truss structure building cases, structures used for purposes involving relatively large numbers of people, such as research and exhibition facilities, were selected. The focus was on buildings in which axial forces were the primary stresses, enabling the creation of relatively long-span spaces. Accordingly, we selected case studies that utilized Warren trusses based on triangular grids and Vierendeel trusses based on orthogonal frames to expand spatial freedom and create column-free spaces, as well as a case study that implemented a Space Lattice Truss based on a three-dimensional spatial grid to create a circular plan, and examined the relationship between structure and form.

This study examines how structural engineering expertise influences floor plans and external forms by selecting cases in which structural elements are actively expressed using new materials and construction methods. It also views early modern architecture as the beginning of system-integrated design, where facilities and structural systems are actively expressed in architectural form, much like contemporary high-tech architecture.

2 THE RELATIONSHIP BETWEEN THE DEVELOPMENT OF ARCHITECTURE AND STRUCTURAL SYSTEMS

2.1 The Development of Building Materials and Changes in Structural Systems

Architectural styles change with the sociocultural context, scientific technology of the era, emergence of new materials, and development of construction methods.

Ancient Western architecture used stone as its main material and iron as the connection method. Iron used in the past lacked sufficient tensile strength and was not used as a building material.

The structural system of Greek architecture consists of columns and beams made of stone. They moved beyond past masonry methods to construct magnificent temple buildings. The stone was cut to size, and the individual members were grooved to insert fasteners for use as columns and beams. For buildings intended for large gatherings, such as assembly facilities, a mixed structure was used with wooden trusses for the roof and stones for the walls. Thus, the structural system takes form through appropriate selection and combination according to the scale and use of the space.

Roman architecture is characterized by arches, vaults, and dome structures. It has changed from architecture for the gods in Greece to architecture for people, requiring larger interior spaces than Greek buildings. The material was stone, as in the Greek era; however, the main structural system responded to compressive forces. A representative example is an aqueduct, which is a water-supply system in which continuous arches are stacked in layers.

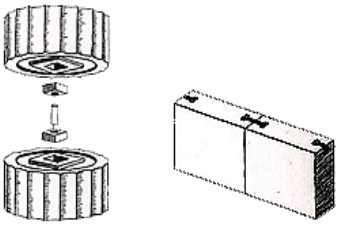


Figure 1 Method for joining columns and beams

The shape of the arch structure causes vertical forces to flow along the curve as compressive forces. Stone is an appropriate material to respond to these compressive forces, and an example of a dome structure is the Pantheon. The interior has a circular plan with a diameter of 43.6M and uses a dome structure for the roof.

Gothic architecture is represented by cathedral buildings, where pointed arches, rib vaults, and flying buttresses were used as structural elements. The pointed arch

can guide compressive forces more vertically, reducing horizontal forces. The rib vault is often seen as the starting point of the Gothic style, and it can reduce the load of the ceiling by concentrating the roof load onto the ribs, which are structural members. A flying buttress is an arch-shaped structural member that prevents the buckling of walls. It distributes the lateral force (thrust) generated on the nave side of the cathedral to the outside.

Gothic architecture can be seen as a case in which a structural solution is rationally applied to symbolically express the Christian faith and the aspiration to rise high towards the heavens. In pre-modernization Western architecture, stone was used as the material, stone units were cut into individual members, and the method of connecting and constructing the members had a significant influence on the structural system and architectural form. With the increase in time and demand, large-scale and unified interior spaces have become necessary, and various construction methods have been applied to create larger and wider spaces. In other words, the main architectural challenge before modernization was creating large interior spaces using stone as the material.

2.2 The Development Process of Modern Architecture and Structural Systems

Ancient Western buildings were mostly temples and cathedrals. Through social transitions, such as the Industrial and civil revolutions, architecture has changed from being for gods to being for people.

The reason for the interest in modern architecture is that it actively applies new construction methods and structural systems to buildings using the new materials of that era: plate glass, iron, and concrete.

19th-century architecture was a transitional period to modern architecture, where the main material, stone, was replaced by new materials such as reinforced concrete and plate glass. The use of steel structures with iron as the main structural frame began in the late 19th century. Early 20th-century buildings used the skeleton structure method with reinforced concrete as the material [11].

A combination of Rahman structures built with columns, beams, and partial load-bearing walls was used. The formal characteristic of such buildings is the repetitive use of forms owing to the modularization of the structural system, and they are often constructed with straight lines rather than curves, considering the workability of concrete.

Late 20th-century buildings used truss structures. Using a truss structure, various stresses occurring in the building are converted into axial forces, thereby reducing the self-weight, and the corresponding members can be thin and long linear elements. This has made it possible to create relatively wide and uniform long-span buildings.

In buildings, structural members typically experience axial forces, bending moments, or a combination of these stresses. Structural members are more efficient when subjected to axial forces than to bending moments. This is because the stress caused by the axial force is distributed almost uniformly across the entire cross-section of the member. This stress distribution allows the entire cross-section to be utilized up to the allowable limit of the material. In contrast, the bending stress distribution exhibited maximum values at the top and bottom flanges of the section and minimum values at the neutral axis. Furthermore, the bending stress requires both the cross-sectional area and the distance between these areas. Therefore, the most critical factor in the design of a bending member is its thickness in the direction of rotation [12].

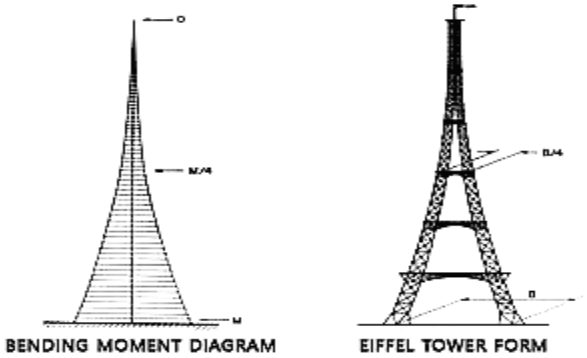


Figure 2 Eiffel Tower response to bending stress

The fact that the bending stress is related to the thickness of the structural member indicates that the bending stress is the primary factor determining the cross-section of the structural member. This is because thickening areas where large moments occur and thinning areas where moments are small naturally form the cross-section of the structural member. In practice, drawing a bending moment diagram based on the loads applied to the structure and shaping the structure yields the form closest to the most rational structural member. For example, drawing a bending moment diagram for wind loads applied to the Eiffel Tower and considering wind loads in both directions revealed that this directly corresponded to the shape of the tower [13].

This study primarily focused on the axial force and bending moment among the various stresses acting on structures. This is because these stresses exert the greatest influence on the cross-sectional shape, connection method, and overall structural form. Therefore, in this study, the stress concentration refers to the support points subjected to high bending moments.

3 STRUCTURAL ELEMENTS AND FORMS IN MODERN ARCHITECTURE

3.1 Skeleton Structure Method

Modern buildings can be divided into skeleton structures, truss structures, and form-active structures, which respond according to the type of structural system.

Each building must have a stable structure. Structural stability refers to the ability of structural members to resist loads while maintaining the form of the structure. Furthermore, a stable structure returns to its original state, even when subjected to minor disturbances, whereas an unstable structure transforms into an entirely different state. Additionally, a building must achieve equilibrium under

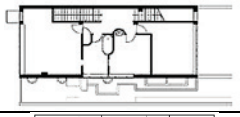
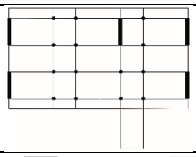
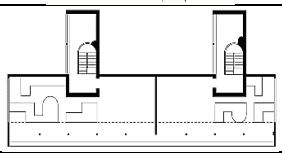
applied loads. Equilibrium implies that all applied loads are precisely balanced by the reaction forces generated at the foundation through the way the loads are transmitted via the internal form of the structure and foundation.

To achieve a stable state, the skeleton structure requires the installation of rigid connections, bracing, or rigid diaphragms. Although rigid connections are highly useful from a spatial planning perspective, they increase the degree of structural indeterminacy [12, p. 23].

The skeleton structure method can be divided into a rectangular-shaped structural method and a tapered-shaped structural method. The rectangular-shaped skeleton structure method consists of a combination of a Rahmen structure made of rigid joints of columns and beams, and load-bearing walls. Its formal characteristics are that vertical and horizontal members are joined at right angles and are composed of linear members, such as columns and beams, and planar members, such as floors, walls, and roofs. The tapered-shaped skeleton structure method responds to the stresses acting on the building with the shape of the members, pursuing appropriate stability, especially at the connection parts where the bending stress is concentrated.

Le Corbusier's dom-ino structural system is a rectangular skeleton structure method that was applied to most residential buildings in the 1920s. A flat slab was used in the lost tile method. A flat slab constructed with lost tiles guides the flow of the load acting on the slab in one direction. The slab sizes of the residential buildings where the dom-ino structure was applied were mainly 5 × 2.5 m and 5 × 5 m [14]. The dom-ino structure allows for the free arrangement of walls on the floor plan because it applies the lost-tile method, which does not require the installation of a ceiling frame. In addition, because the columns and slabs bear the load instead of load-bearing walls, they exhibit a characteristic facade with ribbon windows.

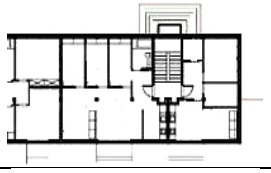
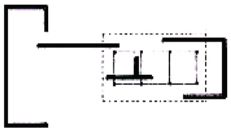
Table 1 Bearing wall and skeleton structure of the domino house

Dom-ino House	Bearing Wall & Frame Structure	Structure System
Guitte house (1926)		skeleton structure + Bearing Wall
Garches house (1927)		skeleton structure + Bearing Wall
Stuttgart-Weissenhof (1927)		skeleton structure + Bearing Wall

The structure of the Mies buildings is a rectangular Rahman structure composed of load-bearing walls, columns, and beams. He also actively used steel materials and employed disjoint in which the connection process of the members was exposed to the external form. It emphasizes the materials that make up the building and the method of connecting members. The connection of the steel members uses a pin joint, which reduces the bending stress at the ends of the members, and the finish is expressed in an external form. Therefore, the emphasis on the connection method shows that Mies's architecture constructed a fluid space using new materials and was distinct from the architectural space before the modern era.

Mies's buildings used members with improved cross-sections. An improved cross-section means using members that are faithful in responding to stress. The Barcelona Pavilion and Tugendhat House used cross-shaped columns and I-beams. The columns used reduced self-weight while performing an appropriate load distribution by welding steel plates where needed and leaving the rest empty.

Table 2 Stabilizing elements of a Mies building

Mies's Residential Architecture	Bearing Wall & Frame Structure	Structure System
Weissenhof Siedlung (1927)		Frame Structure ＋Bearing Wall
Barcelona Pavilion (1929)		Frame Structure ＋Bearing Wall
Tugendhat House (1930)		Frame Structure ＋Bearing Wall

The tapered-shaped skeleton structure method is mainly used for columns and roof shapes, such as those of stadiums, whereas the Stadium in Florence, designed by Pier Luigi Nervi, has a thicker shape at the beginning of the cantilevered canopy than at the end. The length of the long side was 220 m, and it had a rigid joint with a spectator seating slab. The spectator seating slab is inclined according to the viewing direction and acts as a brace.

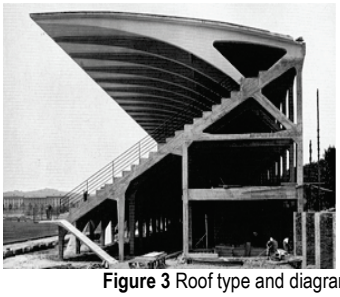
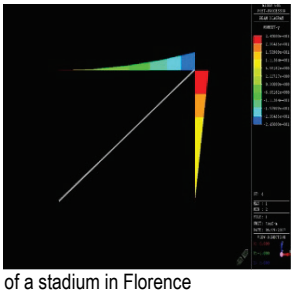



Figure 3 Roof type and diagram of a stadium in Florence

The upper part of the vertical member is the red zone, which corresponds to the region of maximum positive bending moment. The beam-column joint is represented by the blue zone, which corresponds to the region of maximum negative bending moment. In other words, the maximum bending moment of the structure occurs at the beam-column

joint. In the case of skeleton structures, the thickness of the members is increased to achieve structural stability, and in this study, such a configuration is referred to as a tapered structural system.

An example of reinforcement in a column capital can be found in the columns of ancient Greek temple architecture, where an abacus was used as the capital for reinforcement. The internal columns of the Johnson & Wax Building have a wide concentric circular shape in the capital. This is because the bending stress was formed in a concentric circular pattern when the stress flowed from the slab to the column. When columns are joined to a slab, they become load-flow paths, and isostatic lines are formed depending on the column positions. The shapes of these isostatic lines are expressed as waffle slabs or rib bolts.

To elucidate the process of ground and superstructure load transfer in a flat slab system, an analysis of the maximum bending moment distribution revealed that, as loads are applied, severe stress concentration occurs around the column capitals. This is visualized in the diagram as a circular contour pattern of red and orange lines centered on the columns. This area forms the structurally most vulnerable critical section within the slab system. It is at risk of punching shear failure, where the column pierces through the slab. Therefore, enlarging the column capital in a concentric circular pattern ensures structural stability against punching shear forces and directly visualizes the stress distribution.

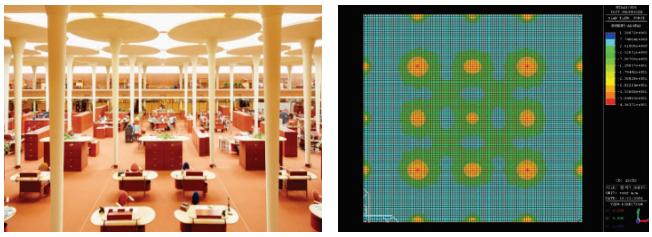


Figure 4 Capital shapes and column diagrams

The tapered shape of the pilotis columns was well expressed in the 1958 Berlin housing complex. The tapered legs that leaned into the building and columns that leaned out of the building were arranged repeatedly. A bending moment that rotates outwards occurs. Therefore, an axial force towards the inside occurs in the cantilever beam. This inward axial force is countered by the pilotis columns that are inclined outwards. The outward bending moment generates an outward axial force in the cantilever beam. To respond to this axial force, pilotis columns inclined inward were used, and the intersecting column shapes of the same structure were related to wind loads.

The cantilever structure was analyzed in two scenarios: one where the vertical load is applied at the outer end of the cantilever and another where it is applied at the inner end. Since both cases inevitably generate horizontal thrust, the direction in which the column tilts directly affects the configuration of the upper beam and the column.

When a vertical load occurs inside the cantilever, a rotational force toward the interior is generated at the joint between the beam and the column. Consequently, a horizontal thrust toward the exterior is generated in the horizontal direction of the beam. Therefore, a column configuration that leans inward is structurally advantageous. When a vertical load is applied to the outer end of the cantilever, a thrust is generated toward the interior of the beam, making it advantageous for the column to be inclined outward. In this way, the forces exerted by the inward-leaning and outward-leaning columns—pushing and pulling the beam outward—cancel each other out across the entire planar structure, allowing the building to withstand lateral drift, such as tilting to one side or spreading outward.

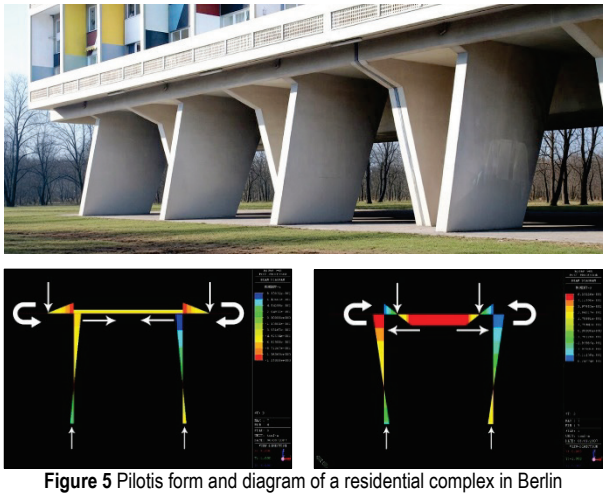


Figure 5 Pilotis form and diagram of a residential complex in Berlin

Table 3 Taper-shaped skeleton structure

Building Name	Structural System
· Hat Factory of Herrmann Company, Erich Mendelsohn, Paris, France, 1921-23	Skeleton Structure
· Stadium in Florence, Pier Luigi Nervi, Florence, Italy, 1930-32	Skeleton Structure
· Johnson&Wax Building, Frank Lloyd Wright, Wisconsin, USA, 1939	Skeleton Structure
· Stadium in Caracas, Villanueva, Venezuela, 1949	Skeleton Structure
· The Unesco Building, Pier Luigi Nervi, Paris, France, 1953-58	Skeleton Structure
· Unite d'Habitation, Le Corbusier, Berlin, Germany, 1958	Skeleton Structure
· Planalto Palace, Oscar Niemeyer, Brasilia, Brazil, 1960	Skeleton Structure

3.2 Trussed Structure Method

A truss is a simple strut assembly with pinned joints and triangular configurations that is subjected to tensile or compressive stresses [15]. The Vierendeel truss can be considered as an example of improving efficiency by altering the internal shape of members to remove portions of the cross-section in areas with low stress. In other words, by considering the stress distribution and eliminating low-stress areas from the cross-section or axial shape of the members, the efficiency of cross-section utilization is enhanced [12, p. 49].

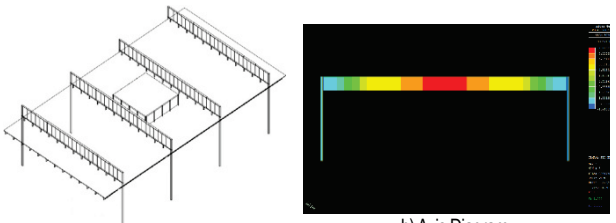
The advantage of truss structures over other form-active structures is that they generate only axial forces even without special shapes [12, p. 51]. Structures that generate only axial forces allow for reduced member thicknesses; when

members are arranged along the load path, no bending moments occur, making them structurally efficient.

Representative buildings using the truss compression structure method are Mies' Crown Hall and Buckminster Fuller's US Pavilion at EXPO 67. Crown Hall realized a column-free space inside. To achieve this, it used a construction method in which the roof was suspended and supported by lateral plate girder installed on the roof as the main structure. Four plate girder are installed and supported by two H-beam columns. In the detailed connection method, the roof purlins were joined to the bottom of the trusses using a shear connection method. The span of the columns was 18.3 m, and the load transmitted from the mullions at 3 m intervals was effectively distributed through the roof trusses. Using a glass wall, the interior appears transparently penetrated, and a column-free space can be created inside, realizing the concept of a universal space.

The roof structure used in Crown Hall is a portal moment frame system that combines plate girders with perimeter columns. It is structurally similar to a Vierendeel truss in that the column-head joints are designed as rigid joints to resist bending and shear, thereby creating a column-free interior space.

The stability of the plate girders is ensured by connecting them to purlins at their lower ends, thereby safely distributing the roof load. The 11 purlins used in the roof distribute the load to the plate girders connected above and transfer it to the H-shaped columns installed at both ends. The stiffeners of the H-shaped columns are aligned with the building's longitudinal direction to prevent overturning.



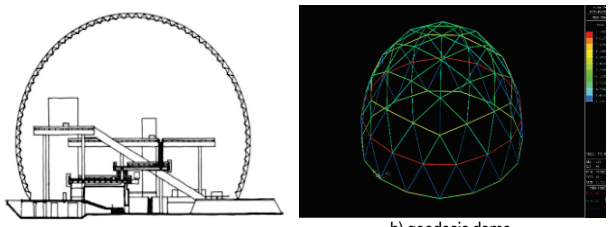
a) Roof truss b) Axis Diagram

Figure 6 Roof truss and stress diagram of Crown Hall

Buckminster Fuller's US Pavilion at EXPO 67 is a geodesic dome in which triangular trusses are repeatedly joined to convert all stresses occurring in the dome into axial forces. It is a dome structure with two layers consisting of an outer layer using triangular trusses, an inner layer with hexagonal trusses, and a web structure that connects the inner and outer layers. The final dome shape had a diameter of 76m and a height of 60.8m, and the structure used tubular steel connected by nodes [16]. The longitudinal compressive and transverse tensile forces of the dome were countered using triangular trusses.

The results of the diagram analysis show that, since the nodes of the geodesic dome structure are pin-connected, no bending moments occur. Therefore, all members can be composed of a combination of straight and diagonal members and are subjected to compressive and tensile forces.

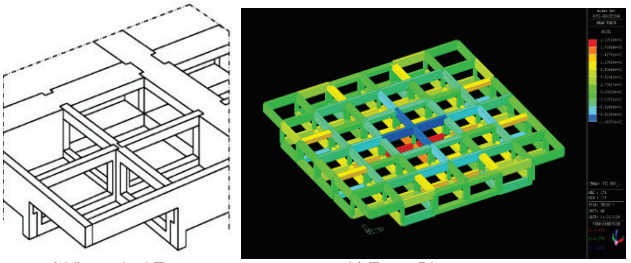
The hoop members at the bottom of the dome structure, shown in red, are subjected to maximum tensile force. Consequently, the diagonal members—which serve as vertical supports connecting the foundation to the lower hoop members—experience maximum compressive force and are shown in the darkest blue on the diagram. The remaining ribs and diagonal members are arranged to form the dome's shape and are shown in yellow or green, indicating sections where tensile and compressive forces coexist.



a) Dome structure cross-section b) geodesic dome

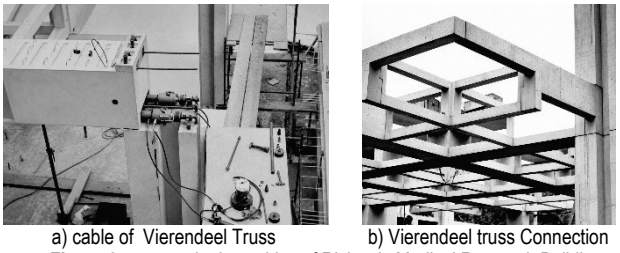
Figure 7 Dome structure and cross-section of the US Pavilion at EXPO 67

As examples of using a Vierendeel truss structure, Louis I Kahn's Richards Medical Research Building and Salk Institute Laboratory Buildings use post-tensioning methods and precast concrete to create column-free spaces inside the laboratories. The floor plan of the Richards Medical Research Building comprises five laboratories, two auxiliary laboratories, and a service cluster that includes ventilation ducts and stairwells. Each laboratory cluster was supported by eight columns, and the slab was constructed using four transverse Vierendeel precast concrete beams. The slab at the end is cantilevered. The Vierendeel truss section was exposed at the end cantilever in the laboratory and constructed using the disjoint method. The interior of the post-tensioned beam accommodates the service piping, and the structural system, service system, and space are mutually integrated. The integration of service facilities through structural systems is a common feature in most of Kahn's buildings [17].



a) Vierendeel Truss b) Truss Diagram

Figure 8 Truss of Richards Medical Research Building



a) cable of Vierendeel Truss b) Vierendeel truss Connection

Figure 9 post-tensioning cables of Richards Medical Research Building

The girder used in the building employs a Vierendeel truss structure. Since there are no diagonal members, the structure resists loads through a combination of bending moments and shear forces supported by the rigid joints.

The lower central section shown in the diagram is marked in red, indicating where the maximum bending moment is concentrated. In contrast, the upper central section is marked in dark blue, indicating where the maximum compressive force occurs. The remaining internal frames, which serve as stress transfer and moment distribution zones, are shown in orange and yellow.

To counteract the central section of the Vierendeel truss, post-tensioning cables are inserted into the truss using the post-tensioning method.

Table 4 Buildings and structural systems using truss structural methods

Building Name	Structural System
· Eames House, Eero Saarinen, California, USA, 1945-47	Truss Structure
· Ford Rotunda, Buckminster Fuller, Michigan, USA, 1953	Truss Structure
· IIT Crown Hall, Mies van der Rohe, Chicago, USA, 1956	Truss structure + Plate Girder
· Richards Medical Research Building, Louis I Kahn, Pennsylvania, U.S.A, 1957-64	skeleton structure + Vierendeel truss
· Salk Institute Laboratory Buildings, Louis I Kahn, California, U.S.A, 1959-65	Bearing Wall + Skeleton Structure + Vierendeel truss
· US Pavilion at EXPO 67, Montreal, Canada, 1967	Truss Structure

4 STABILIZATION METHODS AND ARCHITECTURAL UTILITY IN MODERN ARCHITECTURE

4.1 Structural Systems for Form

The ultimate goal of a building's structural system is structural stabilization. The structural stability can be divided into external and internal equilibria. External equilibrium refers to the state in which stress from above is transmitted along the main structural members to the foundation and subgrade, and the reaction at the foundation balances the upper stress, meaning that it maintains its original state when an external disturbance occurs in the balanced building [13, p. 33]. Internal equilibrium occurs when the strength and stiffness of the members can withstand the internal forces generated within the structure.

The truss structure is a structural system used for roofs. The truss structures for roofs include lattice trusses and geodesic dome structural systems. Both structural systems distribute loads through a triangular structure and respond to the stresses generated on the roof as axial forces. The lattice truss mainly connects relatively short spans within 10m and is constructed in a linear form. The geodesic dome connects relatively wide spans and exhibits a curved shape. Such structural systems respond to axial forces, adopt a disjoint connection method, express details in a technological form, and describe the construction process.

This space is often used for facilities such as exhibitions or assembly halls that require a wide span and a uniform space without columns. The skeleton structure method uses a tapered structural system. The tapered-shaped skeleton structure secures structural stability in places where stress is concentrated to respond to the stress. A tapered-shaped skeleton structure is used in places where stress is concentrated, such as the connection part of columns and beams or the roof end of stadium spectator stands. The member connection method followed the rigid-joint connection method. The joint method visually conceals the connection points and thickening areas where the bending moments concentrate, and conforms to the shape of the bending moment to form the structure.

4.2 Structural Systems for Space

Structural systems for space include the rectangular-shaped skeleton and Vierendeel truss structures. The rectangular-shaped skeleton structure is used as an important structural system for residential architecture as a dom-ino system. In terms of spatial composition, it secured views by placing load-bearing walls on the sidewalls of the residential building and pursued efficiency in the plan composition and spatial arrangement.

A dom-ino slab was placed in the interior space to dimensionalize the residents' range of activities.

This dom-ino structural system has a rectangular form, and the connection of the members adopts a joint method that is not exposed to the outside, thereby influencing the aesthetic form composed of linear elements.

In residential buildings, torsion occurs under horizontal loads depending on the core placement. The location of load-bearing walls or bracings can be determined to minimize torsion while considering architectural planning and space utilization. It is advantageous to position load-bearing walls parallel to the direction of the horizontal loads, and longer walls are more effective than shorter walls.

The rectangular skeleton structure buildings used by Mies employed members with improved cross-sections, such as H-beams and I-beams. Individual members were connected in a disjointed manner and primarily used as columns. Mies pursued a multipurpose space that was diverse, flexible, and flowing. Early residential buildings used load-bearing walls, making it impossible to use large, wide windows; however, as columns and walls were separated, it became possible to construct a fluid space.

The H- and I-shaped steel columns have flange directions aligned with the structural axis. Therefore, the moment connections in the flange direction are structurally advantageous. When steel columns are continuous, the direction of maximum load is often designated as the structural axis. When the columns are arranged discontinuously, the direction of the asymmetry is frequently designated as the structural axis.

The Vierendeel truss, primarily used in Kahn's buildings, serves as a structure for beams, supporting spans of approximately 9 m in laboratories and providing a square-shaped laboratory space. A truss structure for a beam is created by connecting the upper and lower chords using H-beams or I-beams and connecting the two members with braces. This type of truss structure is advantageous in terms of its self-weight because it is lightweight. Moreover, because there is almost no bending stress, the volume of the members can be reduced.

The Vierendeel truss structure is a representative truss structural system for beams with a compressive force at the top and tensile force at the bottom. Reinforced concrete was used for the upper and lower members, which were connected using rigid joints. This is structurally efficient because it distributes and reduces the bending stress inside the members. The spaces in which they were used were mainly places that required a relatively long span of space, such as laboratories or research facilities.

In structural systems for space, the efficient placement of load-bearing walls, the use of steel members with improved cross-sections, and trussed beams are used as elements of stability. In structural systems for space, a structure acts as a medium between spaces. This case differs from pre-modern architecture in that technology directly influences architectural form and space owing to the development of materials and structural systems.

5 CONCLUSION

Architecture has evolved with changes in building materials, joining methods, and structural systems. Advances in structural systems have led to the development of new spaces and forms.

Modern architecture requires relatively large and uniform interior spaces to meet social and cultural needs. Constructing such spaces requires structural elements such as slabs and roofs, with roofs particularly influencing the form of the buildings. Particularly in early modern architecture, building forms were heavily influenced by geometric shapes suitable for steel or reinforced concrete structures. Therefore, early modern buildings are viewed as the origin of contemporary high-tech architecture, which actively exposes structural engineering projects to their external forms.

This study categorizes structural systems into structures for form and structures for space.

Structures for form include truss structures for roofs and tapered frame structural systems. Truss structures for roofs comprise lattice trusses and geodesic domes. The material used was steel, and the connection method employed a disjointed approach, exposed externally, to effectively express the formative characteristics by utilizing fin joints. Lattice trusses combine upper and lower chords to form triangular or square grid patterns, functioning as linear support systems that are used as girders. Geodesic dome trusses combine triangular grids to expand into curved surfaces, thereby creating large-span shell-shaped roofs without intermediate supports. The stability elements of a structural system primarily place the members under axial force conditions. Lattice trusses primarily achieve stability through the arrangement of two-dimensional members by combining bending and axial forces. Geodesic domes achieve stability through three-dimensional curved surfaces and triangular lattices.

Structures designed for this form include tapered-frame systems. They are also called form-active structures and represent the most natural form of response to stress flow, conveying visual stability. The connection method uses joints to ensure that connection points remain concealed. Structural stability is achieved by thickening sections where stress is concentrated, or by arranging members along the load flow paths.

Structures for space include right-angled frame systems and Vierendeel truss structures applied to girders. The rectangular-frame structure efficiently positioned the load-bearing walls to secure daylighting and views, with the walls acting as rigid diaphragms to provide structural stability. Their interior use was predominantly residential. Furthermore, when steel members were used, improved cross-sections were employed by removing sections with low stress, considering the stress distribution. This reduced the self-weight of the members and enhanced their structural stability. Vierendeel trusses eliminate the central portion of the girder because maximum stress occurs at each end. Furthermore, because the bending moments were concentrated in the central section of the Vierendeel truss girders, prestressing was applied to achieve structural stabilization. This system has been applied to interior spaces that require relatively long spans, such as exhibition halls or laboratories.

Buildings constructed using the skeleton structure system employ tapered members and shear walls as stabilizing elements. Tapered structures are primarily applicable to the roofs of buildings requiring long spans, such as auditoriums and assembly halls. Additionally, shear walls serve as stabilizing elements that bear vertical loads while resisting horizontal forces. This system is primarily used in residential buildings, and structural stability is achieved by arranging shear walls parallel to the horizontal forces. Therefore, for buildings with rectangular floor plans, such as residential structures, it is advantageous to prioritize the skeleton structure system; furthermore, incorporating tapered cross-sections allows the structure to adequately resist bending moments.

Buildings using a truss structure minimize bending moments and ensure structural efficiency by directing axial forces into the primary stress system. Through the repetitive combination of linear members, they simultaneously achieve structural stability and architectural form. Therefore, for circular floor plans or when column-free interior spaces are required, prioritizing an appropriate truss system will enable efficient integrated design.

This study examined how the stabilizing elements of skeleton and truss structures influence the form and spatial design of buildings, focusing on early Modern architecture. However, the study has limitations in that the case studies were confined to combinations of two-dimensional linear members and some three-dimensional lattice trusses, and thus did not fully encompass the mechanical behavior of the various non-standard curved surfaces and column-free spaces found in contemporary architecture. To address this, future research will build upon the methodology of this study and focus on large-scale, column-free buildings by analyzing three-dimensional stress flow. Specifically, we plan to investigate how form-active and surface-active structural systems—such as cable, membrane, and shell structures—control and interact with spatial and exterior forms.

6 REFERENCES

[1] Sev, A., & Tugrul, F. (2014). Integration of architectural design with structural form in non-orthogonal high-rise buildings. *Journal of Sustainable Architecture and Civil Engineering, 7*(2). https://doi.org/10.5755/j01.sace.7.2.7046

[2] Uihlein, M. S. (2015). Characteristics of structural and architectural integration as identified by U.S. architectural practitioners. *Journal of Architectural and Planning Research, 32*(3), 258-270. https://www.jstor.org/stable/44113113

[3] Li, L., Wang, J., & Wu, J. (2015). Experiments based integral design of architecture and structure. *Proceedings of the International Conference on Structural, Mechanical and Materials Engineering (ICSMME 2015)*, 98-102. https://doi.org/10.2991/icsmme-15.2015.23

[4] Jin-ho, C., Yeon-soo, P., & Byung-hae, L. (2000). A design object model for the implementation of an integrated structural design system for building structures. *Journal of the Computational Structural Engineering Institute of Korea, 13*(1).

[5] Miloshevska Janakieska, M. (2022). The structural system's influence on the architectural features of the buildings. *International Journal of Technical and Natural Sciences, 3*(1), 3-10.

[6] El-Khadrawy, A. (2023). Role of Structures Inspired by Nature to achieve Sustainable Architectural Buildings. *Texas Journal of Engineering and Technology*.

[7] Kookalani, S., et al. (2024). Trajectory of building and structural design automation from generative design towards the integration of deep generative models and optimization. *Journal of Building Engineering, 97*. https://doi.org/10.1016/j.jobe.2024.110972

[8] Kozan, F., et al. (2024). The Effect of Structural System Elements on the Dynamic Behavior of High-Rise Buildings. *International Journal of Integrated Engineering, 16*(9), 80-95. https://doi.org/10.30880/ijie.2024.16.09.006

[9] Acican, O., & Luyten, L. (2025). Structurally integrated architectural design education: A computational design approach. *Architectural Engineering and Design Management*, 1–23. https://doi.org/10.1080/17452007.2025.2553002

[10] Rinke, M., & Pacquée, R. (2022). Structures and change–Tracing adaptability based on structural porosity in converted buildings. *Architecture, Structures and Construction, 2*(4), 699-710. https://doi.org/10.1007/s44150-022-00054-9

[11] Sang-heon, L. (2002). *Iron and the development of modern architectural theory*. Baleon Press.

[12] Macdonald, A. J. (1994). Structure and architecture. Butterworth-Heinemann.

[13] In-Seon, H. (2003). *Structure of the structure*. Baleon Press.

[14] Ford, E. R. (1996). *The details of modern architecture* (Vols. 1–2). MIT Press.

[15] Ching, F. D. K. (2014). *Building structures illustrated*. Wiley.

[16] Wright, S. H. (1989). *Sourcebook of contemporary North American architecture: From postwar to postmodern*. Van Nostrand Reinhold.

[17] Nakjoong, K. (1999). A study on the structural characteristics of Louis Kahn's architecture (*Doctoral dissertation*, Seoul National University).

Author's contacts:

Sung Hyun Jo
School of Architectural Engineering, Gyeongsang National University,
501 Jinju-daero, Jinju-si, Gyeongsangnam-do, 52828, Republic of Korea
2012yu@naver.com

CC BY

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20260504173433

Original scientific paper
Received: 2026-05-04, Accepted: 2026-05-12

The Influence of Focal Length and Aperture on the Perception of Business Portraits

Vesna Uglješić*, Miroslav Mikota, Ivana Pavlović, Jana Žiljak Gršić

Abstract: Business portrait photographs are often perceived as neutral representations of professionals, although their impression is shaped by a series of technical decisions made during shooting. This paper examines how combinations of lens focal length and aperture, i.e. depth of field, influence the selection of business portraits regarding different communication effects. Six models, three male and three female, were photographed under controlled studio conditions using nine parameter combinations: 28 mm, 85 mm and 180 mm, and f/2.8, f/5.6 and f/11. For each model, respondents viewed a composite image made of nine photographs of the same person and selected the photograph that best matched the given impression. The analysis included selection frequencies, post-hoc comparisons and a robustness check using a logistic model. The results indicate that individual parameter combinations are systematically associated with different evaluations of the business portrait, opening space for empirically grounded shooting guidelines.

Keywords: aperture; business portrait photograph; depth of field; face perception; focal length

1 INTRODUCTION

Business portraits are a specific form of photography at the intersection of art and a clearly expressed function. In this paper, the term *business portrait photograph* refers to a photographic portrait intended for the professional identification and representation of a person in business and organisational contexts, such as organisational websites, professional networks, job-search platforms, annual reports and similar communication channels. Unlike general portrait photography, documentary portraiture, actor headshots or informal profile photographs, the business portrait is not primarily narrative, expressive or promotional in a personal sense, but is expected to communicate professional presence, credibility and suitability within a specific institutional or business environment. In professional literature, business portrait photographs become more clearly distinguished in the early 2000s, when they begin to be defined as a subvariant of the general photographic portrait, with their own principles and requirements [1]. While they follow the basic rules of portrait photography, what separates them from the classical, beautiful or documentary portrait is their pronounced utility.

Although often viewed as a neutral representation of employees or experts in a certain profession, business portraits are never completely neutral. They participate in shaping first impressions of the portrayed person, and these impressions can also be transferred to the organisation that person represents. Business portrait photography therefore does not record only the appearance of the individual, but communicates a range of qualities such as trust, approachability, competence and professional credibility. Literature on employer image and employer branding shows that the visual presentation of people within an organisation can shape the perceived attractiveness of the employer and influence behavioural intentions, for example the decision to apply for a job [2, 3].

Social-cognitive research helps explain why photographs of faces function as such powerful signals. Observers form first impressions of faces after an extremely short time, sometimes after as little as 100 ms [4]. These impressions are mostly based on a small number of socially significant dimensions. One influential framework emphasises warmth and competence as two central dimensions of social cognition [5], while another framework describes face evaluation through valence, or trustworthiness, and dominance [6]. What they share is that we do not observe a face only as a visual object, but very quickly draw conclusions from it about a person's intentions, abilities and social position.

It is precisely here that the photographic medium stops being only a transmitter. Contemporary reviews of social attributions from faces show that judgements of trustworthiness, competence and related traits are sensitive to visual cues, including those produced through photographic mediation [7]. A portrait that appears natural, clear and harmonious can therefore be understood as evidence of the personal qualities of the portrayed person, even though such an effect is at least partly produced by the photographer's technical choices.

Among these choices, lens focal length and aperture, i.e. depth of field, stand out in particular. Focal length, together with the distance between camera and subject, systematically changes the geometry of the facial depiction and the perception of its proportions [8, 9]. The effect of selfie photography, in which the face is photographed from a very short distance and with wide-angle optics, shows that such conditions can emphasise central facial features and disturb the perception of proportions [10]. Depth of field introduces a different type of change: it does not change the geometry of the face, but the relationship between the face and its surroundings. By controlling the distribution of sharpness and blur, depth of field changes how much of the surrounding context remains visible, how strongly the subject stands out from the background, and how easily attention is directed towards the face [11–13].

Despite the relevant findings, there remains a gap between what is routinely adjusted in photographic practice and what has been empirically verified specifically in the business portrait. Previous studies mainly address general face perception, distortion when taking selfies, attractiveness, or isolated technical variables, while the

business portrait is often assumed as a practical format rather than treated as a separate research object. This study therefore examines how different combinations of focal length and aperture are associated with the selection of business portraits with regard to impressions such as warmth, competence, trustworthiness, dominance, business suitability and the naturalness of facial proportions.

2 MATERIAL AND METHODS

2.1 Photographic Material

The research was conducted on a set of business portrait photographs taken under controlled studio conditions, using a full-frame camera (Nikon Z7II), and Nikkor Z 28 mm f/2.8 and Nikkor Z 70–180 mm f/2.8 lenses. Six models were photographed, three male and three female. Each model was photographed in nine parameter combinations, obtained by combining three lens focal lengths and three apertures. The focal lengths used were 28 mm, 85 mm and 180 mm, while the apertures were f/2.8, f/5.6 and f/11. The selected focal lengths and aperture values were chosen to cover a controlled range of conditions relevant to portrait practice on a full-frame camera. The 28 mm focal length represented a wide-angle condition, in which the camera was positioned close to the portrayed person and perspective distortion was expected to be more visible. The 85 mm focal length represented a commonly used portrait setting, while 180 mm represented a longer telephoto condition, associated with greater camera-to-subject distance and reduced perspective distortion. The aperture values were selected to represent a shallow depth of field (f/2.8), a moderate depth of field (f/5.6) and a deeper depth of field with greater overall sharpness (f/11). This range made it possible to examine both changes in facial geometry related to focal length and camera distance, and changes in the distribution of sharpness and blur related to aperture.

The models were photographed from eye level, at distances of 60 cm (for the 28 mm focal length), 190 cm (for the 85 mm focal length) and 400 cm (for the 180 mm focal length).The other shooting conditions were kept constant: neutral grey background, studio lighting, similar framing, gaze directed towards the camera and a mild, unobtrusive expression. In this way, the influence of factors that are not the subject of this paper was reduced, while variations resulting from changes in focal length and depth of field were retained.

For each model, a composite image with nine photographs of the same person was created (as shown in Fig. 1). The position of the photographs within the image was random, so that the position did not follow the order of photographic parameters. Each position was marked with a number from 1 to 9, while a separate coding table recorded which position corresponded to which parameter combination.



Figure 1 Example of composite image

2.2 Respondents and Procedure

The survey was conducted online, using Google Forms. Students participated in the study, because they can be viewed as a group close to early-career job seekers. A total of 119 responses were collected. After applying the basic inclusion and exclusion criteria, 113 respondents entered the analytical dataset. Respondents older than 35, those who did

not pass the attention check or those whose responses were not sufficiently complete were excluded from the analysis. If a respondent stated that they knew a particular portrayed model, not all their responses were excluded from the analysis, but only those related to that model (Tab. 1).

For each model, respondents viewed all nine photographs simultaneously and, for each statement, selected one photograph that in their impression best matched the description. The order of photographs within each composite image was randomised and marked with numbers from 1 to 9. For selected constructs, negatively formulated statements were also included, in which respondents were asked to select the photograph that least matched the description. The same photograph could be selected for several statements.

Table 1 Sample structure and data cleaning

Element	Value
Total number of responses	119
Number of valid respondents	113
Number of portrayed models	6
Number of photographs per model	9
Total number of analysed model evaluations	666
Total number of individual selections	6660

A forced-choice procedure was used because the aim of the study was not to measure absolute ratings of individual photographs, but to identify relative preferences among technically different versions of the same business portrait. By presenting nine photographs of the same person simultaneously, the influence of differences between portrayed individuals was reduced, while respondents were directed to compare subtle variations resulting from focal length and aperture. This approach made it possible to examine which parameter combinations were more often associated with a particular impression within the same portrayed model.

The statements referred to the following impressions: the most positive general impression, warmth and approachability, competence and expertise, trustworthiness, dominance and authority, suitability as a business portrait photograph, and naturalness and proportionality of the face. Negatively formulated statements referred to the lowest suitability, the lowest naturalness and proportionality of the face, and the lowest trustworthiness.

2.3 Data Analysis

The responses from the Google Forms table were translated, using a coding table, into an analytical format: respondent – model – statement – selected position – focal length – aperture. Only valid responses were included in the analysis; when a respondent reported knowing a particular portrayed model, only that respondent's answers for that model were excluded. For each statement, frequencies and percentages of selecting individual focal lengths, apertures and parameter combinations were calculated. This made it possible to compare both isolated parameter levels and their specific combinations. The distributions of selections were tested using chi-square tests, and standardised residuals were used to determine which combinations deviated the most from the expected distribution.

For statements with a positive and a negative pole, a net indicator was also calculated, as the difference between the number of selections of the positive and the negative statement. In addition, a robustness check was conducted using a logistic choice model with clustering by respondent and with the effect of the portrayed model included. In this model, each parameter combination was treated as either selected or not selected for a given statement. This analysis was used to examine whether the patterns found in the frequency-based results remained stable when repeated responses from the same respondent and differences between portrayed models were taken into account. The results of the robustness check are presented as odds ratios (OR) and p-values corrected using the Benjamini–Hochberg procedure.

3 RESULTS

The results show that selections of photographs are not distributed randomly across parameter combinations, but that individual combinations are systematically associated with certain impressions. Since respondents selected one photograph from a set of nine photographs of the same person, the results are interpreted as relative preferences within each model, rather than as absolute ratings of individual photographs. The clearest pattern is visible in the distinction between a warm and positive impression on the one hand, and business suitability, competence and trustworthiness on the other. The most frequently selected parameter combinations for each statement, together with the corresponding percentages and robustness-check results, are shown in Tab. 2. The percentages represent the share of selections within each individual statement, while the robustness-check column shows whether the same parameter combination remained statistically relevant in the logistic model.

Table 2 Most frequently selected parameter combinations by statement

Statement	Most frequent combination	Share of selections	Robustness check
Most positive general impression	85 mm / f/2.8	23,1%	OR = 1,40; pBH = 0,016
Warmest and most approachable	85 mm / f/2.8	27,9%	OR = 2,12; pBH < 0,001
Most competent and expert	180 mm / f/5.6	32,1%	OR = 2,09; pBH < 0,001
Most trustworthy	180 mm / f/5.6	28,4%	OR = 1,79; pBH < 0,001
Most dominant / authoritative	180 mm / f/5.6	30,6%	OR = 1,67; pBH < 0,001
Most suitable as a business portrait	180 mm / f/5.6	32,0%	OR = 2,14; pBH < 0,001
Most natural / proportional face	180 mm / f/5.6	21,6%	not clearly separated
Least suitable	28 mm / f/2.8	36,9%	OR = 1,49; pBH = 0,007
Least natural / proportional face	28 mm / f/2.8	38,3%	OR = 1,68; pBH < 0,001
Least trustworthy	28 mm / f/2.8	28,1%	not clearly separated

The combination 85 mm / f/2.8 was most frequently selected for the most positive general impression (23,1%) and for warmth and approachability (27,9%). The medium focal length combined with a shallower depth of field therefore appears as the setting that most strongly supports an impression of pleasantness, closeness and approachability.

The combination 180 mm / f/5.6 was most frequently selected for competence and expertise (32,1%), trustworthiness (28,4%), dominance and authority (30,6%) and suitability as a business portrait photograph (32,0%). The same combination was also most frequently selected for naturalness and proportionality of the face (21,6%), although in that statement it was not statistically clearly separated from the second most frequent combination (180 mm / f/2.8 – 18,9%).

The opposite pattern is visible in negatively formulated statements. The combination 28 mm / f/2.8 was most frequently selected as the least suitable for a business portrait photograph (36,9%) and as the combination in which the face appears least natural and proportional (38,3%). The same combination was also most frequently selected as the least trustworthy (28,1%), but for this statement it was not statistically clearly separated from the combination 28 mm / f/5.6.

The net indicators for pairs of statements further support the difference between settings associated with desirable and undesirable impressions. For suitability as a business portrait and naturalness/proportionality of the face, combinations with 180 mm show a positive net effect, while combinations with 28 mm, especially 28 mm / f/2.8, show a strongly negative net effect. For trustworthiness, the pattern is similar but somewhat less pronounced, which suggests that trustworthiness is probably shaped by a combination of several visual factors, rather than by facial geometry alone.

The distribution of selections across all statements and parameter combinations is shown in Tab. 3. The heatmap values represent the percentage of selections for each parameter combination within a given statement; therefore, each row shows how selections for one statement were distributed across the nine focal length and aperture combinations. Darker cells indicate combinations that were selected more frequently within that statement, making visible the clustering of positive, professional and negative impressions around different photographic settings.

Table 3 Heatmap: statements × parameter combinations

Statement	28 mm / f/2.8%	28 mm / f/5.6%	28 mm / f/11%	85 mm / f/2.8%	85 mm / f/5.6%	85 mm / f/11%	180 mm / f/2.8%	180 mm / f/5.6%	180 mm / f/11%
Most positive general impression	3,30	1,80	6,61	23,12	11,86	11,41	17,72	13,06	11,11
Warmest and most approachable	3,60	5,26	11,26	27,93	10,36	9,91	15,47	6,31	9,91
Most competent and expert	1,20	1,20	1,65	8,56	12,46	6,76	17,57	32,13	18,47
Most trustworthy	1,65	1,65	1,95	9,91	12,01	8,86	17,42	28,38	18,17
Most dominant / authoritative	1,35	1,50	2,40	5,41	10,81	6,91	20,12	30,63	20,87
Most suitable as a business portrait	2,10	1,65	0,90	9,16	10,36	8,41	17,42	31,98	18,02
Most natural / proportional face	4,05	2,85	1,20	12,61	12,91	9,46	18,92	21,62	16,37
Least suitable as a business portrait	36,94	28,23	22,22	2,70	1,95	1,65	2,25	2,10	1,95
Least natural / proportional face	38,29	27,03	21,32	2,55	2,10	1,95	3,15	1,35	2,25
Least trustworthy	28,08	24,47	23,12	4,50	2,70	4,20	5,11	3,15	4,65

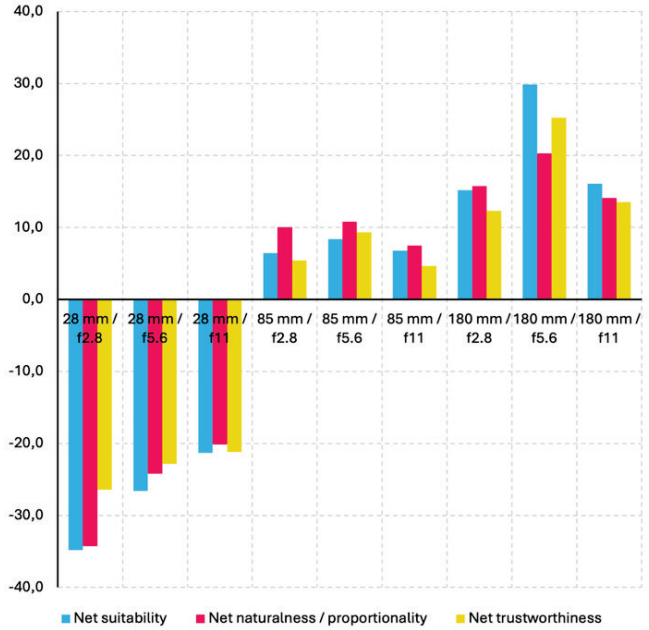


Figure 2 Net selection index by focal length and aperture combination

Fig. 2 shows the net selection index calculated from corresponding positive and negative versions of the same statements. The index was calculated by subtracting the percentage of selections for the negative statement from the percentage of selections for the corresponding positive statement for each parameter combination. Positive values therefore indicate that a combination was selected more often in a favourable context, while negative values indicate that it was selected more often in an unfavourable context. This makes it possible to distinguish combinations that are simply often selected from those that are more clearly polarised towards a desirable or undesirable interpretation.

For suitability as a business portrait and naturalness / proportionality of the face, combinations with 180 mm have a positive net effect, while combinations with 28 mm, especially 28 mm / f/2.8, have a markedly negative net effect. For trustworthiness, the pattern is similar but somewhat less pronounced, which suggests that trustworthiness is probably shaped by a combination of several visual factors, and not only by facial geometry.

4 DISCUSSION

The results of the study indicate that the perception of a business portrait is not determined only by an individual technical parameter, but by the combination of focal length and aperture. In photographic practice, these parameters are most often treated as basic tools of portrait shaping [1], while existing research mainly considers the effects of focal length, perspective distortion or blur separately [8, 10, 13]. This study shows that different combinations of the same parameters can produce different communication effects: one setting may appear warmer and more approachable, while another may be more clearly associated with competence, trustworthiness and business suitability.

The medium focal length of 85 mm in combination with f/2.8 was most often associated with a positive general impression and warmth / approachability. Such a result is understandable because 85 mm preserves relatively natural facial proportions, while the shallower depth of field reduces the amount of visual information and directs attention towards the focus area. In a business context, this combination can be suitable when the aim is to achieve an impression of approachability, pleasantness and closeness.

On the other hand, the combination 180 mm / f/5.6 was most often associated with competence, trustworthiness, dominance and the general suitability of the business portrait. A longer focal length reduces perspective distortion and provides a more stable depiction of facial proportions, while the medium aperture maintains sufficient sharpness and control of the image without excessive blur. In this way, the photograph can appear more professional, formal and secure, but not necessarily the warmest.

The clearest negative evaluations are associated with 28 mm / f/2.8. This is consistent with previous research on the influence of wide-angle shooting and short camera-to-subject distance on the perception of facial proportions [8-10]. In business portraiture, such a combination may be risky because viewers do not necessarily read distortion or unnaturalness of the image as a technical effect, but as an unfavourable impression of the person. The results therefore clearly show that technical parameters do not remain only technical, but participate in the formation of social attributions.

The findings also indicate that there is no single universally best setting for a business portrait. If the communication goal is warmth and approachability, a medium focal length with a shallow depth of field may be a more appropriate choice. If the goal is to emphasise professionalism, competence and trustworthiness, a longer focal length with a moderate aperture shows a more stable pattern. This distinction is important for the practical production of business portraits because it allows shooting decisions to be based not only on the intuition or aesthetic preference of the photographer, but also on the expected communication effect.

For professional photographers, the results suggest that focal length and aperture should be considered as part of the communicative strategy of the portrait, not only as technical or stylistic choices. In practice, this means that the same person may need to be photographed differently depending on whether the intended impression is approachability, authority, reliability or professional neutrality. Such an approach can support a more deliberate dialogue between photographer and client before the shooting process, especially when the portrait is intended for organisational websites, professional profiles or public communication materials.

The findings are also relevant for employer branding and organisational communication. Business portraits often appear as part of a wider visual identity, where they contribute to the way employees, teams and organisations are perceived. If all portraits within an organisation follow the same visual standard, they may support consistency; however, the desired impression may still differ depending on the role, context or audience. For example, portraits of team leaders, consultants, lecturers or customer-facing employees may require different balances of approachability, competence and authority. The results therefore suggest that business portrait guidelines should not define only formal consistency, but also the intended communication effect of the image.

The forced-choice format used in this study was useful because it allowed respondents to compare subtle technical variants of the same portrait directly. Presenting all nine versions of the same person at the same time reduced the influence of differences between portrayed individuals and directed attention towards the effects of focal length and aperture. At the same time, this approach necessarily limits the interpretation of the results. The study does not measure how warm, competent or trustworthy each photograph is in an absolute sense, but shows which technical variant is more often associated with a particular impression in comparison with other versions of the same portrait. It also does not allow the intensity of perceived differences to be measured.

Another limitation concerns the controlled studio setting. The neutral grey background was used deliberately in order to reduce the influence of contextual and visual variables that were not the focus of the study. However, such a background may also have reduced the visible effect of depth of field. Since the background contained little visual detail, differences in blur and background separation were less pronounced than they might be in a more complex environment, such as an office interior, outdoor setting or branded organisational space. This may help explain why aperture did not emerge as the strongest or most consistent factor for all impressions, and why the effects of focal length and facial geometry appear more clearly in some results. The findings related to aperture should therefore be interpreted within the specific conditions of a controlled, visually neutral studio portrait.

A further limitation concerns the sample, which consisted of students. Although this is a relevant group in the context of early career self-presentation and job seeking, the results should be further tested on samples of HR professionals, photographers and a broader professional population. Future research should also examine whether the role of aperture becomes more pronounced when portraits are taken in visually richer environments, where depth of field can more strongly influence the separation of the portrayed person from the background.

Despite these limitations, the study shows that even a relatively simple experimental design can provide useful

guidelines for practice. Systematic variation of focal length and aperture allows a clearer understanding of how technical decisions in portrait photography are translated into perceptual and communication effects. The results therefore show that shooting parameters do not only affect the visual appearance of a photograph, but also indirectly shape the way in which viewers perceive the portrayed person through changes in facial geometry and the distribution of depth of field.

5 CONCLUSION

The study shows that combinations of focal length and aperture have a measurable influence on the choice of business photographic portraits in relation to different impressions. The combination 85 mm / f/2.8 was most often associated with a positive and approachable impression, while the combination 180 mm / f/5.6 was most often associated with competence, trustworthiness, dominance and suitability as a business portrait. In contrast, 28 mm / f/2.8 was most often associated with the least desirable evaluations, particularly with unsuitability and unnatural facial proportions.

The results indicate that a business portrait can be understood as a technically produced communication object, and not only as a neutral representation of a person. The choice of shooting parameters should therefore be connected with the communication goal: different settings may be suitable for an impression of approachability, while others may be more suitable for an impression of professional stability and competence. In this way, the study provides practical guidelines for the selection of technical parameters in the production of business photographic portraits.

Future research should examine whether the same patterns remain stable in other respondent groups, especially among HR specialists, photographers, designers, employers and professionals involved in organisational communication. Such groups could provide insight into whether the impressions found in the student sample are also relevant in contexts in which business portraits are professionally produced, selected or evaluated. It would also be useful to test whether similar effects occur in less controlled environments, such as workplace interiors or branded organisational settings. Such extensions would make it possible to connect controlled photographic parameters more directly with everyday professional portrait practice, employer branding and real-world decisions about visual professional representation.

6 REFERENCES

[1] Amyot, A. (2003). *The art and techniques of business portrait photography*. Amherst Media.

[2] Lievens, F., & Slaughter, J. E. (2016). Employer image and employer branding: What we know and what we need to know. *Annual Review of Organizational Psychology and Organizational Behavior, 3*, 407–440. https://doi.org/10.1146/annurev-orgpsych-041015-062501

[3] Lyons, B. D., & Marler, J. H. (2011). Got image? Examining organizational image in web recruitment. *Journal of Managerial Psychology, 26*(1), 58–76. https://doi.org/10.1108/02683941111099628

[4] Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17*(7), 592–598. https://doi.org/10.1111/j.1467-9280.2006.01750.x

[5] Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11*(2), 77–83. https://doi.org/10.1016/j.tics.2006.11.005

[6] Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences, 105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

[7] Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology, 66*, 519–545. https://doi.org/10.1146/annurev-psych-113011-143831

[8] Třebický, V., Fialová, J., Kleisner, K., & Havlíček, J. (2016). Focal length affects depicted shape and perception of facial images. *PLOS ONE, 11*(2), e0149313. https://doi.org/10.1371/journal.pone.0149313

[9] Banks, M. S., Cooper, E. A., & Piazza, E. A. (2014). Camera focal length and the perception of pictures. *Ecological Psychology, 26*(1–2), 30–46. https://doi.org/10.1080/10407413.2014.877284

[10] Ward, B., Ward, M., Fried, O., & Paskhover, B. (2018). Nasal distortion in short-distance photographs: The selfie effect. *JAMA Facial Plastic Surgery, 20*(4), 272–273. https://doi.org/10.1001/jamafacial.2018.0009

[11] Nefs, H. T. (2012). Depth of field affects perceived depth–width ratios in photographs of natural scenes. *Seeing and Perceiving, 25*(6), 577–595. https://doi.org/10.1163/18784763-00002400

[12] Sadr, J., & Krowicki, L. (2019). Face perception loves a challenge: Less information sparks more attraction. *Vision Research, 157*, 61–83. https://doi.org/10.1016/j.visres.2019.01.009

[13] Enns, J., & MacDonald, S. C. (2013). The role of clarity and blur in guiding visual attention in photographs. *Journal of Experimental Psychology: Human Perception and Performance, 39*, 568–578. https://doi.org/10.1037/a0029877

Authors' contacts:

Vesna Uglješić, senior lecturer
(Corresponding author)
Zagreb University of Applied Sciences,
Vrbik 8, 10000 Zagreb, Croatia
vugljesic@tvz.hr

Miroslav Mikota, PhD, associate professor
University of Zagreb Faculty of Graphic Arts,
Getaldićeva ulica 2, 10000 Zagreb, Croatia
miroslav.mikota@grf.unizg.hr

Ivana Pavlović, PhD, assistant professor
University of Zagreb Faculty of Graphic Arts,
Getaldićeva ulica 2, 10000 Zagreb, Croatia
ivana.pavlovic@grf.unizg.hr

Jana Žiljak Gršić, PhD, full professor
Zagreb University of Applied Sciences,
Vrbik 8, 10000 Zagreb, Croatia
jana@tvz.hr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20260518173334

Original scientific paper
Received: 2026-05-18, Accepted: 2026-06-23

Numerical Modeling of Hydrodynamic Vibrations in the Flow Passage of Centrifugal Pumps

Amir G. Mustafayev, Sanam J. Muslimov*

Abstract: The numerical modeling method enables researchers to study hydrodynamic processes through pump flow passage testing which serves as evidence to support modifications in existing centrifugal pump design methods. Hydrodynamic vibration critically impacts centrifugal pump reliability and performance under off-design conditions. This study presents a numerical model coupling rotor dynamics with hydraulic characteristics to analyze vibration's effect on flow rate and head. The approach integrates mechanical motion equations with vibration-dependent hydraulic loss models and harmonic-response solutions. We propose a physical mechanism where vibration arises from transit-inertial flow interactions in impeller passages, periodically disturbed by the volute tongue. The model predicts predominant vibration at blade-passing frequency, strongest transversely. Validation used vibration data from NM-type oil pipeline pumps at reduced flow. Results show flow rate departures increase hydrodynamic instability, elevating vibration and reducing hydraulic performance. This model enables vibration diagnostics and pump design optimization.

Keywords: centrifugal pumps; energy characteristics; fluid flow; hydrodynamic vibration; head and flow rate

1 INTRODUCTION

One of the most common types of hydrodynamic machines is the centrifugal pump.

During operation, these machines are subjected to a specific hydraulic regime accompanied by structural vibration. This complicates the operation of hydraulic machines, shortens the interval between repairs, and ultimately leads to undesirable failures.

Hydraulic machines are designed for specific operating parameters. One type of hydraulic machine is the centrifugal pump, which is designed on the basis of fluid capacity (flow rate) and the head developed by the pump.

The operating mode of centrifugal pumps is mainly governed by these two parameters. Although studying these parameters is associated with deviations in pump operation and with the minimization of energy losses during service, it also makes it possible to preserve pump components and prevent damage from vibration-induced hydraulic impacts.

Studies show that the higher the power of the pumping unit and the stronger the unsteady pulsations, the greater the flow disturbances in the pump flow passage. In this case, the resulting flows hinder fluid passage through the hydraulic machine, reduce pump capacity, and increase energy losses. During pulsation, the force and impact loads acting on the elements of the pump flow passage intensify significantly and transform into oscillatory motion. The consequences of vibration can be extremely destructive.

A high vibration level arises when pumps are operated outside their normal operating zone. Its negative effects may occur even during short-term operation under such conditions. If pumps are operated in these modes for a long time, the consequences of vibration may become severe and hazardous.

At present, at most oil pipeline transport facilities, due to the transition to a market economy, this situation has changed. However, the acute problem on existing trunk lines has not been eliminated, since they are oriented toward maximum oil supply to the market while not having exhausted their service life or while that service life is being successfully extended.

Hydrodynamic equipment must be operated, under the existing circumstances, together with the means of controlling its operating modes that were available at the time of commissioning.

This problem creates the need to search for methods and means that prevent or substantially reduce the vibration level of centrifugal pumps operating under reduced flow conditions.

The problem under consideration is relevant to all industries in which centrifugal pumps are used.

Under the current conditions, these methods and means impose constraints on the design of hydraulic machines. This is because, on the one hand, the phenomena under study occur either within these machines themselves, and on the other hand, the existing methods for controlling the operating modes of centrifugal pumps do not make it possible to influence these phenomena in the required directions. Throttling the flow at the station outlet and reducing the number of pumps connected in series only make it possible to reduce the volume of liquid passing through the pumps.

2 THEORETICAL PART

2.1 Differential Equations Describing the Dependence of Pump Flow Rate on Vibration

To assess the dependence of the flow rate Q on vibration in the vertical, transverse, and axial directions, a dynamic model of the following form is usually used:

$$Q = Q_0 - k_v A_v - k_t A_t - k_a A_a \quad (1)$$

Where, A_v, A_t, A_a - are the vibration amplitudes in the vertical, transverse, and axial directions, and k_v, k_t, k_a - are the coefficients describing the influence of vibration on the hydraulic efficiency of the pump.

Such a linear approach is based on rotor efficiency degradation models due to vibration proposed in the classical theory of centrifugal pumps and confirmed in rotor dynamics studies [1-5].

However, for accurate modeling, a system of differential equations is used:

$$m\ddot{x} + c\dot{x} + kx = F_{\text{hyd}}\left(Q, H\right) + F_{\text{unbalance}}, \quad (2)$$

where: $x(t)$ - the rotor displacement; c - the damping coefficient; k - the stiffness; F_{hyd} - the hydrodynamic force depending on flow rate and head; $F_{\text{unbalance}} = me\omega^2\sin(\omega t)$ - the force caused by imbalance [6, 7].

The hydrodynamic force is determined through a linear approximation of losses:

$$F_{\text{hyd}} = \alpha_1 Q + \alpha_2 H, \quad (3)$$

which is confirmed by studies on hydraulic transient processes [8, 9].

2.1.1 Differential Equations Describing the Dependence of Pump Head on Vibration

For centrifugal pumps, the head H under vibration is determined through a modified Euler equation [2, 10]:

$$H = H_0 - \beta_{\text{v}} A_{\text{v}}^2 - \beta_{\text{t}} A_{\text{t}}^2 - \beta_{\text{a}} A_{\text{a}}^2 \quad (4)$$

The quadratic effect of vibration is associated with losses of kinetic energy due to non-axial flow rotation and is confirmed in the theory of unstable processes [11, 12].

A more advanced model takes pressure dynamics into account:

$$\frac{\mathrm{d}H}{\mathrm{d}t} + aH = bQ - \gamma A(t), \quad (5)$$

where: $A(t)$ is the resultant vibration amplitude; the coefficients α, β, γ are determined from experimental ESP curves [13, 14].

After this, we write the dynamic flow-head-vibration model.

Combining (2), (3), (4), and (5), we obtain the system:

$$\begin{cases} m\ddot{x} + c\dot{x} + kx = \alpha_1 Q + \alpha_2 H + me\omega^2\sin(\omega t), \\ Q = Q_0 - k_{\text{v}} A_{\text{v}} - k_{\text{t}} A_{\text{t}} - k_{\text{a}} A_{\text{a}}, \\ H = H_0 - \beta_{\text{v}} A_{\text{v}}^2 - \beta_{\text{t}} A_{\text{t}}^2 - \beta_{\text{a}} A_{\text{a}}^2. \end{cases} \quad (6)$$

This system is used in ESP vibration models; see [15-17].

Studies show that, due to hydraulic losses and increased recirculation caused by vibration, the flow rate decreases linearly [1, 2, 4], while vibration disturbs the pressure distribution and the axial acceleration of the flow, causing the head to decrease quadratically [10, 11].

This ESP model is used in practice for failure diagnostics and for selecting operating modes [14, 15].

2.1.2 Basic Rotor Dynamics Model

For each spatial component (vertical, transverse, and axial), the motion of the rotor (pump assembly) can be described by the classical equation with external excitation forces:

$$m\ddot{x}(t) + c_x\dot{x}(t) + k_x x(t) = F_x(t) \quad (7)$$

$$m\ddot{y}(t) + c_y\dot{y}(t) + k_y y(t) = F_y(t) \quad (8)$$

$$m_z\ddot{z}(t) + c_z\dot{z}(t) + k_z z(t) = F_z(t) \quad (9)$$

Where: m - the effective mass (for transverse oscillations) or m_z is the axial mass; $c_{(\cdot)}$ - damping; $k_{(\cdot)}$ - the stiffness of the supports/system; $F_{(\cdot)}(t)$ - the external disturbing force set (rotor unbalance, hydrodynamic pulsations, gas action, impacts, contacts, etc.).

Katsuhiko Ogata's book (2010) presents a systematic theoretical study and exposition of methods for the analysis and synthesis of control systems. It develops mathematical models of dynamic systems and introduces optimal control. However, the book is strongly focused on linear systems; nonlinear methods are covered only to a limited extent, and less attention is given to modern, adaptive, and intelligent methods [18].

Typical components of $F(t)$:

Forced forces due to imbalance: $F_{\text{unb}}(t) = me\omega^2\cos(\omega t)$ (e - eccentricity; ω - rotational frequency), hydrodynamic forces depending on Q, H and operating mode:

$F_{\text{hyd}}(t) = f_{\text{hyd}}(Q, H, \dot{Q}, ...)$; random/noise disturbances F_{noise}.

2.1.3 Model of Vibration-Hydraulics Interaction

A pump is an interconnected hydrodynamic-mechanical system. To obtain the relationship between vibration amplitude and changes in Q and H, two groups of equations are usually used:

1) Hydraulic equations (pump state)

Classical static pump characteristic:

$$H = H_{\text{pump}}(Q, \omega) \text{ (for example, } H = H_0(\omega) - aQ^2) \quad (10)$$

But under vibration, hydrodynamic losses and inlet conditions change. A small-signal relationship can be written as:

$$\Delta H = G_{H,Q}\Delta Q + G_{H,x}x + G_{H,y}y + G_{H,z}z + ... \quad (11)$$

$$\Delta Q = G_{Q,H}\Delta H + G_{Q,x}x + G_{Q,y}y + G_{Q,z}z + ... \quad (12)$$

where $G_{H,x}$ is the sensitivity of head to displacement/vibration along x etc. These coefficients are obtained either from CFD/FSI calculations or empirically.

2) Coupled mechanics-hydraulics system

Substituting relations (11) and (12) into Eq. (2), we obtain coupled ODEs with feedback: mechanics excites the flow, and the flow changes the forces in the mechanical subsystem.

In simplified form:

$$m\ddot{x} + c_x\dot{x} + k_x x = me\omega^2\cos(\omega t) + \alpha_x Q(t) + \beta_x H(t) \quad (13)$$

and similarly for y, z, while $Q(t)$, and $H(t)$ are related through the pump characteristic and through the functions G.

2.2 Harmonic Analysis - Analytical Solution for a Stationary Forced Regime

If the excitation is harmonic (for example, imbalance at rotational frequency ω), and if the behavior can be considered linear in the vicinity of the operating point, then for one coordinate:

$$m\ddot{x} + c\dot{x} + kx = F_0\cos(\omega t) \tag{14}$$

The steady-state solution gives the amplitude:

$$X(\omega) = \frac{F_0}{m}\frac{1}{\sqrt{\left(\omega_n{}^2 - \omega^2\right)^2 + (2\zeta\omega_n\omega)^2}} \tag{15}$$

where $\omega_n = \sqrt{k/m}$ - the natural frequency, and $\zeta = c/(2m\omega_n)$ - the quality factor/damping.

Interpretation: at $\omega \to \omega_n$ (resonance) the amplitude grows; damping reduces the peak amplitude.

Relationship between vibration amplitude and changes in Q and H — models and solution examples

Practical engineering approach. Linear approximation

If experiments show that small changes in vibration lead to approximately linear changes in flow rate and head, one may assume:

$$Q(t) = Q_0 + \gamma_x x(t) + \gamma_y y(t) + \gamma_z z(t) \tag{16}$$

$$H(t) = H_0 + \eta_x x(t) + \eta_y y(t) + \eta_z z(t) \tag{17}$$

Here γ_i, η_i are empirical coefficients ($m^3/h\cdot mm^{-1}$ or $m\cdot mm^{-1}$), determined by regression from measurements.

Then the system (2) + (16), (17) can be solved numerically:

First $x(t)$, $y(t)$, and $z(t)$ must be found from (2) (knowing the Q, and H relationships), and then Q, and H are calculated using (16), and (17).

In the frequency domain, the amplitudes are substituted (16), and (17), and the change in the mean Q/H is obtained.

Nonlinear approximation for large amplitudes

If the effect is strong, a polynomial or logarithmic relationship is used:

$$Q = Q_0 + \sum_i \gamma_{1i}x_i + \sum_i \gamma_{2i}x_i^2 \ldots \tag{18}$$

or

$$H = H_0 \cdot \exp\left(-\sum_i k_i A_i\right) \tag{19}$$

where A_i is the vibration amplitude in direction i.

By solving the problem, we obtain the following analytical conclusions:

For the established harmonic response, using Eq. (15) for the vibration amplitudes and Eqs. (16) and (17), the change in the mean Q and H can be obtained:

For example:

$$\Delta Q = \gamma_x X + \gamma_y Y + \gamma_z Z \tag{20}$$

where X, Y, Z are the amplitudes of steady-state oscillations.

For the transient process in time: the solution of variant (2) with any $F(t)$ yields a transient part (terms $e^{-\zeta\omega nt}$) and a forced part. Numerically, it is solved by ODE-solvers (MATLAB, Python) and then substituted into (16), and (17).

For stability (resonance): when the mechanical natural frequency coincides with characteristic hydrodynamic frequencies, hydro-mechanical resonances are possible, in which even small excitation sharply reduces Q and H because of large vibration amplitudes. In the model this appears as a peak in $X(\omega)$ and then a significant ΔQ.

Let us consider a specific model example with imbalance and linear sensitivity.

For simplicity, let us write one coordinate in the transverse direction:

$$m\ddot{y} + c\dot{y} + ky = me\omega^2\cos(\omega t) + \beta Q(t) \tag{21}$$

$$Q(t) = Q_0 + \gamma_y y(t) \tag{22}$$

Substituting, we obtain a nonlinearly coupled ODE (linearized for small γ_y):

$$m\ddot{y} + c\dot{y} + (k - \beta\gamma_y)y = me\omega^2\cos(\omega t) + \beta Q_0 \tag{23}$$

It can be seen here that the hydraulic coupling changes the effective stiffness: $k_{\text{eff}} = k - \beta\gamma_y$. If $\beta\gamma_y$ is positive and close to k, the system may become less stable → vibration amplification and a drop in Q. The analytical solution is the amplitude according to (15) with k_{eff}.

The performed analyses show that, to solve the stated problem, it is advisable to use the following equations:

In mechanics: a second-order ODE (mass-damper - stiffness), Eqs. (7) – (9). Solution: analytical (harmonic) for a sinusoidal force (Eq. (15)) or numerical integration for a general $F(t)$.

In hydraulics: a static Q–H characteristic plus a small-signal linear model relating ΔQ and ΔH to vibration - Eqs. (3), (4) and approximations (16), (17).

Coupling (FSI): substitute the Q, H dependences into the equations of mechanics → obtain coupled equations; analyze them in the frequency domain (to determine amplitudes) and/or solve them numerically in time.

Eq. (15) together with the linear Eqs. (16), (17) provides a simple way to predict the drop in Q/H at a known vibration amplitude; for strong nonlinearities, polynomial (exponential approximations or CFD-based FSI should be used).

When designing dynamic pumps, the geometric parameters of the cross-sectional flow area of the impeller passage are calculated using the numerical value of the

pumped-medium flow rate without considering rotor rotational speed. The numerical value of the circumferential component of flow velocity is used only to calculate the blade installation angle, taking into account the recommended incidence angle. To minimize hydraulic losses, the principle of smooth variation of the translational velocity of the pumped medium is implemented, starting from the initial section of the pump inlet nozzle. The geometry of the inlet flow area across the blade width is calculated taking into account blockage caused by the actual thickness of the blade leading edge, its rounding radius, and the setting angle.

At present, particular attention is paid to current issues such as energy saving and increasing the efficiency of pumping equipment. Energy consumption and environmental safety are indicators of pump unit quality.

The study of the flow of the pumped medium in the pump flow passage using modern computational tools and methods of mathematical modeling has revealed its complex vortex nature [19-22].

Modeling of fluid flow in the blade system was carried out by numerically solving the Navier-Stokes differential equations [23].

To solve the stated problem, it is necessary to identify the energy source causing the vibration, which represents a process of energy dissipation into the surrounding space. As an example, dynamic phenomena in the flow passage of hydraulic machines related to hydrodynamic processes should be considered. The specific types of these machines capable of generating vibration processes with dissipation, that is, with loss of energy, are closely associated with hydraulic energy losses leading to a reduction in the head of pumping units.

Studies show that, in this field, the energy source of vibration of hydrodynamic origin in centrifugal pumps operating at reduced flow is the interaction of the transit fluid flow with the leading edge of the impeller blades, as well as the interaction of different directions of fluid motion within the interblade channels of impellers during hydrodynamic processes [24].

These factors are always present in operating pumps. They are expressed to different degrees depending on the situation. The degree of their manifestation depends on the pump operating mode.

Under the design operating mode of a pump, the fluid flow enters the impeller tangentially to the leading edge of the blades. Achieving this condition is ensured by the appropriate design calculation and structural implementation of this element of the pump flow passage. If the operating mode deviates from the design one, the initial flow velocity field becomes distorted and a mismatch arises between the direction of the relative flow velocity at the impeller inlet tangent to the leading edge (v_1) and the inclination angle of the leading edge α_1 as defined by angle $\alpha_{1\pi}$ (Fig. 1).

As a result, an angle of attack of the flow on the blade is formed, leading not merely to tangential contact between the flow and the blade, but to a force interaction between them. As the pump operating condition (Q), moves away from the design condition (Q_0), the angle of attack i increases and the force interaction between the flow and the blade becomes stronger. With small deviations of the pump operating condition (Q) from the design condition (Q_0), within the pump operating zone, the interaction is slight, but it increases substantially as Q departs from Q_0.

After losing contact with the blade leading edge, the fluid flow enters the interblade channels of the rotating impeller. In this case, the liquid filling the channels is simultaneously forced to exist in several types of motion.

Most of these motions are derivative in nature. Initially, two types of motion are dominant: inertial motion of the liquid in the direction opposite to the impeller rotation, and transit motion under the action of centrifugal force generated by the energy of the pump drive motor.

The forces of these two types of motion, applied to the same mass of liquid, impart a certain resultant motion to the flow. A liquid mass is formed in which regions of disturbance appear and disappear, moving from the inlet of the impeller channels to their outlet. The resulting disturbances impart pulsations to the flow and include unsteady phenomena caused by cavitation processes triggered by separated flow around the confining surfaces, as well as other factors that together create a certain vibroacoustic picture in the pump flow passage. During pump operation within the working zone, the power characteristics are low and can create only a general vibroacoustic background that does not cause significant damage to the equipment. When the pump operating mode leaves the working zone, the situation changes significantly. The vibrations caused by the types of fluid motion considered increase with increasing deviation of the actual mode Q from the design mode Q_0.

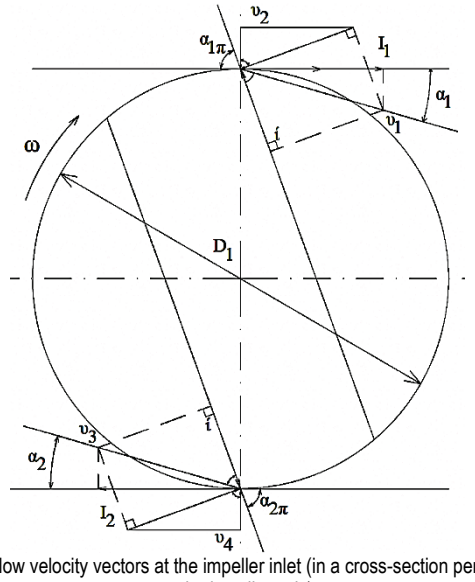


Figure 1 Flow velocity vectors at the impeller inlet (in a cross-section perpendicular to the impeller axis)

Let us consider the interaction between transit and inertial flows using a simplified diagram of fluid motion in the interblade channel of the impeller (Fig. 2).

The liquid located in the interblade channel between blades 1, when the impeller rotates at angular velocity, tends to preserve its state of rest and therefore rotates in the direction opposite to the wheel rotation. The volumes of

liquid located at the wheel outlet are at a greater distance r_2 from the center of rotation than the volumes located near the wheel inlet opening at distance r_1.

The difference in distances leads to the fact that the liquid acquires a higher inertial motion velocity I_2 compared with the liquid at distance r_1, which has velocity I_1.

The difference between I_2 and I_1 causes the liquid filling the interblade channel to rotate relative to the wheel with velocity V_u, directed toward the transit flow entering the interblade channel with velocity V_T after its reflection from the blade leading edge. When the transit and inertial flows meet near the working surface of the blade, their collision leads to force interaction between these flows, which creates the pulsations noted earlier.

When calculating and designing pumps, the interblade channels of the impellers are given geometric shapes intended to ensure favorable interaction between the inertial and transit liquid flows. This can be achieved only within a rather narrow range of flow rates Q. Since the geometric characteristics of the interblade channels become fixed after manufacture, the vector V_u, whose spatial orientation is largely determined by these characteristics, remains almost unchanged and constant. At the same time, the velocity field of the transit flow and, consequently, the orientation of vector V_T changes as the pump operating mode changes. The greater the deviation of Q from Q_0, the greater the mismatch between the transit and inertial flows.

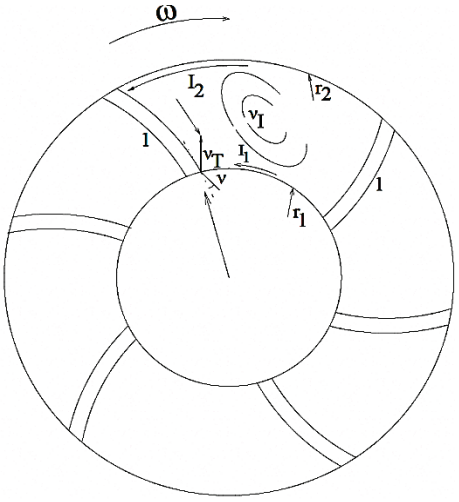


Figure 2 Schematic diagram of the interaction between the transit and inertial flows in the interblade passage of a centrifugal pump impeller

Centrifugal pumps are continuous-flow hydraulic machines operating not periodically, but continuously. The above processes of interaction between the transit flow and the leading edge of the impeller blade and with the inertial flow are continuous. Therefore, they are not accompanied by sufficiently pronounced dynamic impact phenomena that by their nature could cause oscillatory and vibration processes.

According to the studies carried out within this work, the generator of vibration of hydrodynamic origin in the centrifugal pumps under consideration is the tongue of the pump inlet spiral casing (Fig. 3).

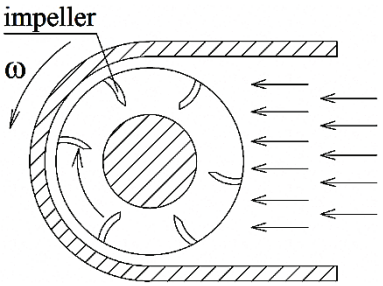


Figure 3 Schematic diagram of fluid flow in the inlet passage of a centrifugal pump with a centrally positioned impeller

The tongue is installed in the inlet spiral casing of the pump directly before the centrifugal wheel inlet to ensure uniform admission of liquid into the interblade channels of the impeller. It fulfills its function by preventing direct entry of liquid into the channels located closer to the pump inlet nozzle, forcing the flow, while moving along the spiral, to fill all the remaining wheel channels sequentially. This provides hydrodynamic balancing of the machine rotor.

During impeller rotation, each channel is alternately blocked by the tongue, which prevents liquid from entering the blocked channel and interrupts the contact of the transit liquid flow entering it with the leading edge of the corresponding impeller blade and with the inertial flow in the blocked interblade channel. After the blocked channel exits the hydrodynamic shadow of the tongue as the wheel continues to rotate, liquid admission into the channel resumes and the transit flow, entering the reopened channel, comes into contact with the blade leading edge and with the inertial flow. In this case, the renewed contact is dynamic and impact in nature. The force of impact between the transit flow and the obstacles in front of it, in the form of the leading edge and the inertial flow, becomes greater as the mismatch between the velocity fields of these flows and the velocity field of the transit flow with the geometric characteristics of the blade leading edge increases. As noted above, this mismatch is almost zero at $Q = Q_0$ and is quite low when Q is within the pump operating zone, but it increases significantly as Q moves away from the boundary of the machine operating zone.

When the transit flow comes into force contact with the leading edge of the impeller blade, the reaction force acts on the blade and, through it, on the impeller and on the pump rotor as a whole. The resulting force is directed normal to the machine rotor axis if the blades are not inclined relative to the impeller discs. When the blades are inclined relative to the wheel discs, impact with the blade generates transverse and axial forces acting on the pump rotor.

Impact between the transit and inertial flows occurs in a liquid medium. The force generated in this case is perceived by the blades and, in addition, by the impeller discs. As a result, the generated transverse and axial forces ultimately act on the pump rotor.

Under the action of these forces, which are periodic in nature, the pump rotor undergoes oscillatory motion. As a result, oscillatory motion is produced throughout the hydraulic machine structure, and the pumping unit enters a vibration regime.

Impact of the transit flow with the leading edges of the impeller blades and with inertial flows in the interblade channels occurs each time access is opened for the transit flow into the next channel opening behind the tongue. This phenomenon occurs at a frequency equal to the number of channels, and therefore blades, in the impeller. If the wheels have a double-sided inlet and matched blades in both halves, then it occurs at a frequency equal to the number of blades in one half of the wheel.

Thus, a physical model of vibration has been developed. According to the proposed physical model of vibration, when centrifugal pumps operate under reduced flow, vibration should be generated in them with frequency φ_Z, for centrifugal pumps operating under reduced flow, equal to the rotational frequency of the machine rotor

$$\varphi_{\text{rot}} = \frac{n}{60} \text{ (rpm)} \tag{24}$$

multiplied by the number of blades in the impeller z, or at the blade-passing frequency

$$\varphi_{\text{rot}} = \frac{n}{60} \cdot z \tag{25}$$

where n – is the pump rotor rotational speed, (1/min).

The angle of inclination of the blades relative to the impeller discs φ may be 0° or, in the case of blades inclined to the discs, of the corresponding order 10° - 20°. With such a wheel design, the force vector arising from interaction of the transit flow with the impeller blades (Fig. 4).

Flow is directed either normal to the pump rotor axis or at some small angle to it. This creates vibration. In the first case, only in the transverse direction; in the second, in both the transverse and axial directions, with the transverse direction predominating.

The main and covering, or front, impeller discs are either equal in diameter or differ only slightly, except in special pump impeller designs. The inner-facing surfaces of these discs likewise have either equal areas or areas that differ only slightly. These surfaces perceive all hydrodynamic consequences of the impact contact of the transit flow with the blade leading edges and with the inertial flow. The consequences of these contacts are pressure pulsations acting on nearly equal-area disc surfaces facing in opposite directions. The resulting forces applied to the discs are nearly equal, act in opposite directions, and mutually compensate each other.

Therefore, in centrifugal pumps operating under reduced flow, according to the proposed model, hydrodynamic vibration should be observed predominantly in the transverse direction; the axial component should also be recorded, but on a smaller scale, mainly in pumps having wheels with blades inclined relative to their discs.

Vibration also occurs in the vertical direction, i.e., a vertical component, because the vibrating massive pump rotor is supported by the machine bearing assemblies. The appearance of the vertical direction is also promoted by the action on the blade that has emerged from the hydrodynamic shadow of the tongue of the pump inlet spiral casing of the above-mentioned pressure pulsations in the corresponding interblade channel of the wheel.

This action is represented by vector v_2 (see Fig. 1).

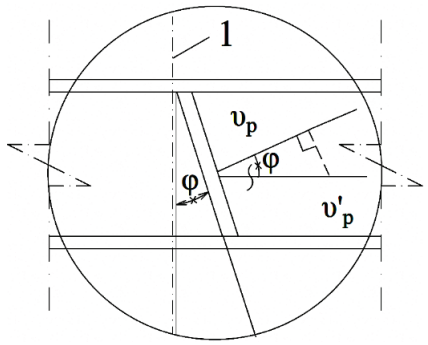


Figure 4 Schematic diagram of the blade positioned between the impeller disks: 1 – direction of the rotation axis of the pump rotor (impeller)

The results of numerous experimental data obtained in vibration tests of pump units confirm the conclusions made above regarding the expected values of vibration parameters in the case considered. As an example, some data from vibration tests of NM-type trunk oil pipeline pumps with a nominal flow rate of 12500 m^3/h. are presented. These pumps have impellers with blades inclined relative to the wheel discs. The number of blades is $Z = 7$. The rotor speed of the pumps is $n = 3000$ rpm.

Table 1 Results of pump vibration tests

Number of the pump	Pump operating mode				Vibration in different directions, 10^{-3} m/s			Vibration frequency, Hz
	Pressure		Head, m	Flow rate Q, m^3/h	Vertical	Transverse	Axial	
	Inlet, MPa	Outlet, MPa						
1	0.98	3.12	258	9700	4.12	13.5	12.7	350
2	0.98	3.14	260	9600	4.2	12.2	11.3	350
3	0.98	3.16	263	9500	4.31	11.6	10.1	350
4	0.98	3.18	265	9250	4.4	10.45	9.18	350
5	0.98	3.20	266	9025	4.55	9.37	7.76	350
6	0.98	3.22	269	8880	4.8	8.3	6.12	350

Tab. 1 presents the results of pump vibration tests under different operating conditions. As the outlet pressure and head increase from 3.12 to 3.22 MPa and from 258 to 269 m, respectively, the flow rate gradually decreases from 9700 to 8880 m³/h. The vibration response is direction-dependent: vertical vibration increases slightly from 4.12 to 4.80 × 10^{-3} m/s, whereas transverse and axial vibrations decrease significantly, from 13.5 to 8.3 × 10^{-3} m/s and from 12.7 to 6.12 × 10^{-3} m/s, respectively. The vibration frequency remains constant at 350 Hz throughout all operating modes.

Based on the above table, we plot graphs of head and flow rate as functions of vibration in the Vertical, Transverse, and Axial directions:

Formula for the vertical line at the optimum: $x_0 = 4.12$ mm/s and $H_{\max} = 269$: $H = -30(x - 4.12)^2 + 269$

Formula for the transverse line at the optimum: $x_0 = 6.12$ mm/s and $H_{max} = 269$: $H = -0.4(x - 6.12)^2 + 269$

Formula for the axial line at the optimum: $x_0 = 8.3$ mm/s and $H_{max}=269$: $H = -0.3(x - 8.3)^2 + 269$

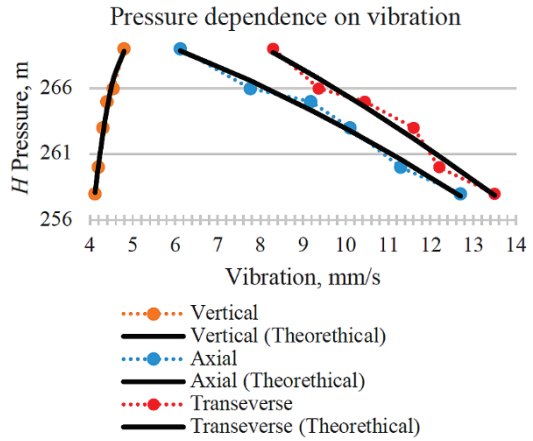


Figure 5 Dependence of pressure *H* on vibration

Formula for the vertical line at the optimum: $x_0 = 4.12$ mm/s and $Q_{max} = 9700$: $Q = -1800(x - 4.12)^2 + 9700$

Formula for the transverse line at the optimum: $x_0 = 6.12$ mm/s and $Q_{max} = 9700$: $Q = -35(x - 6.12)^2 + 9700$

Formula for the axial line at the optimum: $x_0 = 8.3$ mm/s and $Q_{max} = 9700$: $Q = -25(x - 8.3)^2 + 9700$

The plotted graphs show that in all three vibration directions (vertical, transverse, and axial) one main pattern is observed: the higher the head and flow rate, the higher the vibration in all directions. As the head and flow rate decrease, the vibration also decreases. This is due to the fact that, when the pump operates closer to the maximum of the characteristic curve, the following increase: hydrodynamic forces on the impeller and blades, turbulence, axial and radial loads, cavitation micropulsations and vibration of the shaft and supports.

Within the head interval of 258 - 269 m, vertical vibrations as a function of head increase almost linearly from 4.12 - 4.8 mm/s. This small increase indicates uniform distribution of axial forces and good rotor balancing.

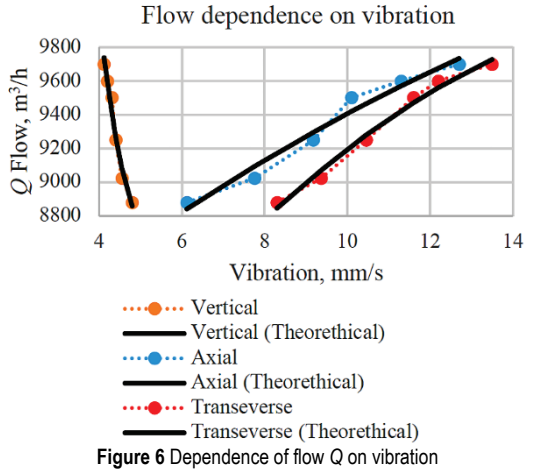


Figure 6 Dependence of flow Q on vibration

Thus, vertical vibration is the most stable and least sensitive to changes in operating conditions. In the transverse direction, vibration decreases from 13.5 - 8.3 mm/s. This is because, as head increases, the pump moves away from the zone of hydraulic instability; that is, in the region of maximum flow the flow becomes more stabilized and the lateral (radial) forces on the shaft decrease. Thus, transverse vibration is the most sensitive parameter indicating the hydraulic stability of the operating mode. In the axial direction, vibration decreases sharply from 12.7 - 6.12 mm/s. This is explained by the fact that, as head increases, the shaft operates closer to its optimum position, axial flow imbalance decreases, and pressure pulsation is reduced. Axial vibration decreases most strongly, which is a sign that the pump is operating closer to the optimum mode in terms of axial forces.

The graphs also show the dependence of vibration on head. In this relationship, the head varies within 258-269 m. The graph shows that vertical vibration grows almost linearly, from 4.12-4.8 mm/s. This slight increase means a uniform distribution of axial forces and good rotor balancing. Vertical vibration is the most stable and least sensitive to changes in operating conditions. Transverse vibration decreases from 13.5-8.3 mm/s. This is because, as head increases, the pump leaves the zone of hydraulic instability; that is, in the region of maximum flow the flow becomes more stabilized and the lateral (radial) forces on the shaft decrease. Transverse vibration is the most sensitive parameter indicating the hydraulic stability of the regime.

The dependence of vibration on flow rate shows that as the flow rate decreases from 9700 to 8880 m³/h, vertical vibration increases, while transverse and axial vibration decrease. The reason is high flow rate, low head, and unstable hydraulics. All this leads to increased turbulence and lateral forces, as well as to flow non-uniformity. Therefore, at high flow rates, transverse and axial vibrations are maximal. As the flow rate decreases, the pump operates closer to its optimum head regime, and therefore vibrations in these two directions fall. The interpretation by direction is given in Tab. 2.

Table 2 Interpretation by direction

Vibration direction	Trend	Reason
Vertical	Slightly increases with increasing head	Increase in the overall load on the bearing
Transverse	Noticeably decreases	The flow becomes more stable, resulting in lower radial forces
Axial	Decreases sharply	Optimization of axial forces and reduction of reverse thrust

Tab. 2 shows that vibration behavior varies depending on the direction of measurement as the pump head increases. Vertical vibration exhibits a slight increase due to the higher load acting on the bearings. In contrast, transverse vibration decreases noticeably because the flow becomes more stable, reducing radial forces. The most significant reduction is observed in the axial direction, which is associated with the optimization of axial forces and the mitigation of reverse thrust.

3 DISCUSSION

Pump operation at high head and lower flow rate reduces transverse and axial vibration and ensures a more stable flow. At high flow rate and low head, however, radial and axial pulsations increase, vibration grows, and stability decreases. NM-type oil pipeline pumps are significantly underloaded in terms of flow at a design capacity of 12500 m^3/h, they operate at 8600 - 9600 m^3/h, and at the same time experience increased vibration. The vibration is evidently hydrodynamic in nature. As noted earlier, the pumping units had no remarks concerning the mechanical part. The recorded vibration exceeds the normal level for this type of machine, characterized by a vibration velocity of 7.1×10^{-3} m/s and occurs at a frequency equal to the blade-passing frequency of the pumps. The same vibration frequency is obtained by calculating it according to the model under consideration, using Eq. (25):

$$\varphi_z = \frac{n}{60}\cdot z = \frac{3000}{60}\cdot 7 = 350 \text{ , Hz} \tag{26}$$

The vibration levels obtained in the three directions show that oscillations predominate in the transverse and axial directions, with the transverse direction prevailing. Vibration is also recorded in the vertical direction, but to a lesser extent compared with the transverse and axial directions.

The results obtained reflect the hydrodynamic processes occurring in pumps with specific design features, containing before the impeller inlet a structural element that fully or partially blocks the entrance to the interblade channels of the wheel. The tongue of the pump inlet spiral casing is characteristic of pumps equipped with double-entry impellers and with the impeller located in the middle of the shaft of hydrodynamic machines; this is one of the most common types of centrifugal pumps. These include main oil pipeline pumps (NM).

A similar element is present in some overhung-type pumps. It is made in the form of a rib along the inner generatrix of the pump inlet nozzle, which simultaneously performs the function of the machine inlet. Its purpose is to prevent or reduce flow swirl before the flow enters the impeller.

4 CONCLUSIONS

1) Testing confirms the expediency of changing the traditional approach to designing the flow passage of the pump impeller and significantly reducing the flow area of its passage section.
2) Increasing the efficiency of a dynamic pump is achieved by minimizing hydraulic losses through optimization of the impeller flow passage.
3) At reduced flow and underloading of centrifugal pumps, increased vibration of hydrodynamic origin arises at a frequency equal to the blade-passing frequency and is recorded in the transverse, axial, and vertical directions. In this case, the predominant oscillations occur in the transverse direction.
4) The presented numerical model adequately reflects the phenomenon under study and explains the nature of the occurrence of the vibration type considered, reveals its specific energy sources and the factors generating the oscillatory motion itself, and explains its distribution along the spatial coordinates.
5) To develop practical recommendations for optimizing the geometry of the impeller flow passage, additional computational and experimental studies are required.
6) Hydrodynamics can modify natural frequencies and damping, creating a resonance risk.
7) The main reason for the strong sensitivity of Q/H to vibration is the presence of resonances and nonzero hydraulic sensitivity to displacements.
8) If the natural frequencies of the mechanics intersect with the operating frequencies, it is necessary to change stiffness/damping or choose another operating mode (for example, according to Pontryagin).

Acronyms

CFD	Computational Fluid Dynamics
FSI	Fluid-Structure Interaction
ODE	Ordinary Differential Equation
ESP	Electric Submersible Pump
NM	Type of mainline oil pipeline pumps
MPC	Model Predictive Control

Acknowledgements

The author(s) special thanks go to the staff of the Scientific-Research Institute "Geotechnological Problems of Oil, Gas and Chemistry", Baku, Azerbaijan for providing useful suggestions.

5 REFERENCES

[1] Chugaev, R. R. (2014). *Hydraulics, hydraulic machines and hydraulic drives*. Mashinostroenie.
[2] Stepanov, A. J. (1993). *Centrifugal and axial flow pumps: Theory, design and application*. Wiley.
[3] Gülich, J. F. (2010). *Centrifugal pumps*. Springer. https://doi.org/10.1007/978-3-642-12824-0
[4] Lebedev, A. A. (2015). *Vibrations in engineering and technologies*. Mashinostroenie.
[5] Ehrich, F. (1999). *Handbook of rotordynamics*. McGraw-Hill.
[6] Rao, J. S. (2011). *Rotor dynamics*. New Age International. https://doi.org/10.1007/978-94-007-1165-5_14
[7] Tymchak, V. S. (2008). *Dynamics and vibrations of rotor machines*. Tekhnika.
[8] Tokarev, M. P. (1983). *Transient processes in hydraulic systems*. Mashinostroenie.
[9] Babakov, A. M. (2000). *Hydraulics of oil pumping systems*. Nedra.
[10] Zheverev, S. V. (2012). *Theory of centrifugal pumps*. Politekhnika.
[11] Monin, A. S., & Yaglom, A. M. (2000). *Statistical hydromechanics*. Nauka.

[12] Nayfeh, A. H., & Mook, D. T. (1995). *Nonlinear oscillations*. Wiley. https://doi.org/10.1002/9783527617586
[13] Basarygin, Yu. M., & Vlasov, A. L. (2015). *Dynamics of ESP operation*.
[14] American Petroleum Institute. (2020). *API RP 11S5: Recommended practice for ESP system vibration*. API.
[15] Martyanov, A. V. (2017). *ESP: Theory and practice of operation*. TyumGNIIPI.
[16] Schlumberger. (2016). *ESP design manual*. Schlumberger.
[17] Baker Hughes. (2018). *ESP system reliability and vibration analysis*. BHGE.
[18] Ogata, K. (2010). *Modern control engineering*. Prentice Hall.
[19] Kochevsky, A. N., & Nenya, V. G. (2014). *Modern approach to modeling and calculation of fluid flow in blade hydraulic machines*.
[20] Sokolov, E. V., & Ankudinov, D. T. (2007). Numerical simulation of working processes in centrifugal pumps as a method for improving their efficiency. *Lesnoy Vestnik, 11*(8), 165–169.
[21] Slobodkina, F. A., Malinin, A. V., Malinin, V. V., & Shigapova, D. Yu. (2008). Fluid flow in a centrifugal pump stage. *Mathematical Modeling, 20*(10), 51–62.
[22] Borovsky, B. I. (1989). *Energy parameters and characteristics of high-speed blade pumps*. Mashinostroenie.
[23] Anderson, D., Tannehill, J., & Pletcher, R. (1990). *Computational fluid dynamics and heat transfer*. Mir.
[24] Sokolov, E. V., & Ankudinov, D. T. (2006, September 20–21). Modeling of fluid flow in centrifugal pumps. In *Proceedings of the Eurasian Symposium "Wood Processing: Technologies, Equipment, Management of the XXI Century"* (pp. 217–221). Ural State Forest Engineering University.

Authors' contacts:

Amir G. Mustafayev, Associate Professor
Faculty of Mechanical Engineering,
Azerbaijan State Oil and Industry University,
Azadliq avenue 43, AZ 1118 Baku, Azerbaijan
mustafaev-1959@mail.ru

Sanam J. Muslimov, PhD Researcher
(Corresponding author)
Department of Industrial Machines,
Scientific-Research Institute "Geotechnological Problems of Oil, Gas and Chemistry",
Azadliq avenue 43, AZ 1118 Baku, Azerbaijan
sanammuslimov@gmail.com

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250502102118

Preliminary communication
Received: 2025-05-02, Accepted: 2025-05-12

Industry 4.0 Readiness in Electrical Companies: Analysing Product Manufacturing and Communication in the Customer and Supplier Chain

Andrea Benešová*, Martin Hirman, František Steiner, Jiří Tupa

Abstract: This study examines the readiness of electrical companies to adopt Industry 4.0 principles, focusing on the degree of product customization and communication practices within customer and supply chains. Using a questionnaire survey via GoogleForms, data was collected from 28 companies in the electrical industry. The results showed that 63% of companies offer full product customization, indicating advanced technology integration, while 23% of companies only allow partial customization, indicating a transitional phase towards Industry 4.0 adoption. However, 14% stick to fixed product portfolios, reflecting limited flexibility and readiness for digital transformation. The communication analysis further revealed that most companies still rely on traditional methods such as electronic and telephone communication to interact with suppliers and customers, with only 10% achieving fully automated digital integration. This study highlights significant gaps in digitalisation and provides useful insights in terms of technology adoption and communication strategies in line with Industry 4.0 principles.

Keywords: communication; customer; Industry 4.0; product; readiness; supplier chain

1 INTRODUCTION

The Industry 4.0 concept is a revolutionary approach to manufacturing processes that combines digital technology, automation and intelligent systems to achieve greater flexibility, efficiency and personalization. A key aspect of this revolution is the individualisation of products and production, which is closely linked to communication between customers, suppliers and society. Industry 4.0 leverages technologies such as the Internet of Things (IoT), artificial intelligence (AI) and machine learning, cloud computing and big data to enable efficient data collection and analysis, quality management automation and predictive maintenance. These technologies enable efficient processing and analysis of large volumes of data, leading to a better understanding of not only manufacturing processes, but can be used to identify customer preferences. [1] In fact, digital platforms and technologies enable direct communication with customers, leading precisely to a better understanding of their needs and preferences. In addition, the technologies allow for better integration of suppliers into the production process, leading to faster and more efficient product delivery. However, the implementation of these technologies is associated with significant economic costs and, at the same time, the increasing volume of production data increases the need to ensure adequate security measures to protect this information from unauthorised access. Although the transition to Industry 4.0 requires significant investment in new technologies and employee training, it offers companies the opportunity to become more competitive, flexible and customer-oriented. [2]

For these reasons, the aim of the publication was to determine the readiness of enterprises for this concept in terms of the degree of individualisation of products and production and the way in which enterprises communicate with customers and suppliers. Electrical companies operating in the Czech Republic were selected for the research for several reasons. These enterprises are often at the cutting edge of innovation as they are closely linked to the development and application of advanced technologies. In addition, they depend on precise and fast production processes and can therefore benefit most from the innovations associated with Industry 4.0. The electronics sector also faces increasing competition and changing customer demands, which requires the adaptability and flexibility that Industry 4.0 can provide. In addition, businesses in this sector must address safety and sustainability issues, which are also key aspects in implementing the concept. So by examining the readiness of these businesses for Industry 4.0, their needs can be better understood and areas for improvement identified. In contrast to other studies already conducted, this publication offers unique empirical data obtained directly from companies operating in the electrical industry in the Czech Republic. These data provide valuable insights into the manifestation of Industry 4.0 in this specific sector and region, allowing for a deeper understanding of its impacts and applications in the given context. In this context, the following research questions were set:

- Question 1: What is the extent of product and process individualization in electronics manufacturing companies?
- Question 2: How does communication occur within the supplier-customer chain in electrical companies?

2 STATE OF THE ART

The transformation of industrial production in three distinct revolutionary phases has not only fundamentally changed the technological capabilities of production, but has also redefined customer expectations and communication paradigms in the supply and demand chain. This evolution reflects the complex interaction between technological innovation, socio-economic factors and the evolution of market mechanisms, with each industrial epoch bringing

about a specific transformation of the producer-customer relationship; industrial production is currently facing another revolutionary phase that will bring further changes in these areas, as summarised in Tab. 1 below.

Table 1 Comparison of industrial revolutions in terms of communication with customers and suppliers

	Industry 1.0 [3]	Industry 2.0 [4]	Industry 3.0 [5]	Industry 4.0 [6]
Main difference	Craft excellence a local communication networks	Mass production and standardization of communication	Computerisation of production and segmented individualization	Hyper-personalization and integrated communication
Production	Dominated by craft production supplemented by the beginnings of mechanisation	Electrification of production and assembly lines	Computer-controlled manufacturing systems transformed production processes	Characterised by cyber-physical systems and artificial intelligence
Customer expectations	Functionality and durability of products, with personalization being realised through direct interaction with the craftsman	Shifted from artisanal uniqueness to a prioritisation of affordability and product reliability, with personalization taking a back seat	A significant transformation - from acceptance of uniform products to demands for differentiation and market segmentation. Customers have increased their focus on quality and have started to prefer more customised products	Personalised product - manufacturers will be able to anticipate market needs through predictive analysis
Changes in the supply chain	Personal contact or correspondence, with delivery times determined by the physical constraints of transport	The expansion of telegraph and telephone communication accelerated information flows in the supply chain and laid the foundations for global business networks	Communication in the supply chain has undergone a fundamental transformation with the implementation of electronic data exchange, which has eliminated the dependence on paper documentation and manual order processing.	Integrated digital platforms should facilitate an unprecedented level of collaboration between manufacturers and their suppliers through real-time data sharing.

As the table shows, the new revolutionary phase, often referred to as Industry 4.0, is characterised not only by digitisation and automation of production, but above all by a significant shift in customer expectations. The emphasis is shifting to the individualisation of products and services, as confirmed by recent surveys. In fact, McKinsey & Company research from 2021 shows that 71% of customers expect a personalised experience, while 76% express frustration when their individual preferences are not reflected. [7] Deloitte statistics show that 36% of customers are even willing to provide personal information in exchange for a personalised product, while 22% accept a higher price for a personalised solution. [8] However, when it comes to communication and ensuring a stable supply chain, the experience of the Covid-19 pandemic has necessitated the adaptation of these systems to exhibit sufficient levels of stability and resilience to global disruptions. A 2022 World Economic Forum study identifies a critical need to implement hybrid supply chain models that balance the efficiency of Just-in-Time principles with the robustness of more traditional, higher-volume approaches. An analysis of supply chain disruptions during the pandemic showed that companies with advanced digitalization of SCM processes were able to adapt their supply structures to changing conditions faster than companies with lower levels of digitalization. [9] Current trends are therefore towards the implementation of predictive risk models that continuously assess potential vulnerabilities in supply chains and automatically suggest alternative scenarios when imminent disruptions are identified.

In terms of companies' readiness to implement changes not only in production but also in the supply chain or communication with customers related to Industry 4.0, a number of models have been developed since 2015, which differ in their complexity and focus. Some models comprehensively assess the entire enterprise, while others assess only a specific area. The most famous models include RAMI, IMPULS or SIMMI 4.0. In terms of the current state of the art, for example, the GUVEI-Model was introduced in 2022, integrating improvement strategies and Industry 4.0 technologies to optimize key performance indicators (KPIs) and support the long-term sustainability of manufacturing organizations. [10] In addition, the Industry 4.0 Readiness Model, this model includes 10 dimensions and 37 indicators that allow companies to more accurately assess their readiness to implement Industry 4.0 technologies. [11]

Another a comprehensive view of companies' readiness in the area of supply chain digitalization is provided by the "Digital Supply Chain Readiness" diagnostic tool developed by EDGE Digital Manufacturing, Digital Catapult and AI Idea Factory, which is based on the principles of the British Standard PAS 1040 for the adoption of digital technologies in manufacturing. [12] In line with this approach, a recent 2023 study revealed that leadership, processes and culture are key factors for successful digital transformation. [13] Further research from 2024 highlights that digital maturity is essential for optimising production processes and increasing competitiveness. [14] This claim is confirmed by a structured survey [15] conducted among 364 supply chain professionals. The data collected was analysed using structural equation modelling using the Partial Least Squares Method (PLS-SEM). The empirical findings confirm and highlight the importance of real-time digital integration, continuous innovation and improved visibility in leveraging I 4.0 for competitive advantage. Further research [16] has focused on identifying nuanced trends that contribute to supply chain efficiency, waste reduction and environmental footprint. The findings highlight key insights: the adoption of management practices, particularly total quality management

and just-in-time, contribute significantly to reducing waste, increasing efficiency and minimizing the environmental footprint in supply chains. At the same time, the integration of Industry 4.0 technologies, such as blockchain, IoT, big data, strengthens data-driven decision-making, transparency and traceability, reinforcing sustainability efforts. Another, empirical research shows that firms implementing advanced supply chain simulation models have higher success rates in implementing individualized manufacturing strategies. [17]

There has been a significant shift in product customization and customer communication in recent years, mainly due to advances in technologies such as artificial intelligence (AI) and big data. The authors of publication [18] highlight the role of these technologies in effectively tailoring products to individual customer needs. Their study also recommends comprehensive models including internal and external integration, technology readiness and statistical validation for successful implementation of product personalization and improved customer experience. In addition, the 2024 3D EXPERIENCE Platform has already been created to enable manufacturers to transform data from connected Industry 4.0 ecosystems into actionable insights for successful mass customization. The platform provides a digital thread that integrates people and information across the organization and value chain, enabling continuous improvement of products and operations. [19] In further research [20], the authors believe that current models do not place enough emphasis on avoiding the development of redundant solutions and focusing on customer needs. Therefore, the presented approach enables organisations to address customer requirements with digital solutions without creating redundancy. In another publication [21], the authors propose a methodology that could help create new software services in line with Industry 4.0 that will change the way customers interact with businesses and the experience they have in doing so. The methodology details a set of phases that are divided into stages, which in turn consist of activities. It places particular emphasis on the integration of semantic descriptions and 3D visualisation into the implementation of these new services.

3 METHODOLOGY

To investigate the degree of product customization and to analyse the communication tools used to communicate with suppliers and customers, the research used the methodological framework shown in the figure below.

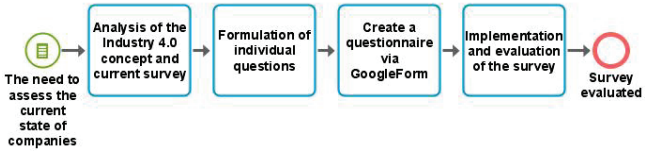


Figure 1 Description of the methodological framework

For the purpose of our investigation, a questionnaire was created using the Google Forms platform. This option was chosen because of its efficiency and speed in reaching a large number of respondents. Google Forms allows for easy distribution of the questionnaire through a shared link that can be effectively disseminated through emails, social networks and other communication channels. In this way, a wide range of respondents can be reached in a minimal timeframe. Another significant advantage of Google Forms is the ability to monitor the current status of questionnaire completion in real time. This feature provides an overview of the number of respondents who have already completed the questionnaire and the time at which completion occurred. Example questions from the survey are shown in the following figure.

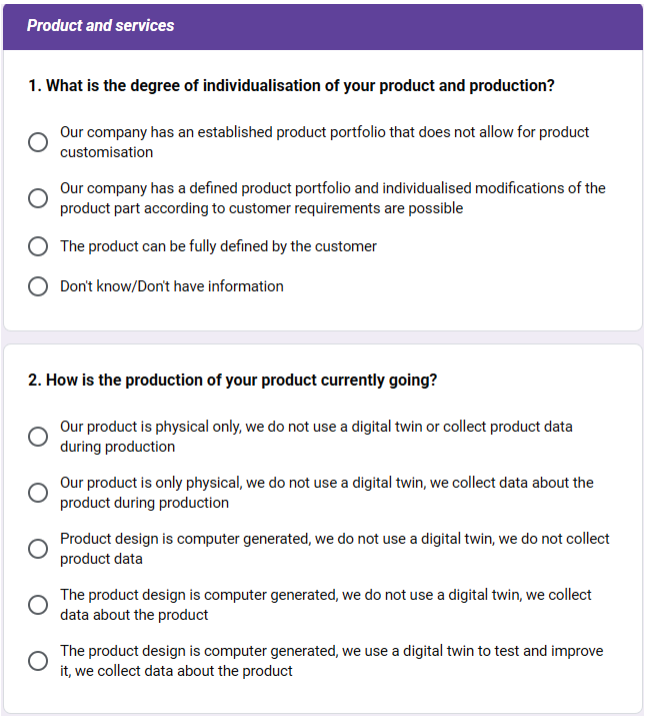


Figure 2 Example of questionnaire

4 RESULTS AND DISCUSSIONS

For the purpose of our empirical investigation, a questionnaire was developed and sent by email to approximately 150 companies specialising in the electrical industry. This selection of respondents was chosen because of the relevance of their expertise and experience in the field of Industry 4.0. In the end, 55 respondents completed the questionnaire on behalf of the companies concerned, but 5 companies were excluded because their main focus did not match the criteria for the electrical industry set for our survey and a further 22 companies were excluded because they answered that they did not have sufficient experience and knowledge in Industry 4.0. This left 28 valid respondents who met the established criteria. Of these respondents, 8% were micro enterprises, 22% were small enterprises, 28% were medium enterprises and 42% were large enterprises. This diversity of respondents by enterprise size provides valuable information on the structure of the electrical industry and allows for a deeper analysis of the responses in the context of enterprise size. This provides a better understanding of the needs and trends in the sector and a comprehensive view of the characteristics and strategies of enterprises in the electrical sector. The following chart

provides an overview of the ability of companies to adapt their products to customer requirements, a key aspect of the modern industrial environment, especially in the context of Industry 4.0.

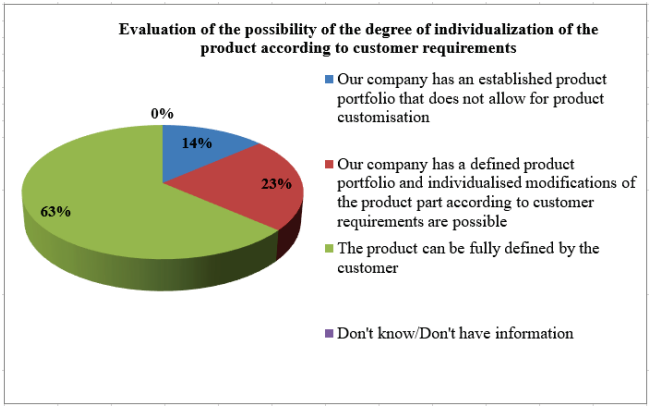


Figure 3 Evaluation of the degree of individualization of the product

According to the chart, 63% of respondents said their products can be fully defined by the customer. This result suggests that most companies have already achieved a high level of flexibility in production, which is in line with the principles of Industry 4.0. The ability to fully tailor products to individual customer needs often requires the implementation of advanced technologies such as digital twins, modular manufacturing systems or sophisticated product design software tools. These companies are likely to be well prepared for the challenges of digitalization and automation of production. A further 23% of enterprises reported that they have a defined product portfolio but allow customization of certain parts of their products according to customer requirements. This approach represents a compromise between standardisation and flexibility. In terms of Industry 4.0, this group of companies can be considered as a transitional phase on the way to the full integration of its principles. They are likely to gradually introduce technologies and processes that allow a greater degree of customization, but still face certain technological or organisational constraints.

Only 14% of respondents said that their company has a fixed product portfolio that does not allow for any form of customization. This group of companies is the least prepared in terms of Industry 4.0 for market changes requiring flexibility and personalization. The absence of the possibility of customization suggests a high degree of standardisation of production, which may be effective in traditional industrial models but less competitive in an environment requiring a rapid response to changing customer needs. An interesting finding is the fact that none of the respondents chose the "Don't know/Don't have information" option. This indicates a high level of awareness among respondents of their companies' capabilities in product customization, which is a positive signal for the implementation of Industry 4.0 principles. The results of the survey, shown in the chart below, provide an overview of the current state of how companies are using modern technologies such as digital twins, computer generated design and in-process data collection. Again, the results show the varying levels of implementation of these elements across industry.

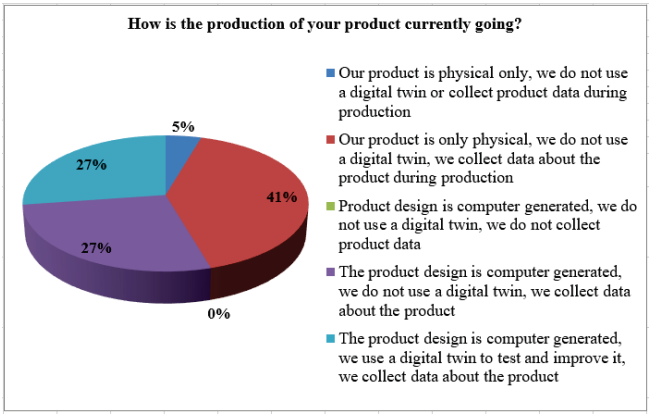


Figure 4 Evaluation of current production

In terms of implementing Industry 4.0 principles, it is clear from the results of the survey that most companies are still at a lower level of digitalization. In fact, the largest share is made up of enterprises with physical production without any data collection (41%), which indicates significant reserves in the area of innovation and technological development. Enterprises that already collect data or use computer generated design (54% in total) show some progress towards digitalization. However, only a small proportion (5%) have achieved full integration of technology through digital twins.

Effective communication with suppliers is a key factor in ensuring smooth production, optimizing the supply chain and minimizing risks associated with material shortages or delivery delays. In the context of Industry 4.0, it is important to analyse how companies use modern technologies to automate and digitise processes, which can significantly increase their competitiveness. Finding out the current state of communication between businesses and suppliers thus provides valuable information on the readiness of businesses for digital transformation and their ability to adapt to new trends.

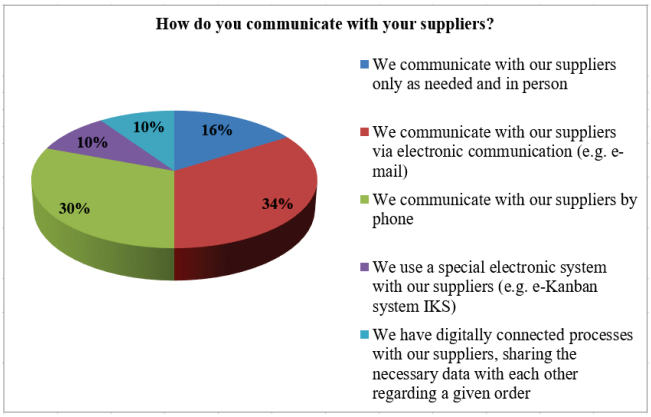


Figure 5 Evaluation of the communication between companies and suppliers

From the results it was found that most companies use a combined method of communication with suppliers, mainly

electronic and telephone communication. In terms of actual shares for each communication option, the largest share (34%) is accounted for by enterprises that communicate with suppliers via electronic communication such as e-mail. While this method is simple, fast and widely available, it often lacks a higher degree of automation and data integration. In second place (30%) are businesses that prefer to communicate by telephone. Telephony allows for quick resolution of urgent situations, but can be time consuming and less efficient when sharing complex data or documents. Only 16% of businesses use face-to-face communication on an 'as needed' basis. This method is suitable for specific situations, such as dealing with strategic issues or building long-term relationships, but is not practical for day-to-day operations.

In addition, 10 % of businesses use specialised electronic systems such as the IKS e-Kanban system. These systems enable automated inventory and order management, which significantly increases process efficiency and reduces the risk of human error. The smallest share (10%) is made up of enterprises that have fully digitally connected processes with suppliers. This approach represents the most advanced level of integration in line with Industry 4.0 principles, where data is shared automatically and in real time between all relevant parties. The way businesses and customers communicate plays a key role in the efficiency of production processes, customer satisfaction and building long-term relationships. In the context of modern manufacturing, especially in the context of the Industry 4.0 concept, it is important not only to ensure direct communication, but also to integrate digital tools that enable automation, personalization and real-time monitoring of processes. For this reason, the next section focused on exploring how companies communicate with their customers.

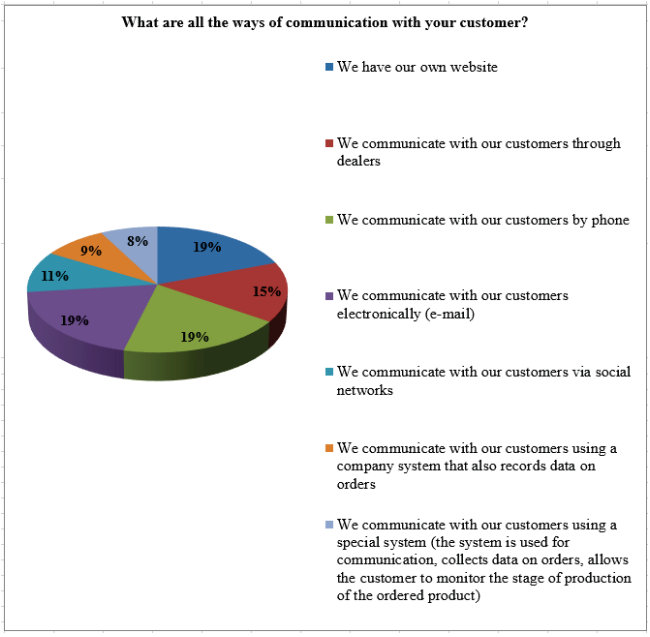


Figure 6 Evaluation of communication between companies and customers

Respondents were able to choose more than one method of communication, most often a combination of telephone and electronic communication. However, the individual results show that the most frequently used methods are the own website (19%), telephone communication (19%) and electronic communication via e-mail (19%). These methods represent the basic channels for direct interaction with the customer, allowing for a quick exchange of information and flexibility. Another important form of communication is the use of social networks (11%), indicating a growing trend towards digitalization and the use of modern technologies for marketing and interaction with clients. Only 9% of businesses use internal company systems that record order data, which can be considered insufficient in terms of digital integration. The lowest share (8%) is for communication through dedicated systems that allow customers to monitor the production status of their orders.

5 CONCLUSION

This study aimed to analyse the readiness of electrical companies for Industry 4.0 by investigating product customization capabilities and communication practices within customer and supplier chains. The research utilized a structured questionnaire distributed via Google Forms to gather data from a diverse sample of electrical enterprises. The findings provide valuable insights into the technological adoption, operational flexibility, and digital communication strategies employed by these companies, highlighting their current position in the journey toward Industry 4.0 integration. The analysis revealed significant variability in the ability of electrical companies to adapt their production and products to customer requirements, a core principle of Industry 4.0. Approximately 63% of respondents demonstrated full customization capabilities, indicating advanced technological integration such as modular manufacturing systems and digital twins. These companies are well-positioned to leverage Industry 4.0 technologies for enhanced competitiveness and responsiveness to market demands. In terms of digitalization levels, the study found that most companies remain at lower stages of technological integration. While 54% collect data or utilize computer-generated design tools, only 5% have implemented advanced solutions such as digital twins—a critical technology for optimizing production processes and bridging physical and virtual systems. This indicates substantial room for improvement in adopting Industry 4.0 technologies across the sector. Communication practices between companies and suppliers and customers still rely mostly on traditional methods such as email and telephone. Only a minority have achieved fully automated digital communication systems that are in line with Industry 4.0 principles. The findings highlight that while some businesses are showing advanced capabilities for adaptation and technology adoption, the majority remain at a lower level of digitalization.

Future research will build on the results of this research and will focus on assessing the economic and operational impacts of introducing advanced technologies, such as digital twins or IoT systems, in electrical manufacturing. And customer-centric innovation, which will mainly focus on analysing how improved communication tools can increase

customer satisfaction and support demand for customised products in line with Industry 4.0 principles.

6 REFERENCES

[1] Khan, I., et al. (2025). Integrating industry 4.0 for enhanced sustainability: Pathways and prospects. *Sustainable Production and Consumption, 54*, 149–189. https://doi.org/10.1016/j.spc.2024.12.012

[2] Hong, Z., et al. (2024). Digital twin-based cross-enterprise production-delivery synchronization in a highly dynamic environment. *Computers & Industrial Engineering, 198*, 110680. https://doi.org/10.1016/j.cie.2024.110680

[3] Bruland, K., et al. (2013). Assessing the role of steam power in the first industrial revolution: The early work of Nick von Tunzelmann. *Research Policy, 42*(10), 1716–1723. https://doi.org/10.1016/j.respol.2012.12.008

[4] Woodrow, W., et al. (2015). Industrial Development. In *The Green Industrial Revolution* (Chapter 2, pp. 13–35). Butterworth-Heinemann. https://doi.org/10.1016/B978-0-12-802314-3.00002-0

[5] Finkelstein, J., & Newman, D. (1984). The third industrial revolution: A special challenge to managers. *Organizational Dynamics, 13*(1), 53–65. https://doi.org/10.1016/0090-2616(84)90031-7

[6] Centobelli, P., et al. (2022). Blockchain technology for bridging trust, traceability and transparency in circular supply chain. *Information & Management, 59*(7), 103508. https://doi.org/10.1016/j.im.2021.103508

[7] McKinsey & Company. (n.d.). *What is personalization?* https://www.mckinsey.com/~/media/mckinsey/featured%20insights/mckinsey%20explainers/what%20is%20personalization/what-is-personalization.pdf

[8] Deloitte. (n.d.). *Omnia AI – Marketing POV*. https://www2.deloitte.com/content/dam/Deloitte/ca/Documents/deloitte-analytics/ca-en-omnia-ai-marketing-pov-fin-jun24-aoda.pdf

[9] World Economic Forum. (2022). *Leaders and supply chains – Davos 2022*. https://www.weforum.org/stories/2022/05/leaders-supply-chains-davos-2022/

[10] Gallego-García, S., et al. (2022). Integration of improvement strategies and Industry 4.0 technologies in a dynamic evaluation model for target-oriented optimization. *Applied Sciences, 12*(3), 1530. https://doi.org/10.3390/app12031530

[11] Ansari, I., et al. (2023). An Industry 4.0 readiness model for new technology exploitation. *International Journal of Quality & Reliability Management, 40*(10), 2519–2538. https://doi.org/10.1108/IJQRM-11-2022-0331

[12] Digital Supply Chain Hub UK. (n.d.). *Digital supply chain readiness*. https://digitalsupplychainhub.uk/get-involved/digital-supply-chain-readiness/

[13] Monshizadeh, F., et al. (2023). Developing an industry 4.0 readiness model using fuzzy cognitive maps approach. *International Journal of Production Economics, 255*, 108658. https://doi.org/10.1016/j.ijpe.2022.108658

[14] UNDP North Macedonia. (n.d.). *Digital readiness assessment*. https://www.undp.org/north-macedonia/publications/digital-readiness-assessment

[15] Emon, M., & Khan, T. (2025). The transformative role of Industry 4.0 in supply chains: Exploring digital integration and innovation in the manufacturing enterprises. *Journal of Open Innovation: Technology, Market, and Complexity, 11*(2), 100516. https://doi.org/10.1016/j.joitmc.2025.100516

[16] Al-Okaily, M., Younis, H., & Al-Okaily, A. (2024). The impact of management practices and Industry 4.0 technologies on supply chain sustainability: A systematic review. *Heliyon, 10*(17), e36421. https://doi.org/10.1016/j.heliyon.2024.e36421

[17] NetSuite. (n.d.). *Digital supply chain*. https://www.netsuite.com/portal/resource/articles/erp/digital-supply-chain.shtml

[18] Pech, M., & Vrchota, J. (2022). The product customization process in relation to Industry 4.0 and digitalization. *Processes, 10*(3), 539. https://doi.org/10.3390/pr10030539

[19] Dassault Systèmes. (n.d.). *Industry 4.0 and mass customization*. https://www.3ds.com/manufacturing/connected-industry/industry-4-0-data-mass-customization-actionable-insights

[20] Lorenz, R., Lorentzen, K., Stricker, N., & Lanza, G. (2018). Applying user stories for a customer-driven Industry 4.0 transformation. *IFAC-PapersOnLine, 51*(11), 1351–1356. https://doi.org/10.1016/j.ifacol.2018.08.345

[21] Ramírez-Durán, V. J., Berges, I., & Illarramendi, A. (2021). Towards the implementation of Industry 4.0: A methodology-based approach oriented to the customer life cycle. *Computers in Industry, 126*, 103403. https://doi.org/10.1016/j.compind.2021.103403

Authors' contacts:

Andrea Benešová, PhD
(Corresponding author)
University of West Bohemia in Pilsen, Faculty of Electrical Engineering,
Department of Materials and Technologies,
Univerzitní 2795/26, 301 00 Pilsen, Czech Republic
benesov2@fel.zcu.cz

Martin Hirman, PhD
University of West Bohemia in Pilsen, Faculty of Electrical Engineering,
Department of Materials and Technologies,
Univerzitní 2795/26, 301 00 Pilsen, Czech Republic
hirmanm@fel.zcu.cz

František Steiner, Assoc. Prof.
University of West Bohemia in Pilsen, Faculty of Electrical Engineering,
Department of Materials and Technologies,
Univerzitní 2795/26, 301 00 Pilsen, Czech Republic
steiner@fel.zcu.cz

Jiří Tupa, Assoc. Prof.
University of West Bohemia in Pilsen, Faculty of Electrical Engineering,
Department of Materials and Technologies,
Univerzitní 2795/26, 301 00 Pilsen, Czech Republic
tupa@fel.zcu.cz

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250508113124

Preliminary communication
Received: 2025-05-08, Accepted: 2025-05-22

Perceived Helpfulness of Learning Methods for Developing Logistics Competencies in the Context of Digitalization, Sustainability, and Industry 4.0

Simona Šinko, Tone Lerher, Brigita Gajšek*

Abstract: This study explores the perceived helpfulness of addressed learning methods in developing competences essential for logistics activities, particularly those related to digitalization, sustainability, and Industry 4.0. A total of 623 respondents, including logistics students, employees, and representatives of logistics associations and leading-edge companies, rated 17 learning methods. The findings reveal a strong preference for experiential formats. Over 95% of respondents rated methods such as real-world problem-solving in companies, the use of professional software, and guided simulations helpful. In contrast, more passive methods like critical reading and independent study were perceived as less effective. Statistically significant differences between students and employees underscore the role of prior industry experience in shaping learning preferences. Survey's results highlight the value of applied learning and closer collaboration with industry in logistics education, offering actionable insights designing training programs that align with the logistics sector's evolving needs.

Keywords: competences; digitalization; Industry 4.0; learning methods; logistics; sustainability

1 INTRODUCTION

Logistics is managing material, and information flows within businesses to ensure that goods are delivered to the end customer in the most efficient and cost-effective manner. It involves the movement, storage, and handling of materials, along with the management of related information. As a fundamental component of the global supply chain, logistics ensures the coordinated and timely movement of goods and materials across suppliers, manufacturers, and consumers by managing transportation, warehousing, inventory, and distribution processes. The logistics supply chain typically includes supply, production, distribution, and reverse logistics [1]. In light of growing pressures to improve sustainability, logistics also faces the challenge of reducing its carbon footprint and adopting greener practices [2].

The logistics sector is transforming rapidly, driven by Industry 4.0 technologies such as artificial intelligence, automation, and data analytics. These technologies not only optimize efficiency but also address sustainability challenges [3, 4]. As a result, logistics professionals must adapt to new needs, requiring new competencies in technology, environmental awareness, and digitalization [5, 6].

Existing literature primarily addresses traditional logistics skills, such as transportation and inventory management, while overlooking the need for expertise in Industry 4.0, sustainability, and digitalization [7]. As part of a broader research agenda on the evolving competency requirements in logistics, this study specifically examines which learning methods most effectively support students in acquiring the skills necessary to navigate the challenges of digitalization, sustainability, and Industry 4.0.

Effective logistics practice requires both specialized knowledge and practical skills. While formal education provides a basic knowledge, hands-on experience and continuous learning are beneficial for progress in this rapidly changing field [8]. Lifelong learning, both formal and informal, is needed for professional growth and organizational success. Training programs should combine theoretical knowledge with practical application, leveraging case studies, simulations, and industry partnerships, to ensure that professionals can effectively transfer their learning in real-world practice [9].

The research by Stevens et al. [10] and Kozma [11] emphasizes the importance of collaboration between educators and students to ensure that learning outcomes are achieved, whether through in-person or online learning modalities. Online learning, in particular, has been found to provide greater flexibility and opportunities for collaboration, fostering higher cognitive engagement and enhancing knowledge retention [12]. A blended approach may be essential for equipping logistics professionals with the diverse skill set required in an increasingly digital and sustainable industry.

This paper addresses a gap in existing research by exploring how different stakeholder groups perceive the effectiveness of seventeen purposefully selected learning methods in fostering the competencies needed in modern logistics, necessary to navigate the challenges of digitalization, sustainability, and Industry 4.0. Three stakeholder groups were defined, namely full-time students, logistics employees, and representatives of leading-edge companies and logistics associations. Based on a large sample, this study provides insights into opportunities to improve logistics education and training.

Despite increasing focus to sustainability, digitalization, and Industry 4.0 in logistics, there remains limited understanding of which learning methods are most effective in preparing students to implement related activities. Existing educational programs often rely on traditional instructional formats, which may not align with the sector's evolving needs. This gap motivated the present study, which seeks to identify how different stakeholder groups perceive the effectiveness of addressed learning methods in developing relevant logistics capabilities.

Building on previously defined stakeholder groups, this research explores how perceptions regarding the effectiveness of learning methods differ among full-time students, logistics employees, logistics pioneers in Industry 4.0, digitalization and sustainability, and representatives of

national logistics associations, with a focus on preparation for real-world logistics challenges. This study provides actionable insights to help industry and educational institutions design more targeted training programs tailored to the future needs of the logistics industry.

The paper continues with an overview of the research methodology, followed by the presentation of results that identify key activities and skills essential for addressing the challenges posed by Industry 4.0, sustainability, and digitalization. It highlights the learning methods considered most effective for acquiring practical knowledge, based on insights from both educators and industry experts. The discussion section offers recommendations for enhancing logistics education, while the conclusion summarises key findings and proposes directions for future research.

2 METHODOLOGY

The methodological approach builds on four interconnected research stages, combining conceptual groundwork and empirical validations to:

- determine learning outcomes relevant for logistics higher education on sustainability and Industry 4.0;
- determine competences that logistics personel shoul have to cope with challenges of introducing new technologies and EU guidelines;
- access readiness of logistics personel to logistics industry because of digital transformation, development towards Industry 4.0 and introducing sustainable practices;
- examine perceived helpfulness of learning methods for developing logistics competences.

An overview of the key research stages and their respective outcomes is presented in Fig. 1. This paper focuses on the results of the final, fourth research phase.

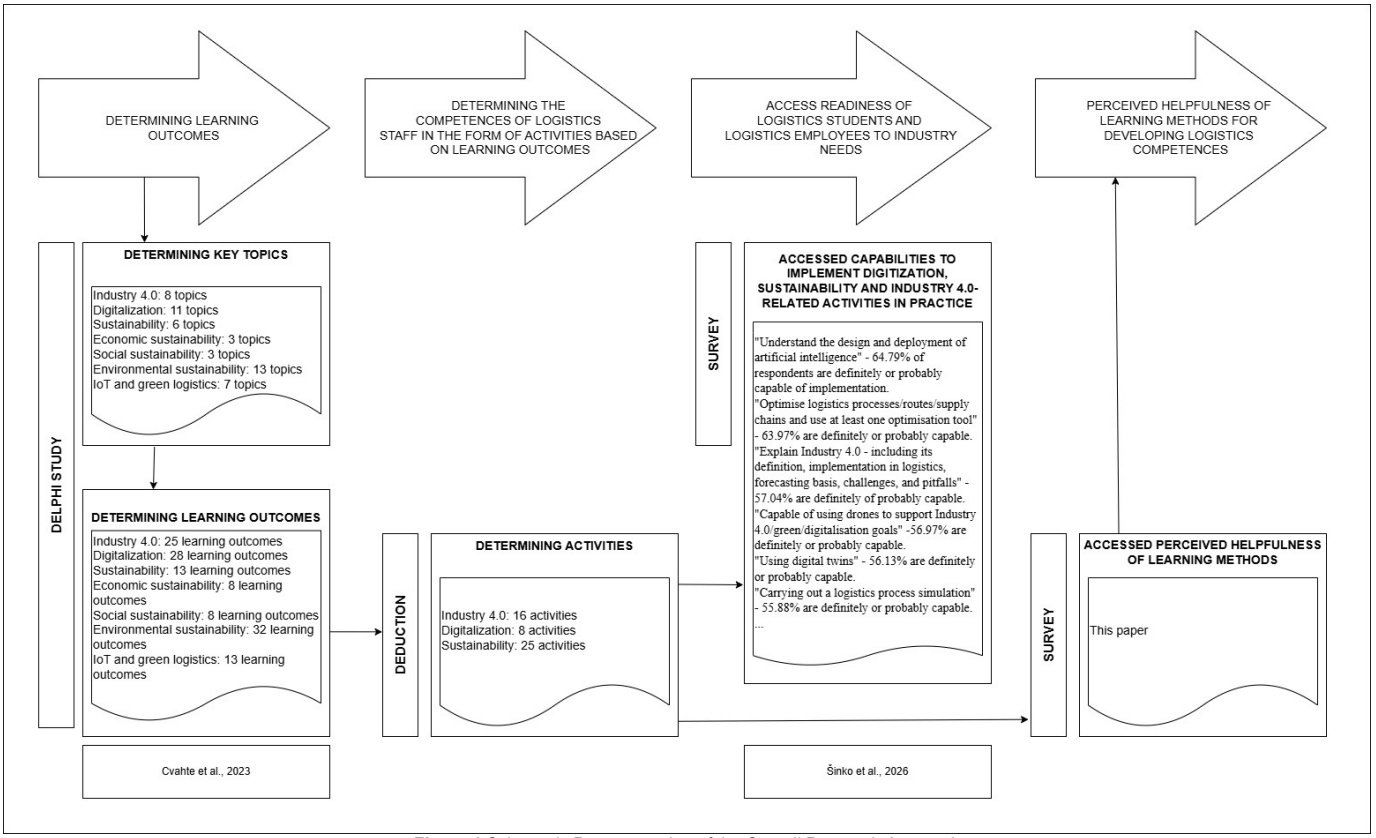


Figure 1 Schematic Representation of the Overall Research Approach

2.1 Research Design

The research process began with determining learning outcomes relevant for logistics higher education on sustainability and Industry 4.0 with the help of Delphi study. The group of competent stakeholders first determined key topics on Industry 4.0, digitalization, Sustainability, IoT and Green, for which they expect to shape the future of logistics work. Based on these topics, an initial list of learning outcomes was generated, reviewed, and iteratively refined.

All further research phases, including the fourth phase discussed in this paper, build on insights from Delphi study results described in [7], which identified 127 unique learning outcomes mandatory for logistics personnel in the context of Industry 4.0, digitalization, and sustainability. These learning outcomes were determined by an international group of experts from six countries. The panel included academics with at least two high-profile publications and company representatives from commercially successful logistics firms.

The following section presents the research design for the fourth phase of the study, which aims to investigate the

perceived effectiveness of learning methods in the development of logistics competencies. To enhance respondents' understanding of the learning outcomes in the survey, the initial list of 127 learning outcomes was systematically reviewed. For each outcome, potential activities that logistics personnel could implement in practice, particularly in the context of Industry 4.0 and digitalization, were identified and considered. The team of four academics and industry representative clustered learning outcomes according to their affiliation to management and implementation of logistics activities. The list of learning outcomes was thus converted into a list of logistical activities. Both academic and industry experts validated the final list of 49 logistics-related activities to ensure conceptual clarity, practical relevance, and alignment with real-world logistics applications.

Based on those 49 logistics-related activities, a survey questionnaire was developed. A detailed description of the questionnaire and its structure is provided in Section 2.4. The questionnaire served as a data collection mechanism for the two thematically related but independently conducted studies. The first part of questionnaire focused on assessing the readiness of logistics students and employees to meet emerging industry needs. The results are presented in [13]. The second part of questionnaire focused on researching how different respondent groups perceive their capability to perform offered logistics activities and assesses the perceived helpfulness of offered learning methods in fostering the development of those capabilities.

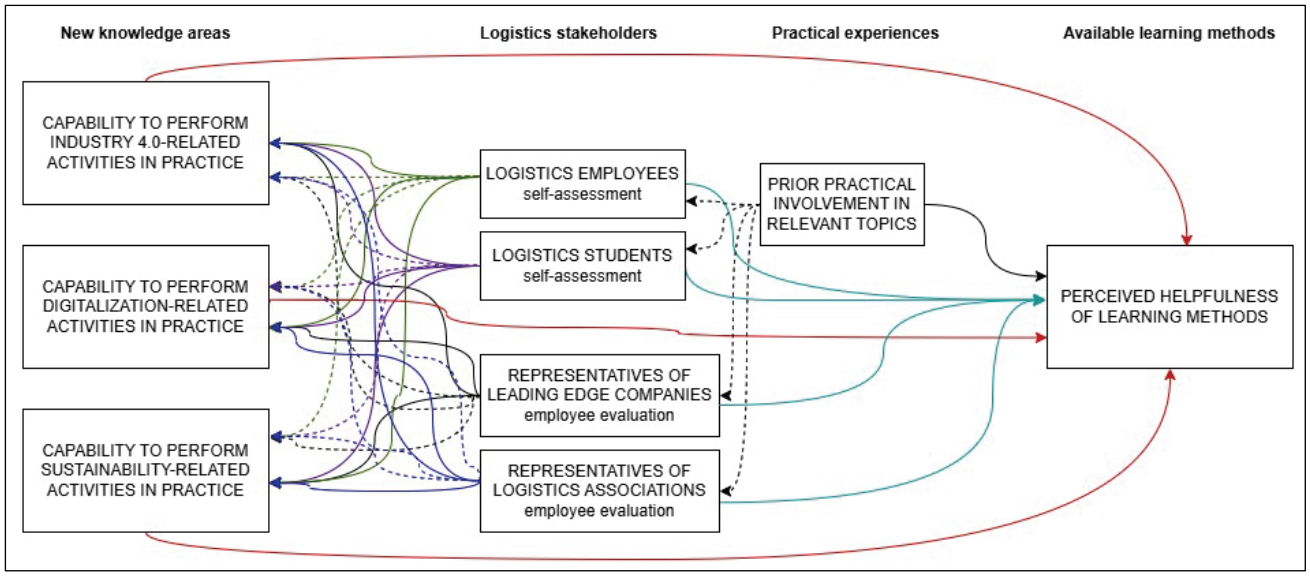


Figure 2 Conceptual Framework Showing Relationships among Key Variables in the Study

2.2 Conceptual Framework of the Research

The conceptual framework of the research is presented in Fig. 2. It presents the researched connections between logistics stakeholders, knowledge areas and learning methods. Presented framework served as a basis for developing the survey instrument and guided the subsequent statistical analysis.

The framework is structured around three main capability domains: the capability to perform Industry 4.0-related, digitization-related, and sustainability-related logistics activities. The framework is organized around three primary capability domains: the ability to perform logistics activities related to Industry 4.0, digitalization, and sustainability. These domains represent essential knowledge areas in contemporary logistics practice. The inclusion of these capabilities aimed to assess their prevalence across different stakeholder groups and to explore potential factors influencing their development.

To capture different perspectives, the framework includes self-assessment of capabilities by logistics employees and students and external evaluations provided by representatives of leading-edge companies and logistics associations regarding the capabilities of the logistics professionals currently active in the field.

In addition, the framework incorporates prior practical involvement in relevant topics, intended to explore whether such experiences are associated with higher capability levels or influence the perceived effectiveness of different learning methods.

Finally, the framework incorporates the perceived helpfulness of learning methods, reflecting how various educational approaches are evaluated in terms of their contribution to developing the specified capabilities. While not the central focus of the study, this variable plays a significant role in linking educational practices with the practical competence of logistics professionals.

2.3 Survey Instrument

The questionnaire was divided into two main sections. The first section assessed respondents' self-reported capabilities in implementing the identified activities in real-world settings. The analysis of this section was presented in prior research [13], which focused on comparing the readiness of students and logistics employees to engage with

digitalization, sustainability, and Industry 4.0-related activities.

The second section of the questionnaire explored respondents' perceptions of the helpfulness of various learning methods in developing the competencies needed to carry out the listed logistics activities. Students and logistics employees were asked: *"How would the following ways of acquiring knowledge help you acquire the above-listed competencies*?*"* Meanwhile, representatives of leading-edge companies and national logistics associations were asked: *"Which ways of acquiring knowledge would help acquire the above-listed competencies*?*"*

A total of 17 learning methods were evaluated using a 4-point Likert scale, ranging from 1 (definitely won't help) to 4 (definitely will help).

Prior to launching the full survey, the questionnaire underwent a two-phase validation process. In the first phase, it was reviewed by 14 validators representing various respondent profiles: three logistics students, five logistics employees, two representatives from national logistics associations, and four representatives of leading-edge companies in Industry 4.0 and sustainability. Feedback from this group informed revisions to the wording and structure of the questionnaire. In the second phase, the revised version was validated by experts from an Italian deep-tech company, who confirmed its clarity and alignment with the purpose of the research.

2.4 Sample and Data Collection

The survey targeted five respondent groups: (1) logistics students (514 respondents, 82.6%), (2) logistics employees (75 respondents, 12.1%), (3) representatives of leading-edge companies in Industry 4.0 (15 respondents, 2.5%), (4) representatives of representatives of leading-edge companies in sustainability (11 respondents, 1.8%), and (5) representatives of national logistics associations (7 respondents, 1.1%). The sample primarily included participants from Italy, Poland, Portugal, Slovenia, and Turkey. An additional 14.68% of respondents were from other European countries, including Austria, Brazil, Bulgaria, Bosnia and Herzegovina, Croatia, Czech Republic, Germany, Hungary, Netherlands, Romania, Russia, Spain, the United Kingdom, and Ukraine.

The online questionnaire was distributed using 1KA Arnes, a Slovenian platform for survey creation and data collection. It was shared with students from various logistics faculties and employees from logistics companies to ensure broad representation. Respondents representing leading-edge companies in Industry 4.0 and sustainability were carefully selected based on their companies' recognized excellence, with a focus on targeting middle or senior management. Heads of national logistics associations were also included to capture insights from a strategic logistics perspective. All respondents were anonymous, compliant with GDPR regulations, and used exclusively for research purposes. In total, 623 valid responses were collected, yielding a response rate of 38.9%.

2.5 Data Analysis and Availability

Data were analysed using IBM SPSS Statistics (29.0.0.0). Descriptive statistics were applied to summarise the responses, and comparisons between groups were conducted to identify trends in the perceived helpfulness of different learning methods.

The results are presented in the third section, with tables included where they are essential for clarity. For full transparency, all raw data and supplementary materials, including detailed analysis outputs, are available via the open-access link: https://osf.io/gx8az/?view_only=d877ce2fea47433fad3998546cbb8c97.

3 RESULTS

3.1 Approaches to Acquiring Knowledge for Competency Development

Tab. 1 presents respondents' perceptions of the helpfulness of various learning methods in acquiring the competencies needed to implement logistics activities related to Industry 4.0, digitalization, and sustainability. Only valid responses are included, excluding "unable to respond" entries.

The method most frequently rated as the most helpful for acquiring knowledge was the *"Study of a real problem in a real company"*, with over 70% of respondents indicating it would definitely help, and an additional 25% rating it as probably helpful. This was followed by *"Performing tasks with software that is widely used in the work environment"* (93.88% of respondents rated it as probably or definitely helpful, *"Use of video materials"* (93.81%), and *"Practical implementation of complex tasks under guidance"* (93.67%). Similarly, *"Guided use of a simulation tool"* (92.89%) and *"Listening to experts from business or professional practice on a specific topic"* (91.83%) were also perceived as highly helpful.

Other methods evaluated as helpful by more than 90% of respondents include *"Team Working with other learners in a course"* (91.27%), *"Face-to-face teaching"* (90.50%), *"Group discussion about a real problem"* (90.43%), and *"Study of a case in a laboratory"* (90.26%).

In contrast, more traditional methods were considered less helpful by a larger proportion of respondents. *"Critical reading and writing as preparation for lectures"*, *"Role-playing"*, and *"Working independently of other learners in a course"* were rated as less helpful by a significant share of participants. Nearly one-quarter of respondents felt that both *critical reading and writing* and *independent work* would probably or definitely not contribute to acquiring relevant knowledge. Additionally, *"Distance learning from home with the help of an online classroom containing materials and pre-recorded lectures"* was deemed less beneficial by 19.8% of respondents, who indicated it would not contribute to acquiring the needed knowledge.

Overall, the findings strongly favor applied, hands-on learning that emphasizes engagement with real-world challenges and expert insights. While traditional classroom-based approaches are still regarded as useful, they are

considered less effective than experiential learning in developing the relevant logistics competencies.

Table 1 Respondents' perceived helpfulness of learning methods

Learning method	1 Definitely won't help (%)	2 Probably won't help (%)	3 Probably will help (%)	4 Definitely will help (%)
Distance learning, online classroom containing materials and pre-recorded lectures	7.18	12.62	44.31	35.89
Lectures with presentations on theoretical knowledge	3.68	9.80	45.34	41.18
Critical reading	5.56	19.44	52.02	22.98
Listening to experts from business/practice on a specific topic	1.73	6.44	29.95	61.88
Group discussion about a real problem	1.51	8.06	35.52	54.91
Use of games for teaching and learning	3.04	9.62	39.75	47.59
Use of video materials	0.99	5.20	45.54	48.27
Role playing	6.91	17.14	41.94	34.02
Practical implementation of complex tasks under the guidance	2.03	4.30	36.71	56.96
Guided use of simulation tool	1.52	5.58	38.83	54.06
Performing tasks with software that is widely used in the practice	1.79	4.34	35.46	58.42
Work on a fictional case study in the classroom	3.07	7.42	45.27	44.25
Study of a case in a laboratory	2.56	7.18	41.79	48.46
Study of a real problem in a real company	0.74	3.72	25.06	70.47
Team Working with other learners in a course	2.49	6.23	42.14	49.13
Work independently of other learners in a course	4.03	19.14	46.60	30.23
Face to face teaching	4.50	5.00	43.25	47.25

3.2 Perception Differences of Learning Methods across Respondent Profiles

Tab. 2 summarizes statistically significant differences ($p < 0.05$) between students, logistics employees, and representatives from leading-edge companies and national logistics associations, as determined by Mann-Whitney U test.

Overall, logistics employees and representatives of leading-edge companies and national logistics associations perceived the learning methods as more helpful for acquiring knowledge compared to students. A statistically significant difference emerged only for the method *"Distance learning, from home with the help of an online classroom containing materials and pre-recorded lectures"*. No significant differences were found among the remaining 16 activities.

Table 2 Statistically significant differences in perceived helpfulness of learning methods across respondent profiles

Learning method	The statistically significant difference expressed with mean ranks	Mann-Whitney U test
Distance learning, from home with help of an online classroom containing materials and pre-recorded lectures	Employees (221.15) - Students (186.30)	7572.50*
Group discussion about a real problem	Employees (31.46) – RLEC Sustainability (44.00)	316.00*
	RNLA (4.83) - RLEC (9.50)	40.00*
	RLEC Industry 4.0 (6.33) – RLEC Sustainability (12.00)	60.00*
	RLEC Sustainability (236.50) - Student (161.15)	680.00*
Performing tasks with software that is widely used in the work environment	Employees (30.79) - RLEC Sustainability (44.50)	320.00*
	RLEC Sustainability (226.50) - Student (159.33)	728.00*
	RNLA (4.83) - RLEC Sustainability (9.5)	40.00*
Practical implementation of complex tasks under the guidance	RNLA (4.00) – RLEC Sustainability (8.29)	30.00*
	RNLA (82.30) - Student (162.75)	1,188.50*
Team Working with other learners in a course	RLEC Sustainability (231.00) - Student (163.36)	756,00*
Face to face teaching	RNLA (95.83) - Student (165.78)	1,378.00*

RLEC - representatives from leading-edge companies; RNLA - representative of national logistics associations; *The significance level is 0.05

Logistics employees differed from representatives of leading-edge companies and national logistics associations in their evaluations of three methods, rating them as less helpful: *"Group discussion about a real problem"*, *"Representative of an outstanding example of sustainability implementation"*, and *"Performing tasks with software that is widely used in the work environment"*.

Differences between students and representatives of leading-edge companies and national logistics associations were observed in the evaluation of five learning methods. Students rated *"Face-to-face teaching"* and *"Practical implementation of complex tasks under guidance"*. As more helpful than another group did. In contrast, representatives of leading-edge companies and national logistics associations rated the following learning methods as more helpful: *"Group discussion about a real problem"*, *"Performing tasks with software that is widely used in the work environment"*, and *"Team Working with other learners in a course"*.

4 DISCUSSION

The study's findings reveal a clear preference for experiential and practice-oriented learning methods, which were perceived as more helpful by a larger share of respondents compared to traditional learning methods, such as independent study or critical reading. This preference aligns a growing body of literature emphasising the value of hands-on learning in developing the competencies and skills

necessary for real-world logistics tasks. As Fabbe-Costes [14] notes, collaboration between academia and industry remains a key success factor in advancing logistics knowledge. Notably, this insight was already highlighted in 2012 [15], yet the present findings indicate that its integration into educational practices still requires further reinforcement.

The positive evaluation of approaches such as group discussions on real-world problems and the use of widely adopted logistics software reinforces this perspective. Both learning methods promote active engagement with practical challenges and help bridge the gap between theoretical instruction and professional practice.

The preference for practical learning is further supported by Chernikova et al. [16], who highlight simulations as effective pedagogical tools for fostering problem-solving skills. In the context of logistics, gamified modelling and simulation frameworks are especially valuable, as they contribute to the optimization of real-world logistics processes and the advancement of planning methodologies [17]. Simulation games not only enhance learning in higher education but also serve as effective tools in workplace training—improving decision-making and providing a dynamic, systems-level understanding of logistics operations [18].

Furthermore, the results suggest that respondents' evaluations of learning methods may be shaped by their personal learning preferences and prior experiences. Notable differences were observed between the perceptions of students and representatives of leading-edge companies and national logistics associations. Students generally rated learning methods such as distance learning and group discussions lower than logistics employees. A plausible explanation for this trend is students' limited exposure to real-world logistics environments, which may lead them to underestimate the relevance of methods designed to simulate or address practical challenges.

Another contributing factor may stem from traditional educational experiences, which often emphasize passive learning. As a result, students may feel unfamiliar or even uncomfortable with more interactive learning methods. Differences in perception could also be attributed to varying levels of familiarity with specific learning methods, such as online learning or group collaboration. For instance, students may have had limited exposure to well-structured digital learning environments or collaborative group work settings, which could influence their evaluations. Additionally, perceived self-efficacy in engaging with collaborative tasks may also play a role. However, these hypotheses require further empirical investigation to better understand their impact on the perceived helpfulness of learning methods.

These findings align with the broader literature, which suggests that traditional classroom methods often lack the practical applications necessary for students to effectively address real-world logistics challenges [19, 20]. Independent study tends to engage students less than experiential learning approaches. Research has shown that game-based learning and simulations are more effective in engaging students by offering interactive, real-world scenarios [21]. In contrast, traditional methods such as critical reading do not sufficiently foster key soft skills, like communication, problem-solving, and teamwork, that are better cultivated through collaborative learning environments [22].

Additionally, students tend to favour methods that incorporate technology and real-life applications, such as logistics simulation tools and learning factories, as these approaches closely mirror real-world logistics tasks [19]. While autonomous learning strategies offer theoretical benefits, they may lack the structure and guidance that some students need to develop practical skills. However, methods involving group work, leadership roles, and peer evaluation have been shown to enhance responsibility and engagement, leading to improved learning outcomes [22].

This study emphasizes the importance of integrating practical, hands-on learning experiences into logistics education. The strong preference for experiential learning, along with the observed differences in perceptions between students and logistics employees, indicates that educational programs should prioritize industry engagement, simulations, and collaborative projects. Universities could consider redesigning curricula to incorporate real-world case studies, internships, or company-based projects, while companies might focus on on-the-job training, mentoring, and guided simulations. Aligning training formats with these preferences will not only enhance the relevance of competency development but also strengthen the link between academia and industry, effectively bridging the gap between theory and practice.

5 CONCLUSION

This study explored how addressed learning methods are perceived in terms of their effectiveness in developing competencies in logistics, with particular focus on digitalization, sustainability, and Industry 4.0.

5.1 Theoretical Contribution

The findings highlight the increasing significance of experiential learning in logistics education. As the sector continues to evolve rapidly, traditional, theory-based models are no longer productive. Competence in logistics now requires not only a solid foundation of theoretical knowledge but also the ability to apply that knowledge in real-world contexts. This study contributes to a deeper understanding of how applied learning approaches support professional readiness in this dynamic field.

5.2 Practical Implications

The study provides clear guidance for education providers and industry stakeholders: move beyond traditional theory and incorporate hands-on training. Practical learning formats, such as simulations, workshops, and collaborative projects, should be embedded into both academic curricula and corporate training programs.

As logistics continues to evolve under the pressures of digitalization, sustainability imperatives, and the rise of Industry 4.0, education systems must move beyond incremental improvements and embrace a fundamental shift in how learning is designed, delivered, and integrated with industry. Future logistics education should not only respond

to current trends but anticipate them by fostering proactive, adaptive, and innovation-driven professionals. This requires embedding learning directly within operational contexts and utilizing advanced technologies, such as AI-powered simulators, digital twins, and immersive virtual environments, to replicate the complexity of real-world logistics challenges.

Furthermore, collaboration between academia and industry must become systemic rather than an incidental part of logistics education. Joint curricula development, co-supervised projects, and long-term partnerships will ensure that educational content evolves in alignment with real-world demands. Education providers should strive to create living laboratories—ecosystems where students, educators, and industry partners co-create knowledge, test innovations, and tackle authentic logistics challenges. To fully realize this vision, educational policy must support flexible, modular, and personalized learning pathways, along with stronger incentives for lifelong learning. By rethinking traditional boundaries between classroom and workplace, theory and practice, and student and practitioner, the logistics sector can cultivate a future-ready workforce capable of adapting to and leading in an evolving industry world.

5.3 Limitations and Future Research Directions

While this study provides valuable insights, several limitations should be acknowledged. The sample is predominantly composed of students, whose responses may reflect educational experiences more than practical industry realities. Although the large number of student responses increases statistical power, it may introduce sampling bias and skew the results toward academic rather than practical perspectives. Furthermore, the geographical scope of primarily covering six European countries may limit the generalisability of findings to regions with different educational systems or different technological maturity.

Another limitation lies in the reliance on self-reported data, which introduces potential bias, as respondents may not always accurately assess the helpfulness of specific methods. Emotions, limited experience with certain learning methods, or personal preferences could influence respondents' evaluations.

Future research should consider more balanced sampling across respondent profiles, using stratified or quota-based approaches. In addition, integrating objective performance-based assessments, such as simulations, task completion, or real-world problem-solving, could provide a more accurate picture of the actual effectiveness of learning methods. Longitudinal studies would further clarify how specific learning approaches contribute to competence development over time.

Finally, combining quantitative surveys with qualitative approaches, such as interviews or focus groups, may yield more profound insights into why specific learning methods are perceived as more or less helpful and how context influences these perceptions. Further investigation into the relative effectiveness of online versus in-person learning, particularly in relation to technical and soft skills, could also support the development of more tailored educational strategies.

Acknowledgements

This research was funded by the Erasmus+ project Sustainable Logistics 4.0: Digital and green skills for boosting innovation and sustainability of the logistics sector [Slog4.0].

6 REFERENCES

[1] Dey, A., LaGuardia, P., & M. Srinivasan, M. (2011). Building sustainability in logistics operations: a research agenda. *Manag. Res. Rev*., *34*(11), 1237–1259. https://doi.org/10.1108/01409171111178774

[2] Montreuil, B. (2011). Toward a Physical Internet: meeting the global logistics sustainability grand challenge. *Logist. Res*. 3, 71–87.

[3] Gammelgaard, B. (2019). Editorial: Congratulations to IJLM on its first 30 years. *Int. J. Logist. Manag., 30*(1), 2–7. https://doi.org/10.1108/IJLM-02-2019-348

[4] Winkelhaus S., & Grosse, E. H. (2020). Logistics 4.0: a systematic review towards a new logistics system. *Int. J. Prod. Res*., *58*(1), 18–43. https://doi.org/10.1080/00207543.2019.1612964

[5] Matt, D. T., & Zsifkovits, H. E. (2021). *Industry 4.0 for SMEs: Challenges, opportunities and requirements*. Switzerland: Springer Nature: Cham.

[6] Bag, S., Gupta, S., Kumar, A., & Sivarajah, U. (2021). An integrated artificial intelligence framework for knowledge creation and B2B marketing rational decision making for improving firm performance. *Ind. Mark. Manag*., *92*, 178–189. https://doi.org/10.1016/j.indmarman.2020.12.001

[7] Cvahte Ojsteršek, T., Šinko, S., & Gajšek, B. (2023). Determining learning outcomes relevant for logistics higher education on sustainability and Industry 4.0. *Teh. Glas*., *17*(3), 447–454, 2023. https://doi.org/10.31803/tg-20230505103326

[8] Thai, V. V., Tran, D. M., Nguyen, L. T. N., Nguyen, T. K., & Nguyen, T.-T. (2024). Revisiting the Role of Supply Chain Managerial Competence in Firm Logistics Performance: Do Experience and Education Matter? *Sustainability*, *16*(23), 10459. https://doi.org/10.3390/su162310459

[9] Becker, C., Rager, R. C., & Wright, F. E. (2013). Update on Validity of Required Competencies for Worksite Health Professionals. *Am. J. Health Educ*., *44*(2), 67–73. https://doi.org/10.1080/19325037.2013.764236

[10] Stevens, G. J., Bienz, T., Wali, N., Condie, J., & Schismenos, S. (2021). Online university education is the new normal: but is face-to-face better? *Interact. Technol. Smart Educ*., *18*(3), 278–297. https://doi.org/10.1108/ITSE-08-2020-0181

[11] Kozma, R. B. (1994). Will media influence learning? Reframing the debate. *ETR&D*, *42*, 7–19. https://doi.org/10.1007/BF02299087

[12] Nguyen, T. (2015). The Effectiveness of Online Learning: Beyond No Significant Difference and Future Horizons, *MERLOT Journal of online learning and teaching*, *11*(2), 309-319.

[13] Šinko, S., Mrugalska, B., & Gajšek, B. (2026). Readiness of Students and Employees in Logistics to Implement Digitization, Sustainability and Industry 4.0 in Practice. In *Sustainable innovations and performance of organizations in the age of digitalization*, S. Umair, U. Waqas, and B. Mrugalska, Eds., in Industrial and systems engineering series. Taylor & Francis Group.

[14] Fabbe-Costes, N. (2018). Logistics knowledge creation in joint industry-academia research projects: the importance of

dialogue and co-construction. *Knowl. Manag. Res. Pract.*, *16*(4), 464–476. https://doi.org/10.1080/14778238.2018.1486788
[15] Thai, V. V. (2012). Competency requirements for professionals in logistics and supply chain management. *Int. J. Logist. Res. Appl*., *15*(2), 109–126. https://doi.org/10.1080/13675567.2012.694859
[16] Chernikova, O., Heitzmann, N., Fink, M. C., Timothy, V., Seidel, T., Fischer, F., & DFG Research group COSIMA. (2020). Facilitating diagnostic competences in higher education—a meta-analysis in medical and teacher education. *Educational Psychology Review*, *32*, 157-196. https://doi.org/10.1007/s10648-019-09492-2
[17] Hubl, A., & Fischer, G. (201). Simulation-based business game for teaching methods in logistics and production. In 2017 *Winter Simulation Conference* (WSC), Las Vegas, NV: IEEE, 4228–4239. https://doi.org/10.1109/WSC.2017.8248129
[18] Jiang, P., Wang, Y., Liu, C., Hu, Y.-C., & Xie, J. (2020). Evaluating Critical Factors Influencing the Reliability of Emergency Logistics Systems Using Multiple-Attribute Decision Making. *Symmetry*, *12*(7), 1115. https://doi.org/10.3390/sym12071115
[19] Fu, H. (2017). Integration of Logistics Simulation Technology and Logistics Learning Factory in a Two-stage Teaching Method for Logistics Management Courses. *Int. J. Emerg. Technol. Learn. IJET*, 12(09), 62. https://doi.org/10.3991/ijet.v12i09.7485
[20] Schinckus, C., & Nguyen, T. T. (2022). Implementation of a real-life educational method for logistics in Vietnam: between Episteme and Techne. *Int. J. Logist. Res. Appl*., 25(11), 1469–1489. https://doi.org/10.1080/13675567.2021.1914562
[21] Pacheco-Velazquez, E., Rodés, V., & Salinas-Navarro, D. (2024). Developing learning skills through game-based learning in complex scenarios: A case in undergraduate logistics education. *J. Technol. Sci. Educ*., 14(1), 169. https://doi.org/10.3926/jotse.2219
[22] Carravilla, M. A., & Oliveira, J. F. (2004). Teaching logistics without formal classes: a case study. *Eur. J. Eng. Educ*., 29(4), 571–580. https://doi.org/10.1080/03043790410001711289

Authors' contacts:

Simona Šinko, PhD, Assistant Professor
Faculty of Logistics, University of Logistics,
Mariborska cesta 7, 3000 Celje, Slovenia
simona.sinko@um.si

Tone Lerher, PhD, Full Professor
Faculty of Logistics, University of Logistics,
Mariborska cesta 7, 3000 Celje, Slovenia
tone.lerher@um.si

Brigita Gajšek, PhD, Associate Professor
(Corresponding author)
Faculty of Logistics, University of Logistics,
Mariborska cesta 7, 3000 Celje, Slovenia
brigita.gajsek@um.si

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250516203854

Preliminary communication
Received: 2025-05-16, Accepted: 2025-05-26

Aeye: Real-Time Age, Gender and Emotion Detection in Browser

Mihael Holi*, Sven Maričić, Irena Jurjević, Adriana Prunk Drmić, Antonija Ružić Baršić

Abstract: This paper presents Aeye, a lightweight, browser-based application capable of real-time detection of human age, gender and emotional state using artificial intelligence. Unlike many existing AI systems, Aeye performs all computation on the client side, preserving user privacy and enabling full offline functionality. The application utilizes face-api.js and Tiny Face Detector models, achieving fast and efficient performance even on low-resource devices. Its modular design supports dynamic interface controls and responsive operation across platforms. The scientific contribution lies in the integration of asynchronous model loading, real-time GPU-accelerated inference using TensorFlow.js, and deployment as a progressive web app (PWA). These innovations make Aeye suitable for educational, experimental, and field use cases where connectivity and infrastructure are limited. The paper outlines the system architecture, implementation, and evaluation results, and discusses potential extensions such as multi-face detection and integration with emotion-driven user interfaces.

Keywords: age estimation; emotion recognition; face detection; gender classification; real-time AI

1 INTRODUCTION

The rapid development of artificial intelligence (AI) and browser-based technologies has enabled the deployment of intelligent systems that operate entirely on the client side. By leveraging lightweight machine learning models and JavaScript-based frameworks, modern applications are now capable of performing real-time computer vision tasks without requiring server infrastructure [1, 2].

Facial analysis, as one of the most active areas in applied computer vision, has seen significant progress in recent years. Applications that estimate age, detect gender, and recognize emotional states are increasingly used in fields such as education, accessibility, digital interaction, and user profiling [3]. The combination of speed, simplicity, and privacy offered by client-side implementations makes them especially attractive for non-invasive, demonstrative, or offline use cases [4].

Aeye is a web-based application designed to detect human faces and estimate their age, gender, and current emotional expression in real time. The system is built using the face-api.js library, which runs on top of TensorFlow.js, and incorporates pre-trained models for facial detection and analysis. The application utilizes the Tiny Face Detector model to identify faces from the webcam feed, after which it applies expression and demographic models to extract meaningful information [5].

All processing is done locally in the browser, which preserves user privacy and ensures compatibility with mobile and desktop platforms. In addition, the user interface is designed to be intuitive and accessible, offering functionalities such as fullscreen mode, camera flipping, dark mode, and screenshot capture. This paper outlines the technical implementation of Aeye and explores its capabilities, limitations, and potential applications in real-world scenarios.

Aeye distinguishes itself from existing solutions by offering a fully client-side implementation without requiring any backend infrastructure. Through asynchronous loading of models and browser caching, the application ensures fast startup and continuous operation even in low-connectivity environments. GPU acceleration is leveraged via WebGL integration with TensorFlow.js, which enables real-time inference across a wide range of devices. These characteristics make Aeye not only a privacy-conscious solution but also highly suitable for educational environments, public exhibitions, and research prototypes in human-computer interaction.

2 PROGRAMMING IMPLEMENTATION OF AEYE

2.1 Technologies and Architecture

The application Aeye is developed using the **face-api.js** JavaScript library, which is built on top of TensorFlow.js and provides a high-level API for performing face detection, facial expression recognition, and age and gender estimation in the browser [3]. The system relies on three pre-trained models: *Tiny Face Detector (TFD)*, *Age & Gender Recognition Net*, and *Face Expression Recognition Net*. All models are loaded asynchronously at application startup and cached in the browser. This enables real-time processing while preserving user privacy, as no data is sent to external servers [5, 6]. The application was designed to run on both desktop and mobile browsers, using only HTML, CSS and JavaScript, with no server-side components required beyond static file hosting.

The use of **TinyFaceDetector** enables significant performance optimization. It strikes a balance between speed and accuracy, especially on low-resource devices, making it ideal for education and public demonstrations [4, 6-8]. TensorFlow.js integration allows the application to leverage GPU acceleration through WebGL, further improving frame rate and responsiveness.

2.2 Workflow and Logic Structure

The application initializes the webcam stream using the browser's **getUserMedia** API and continuously processes the video frames. The primary steps in each processing loop include:

- **face detection**: executed using TFP for lightweight and fast detection [8]
- **age and gender estimation**: estimates demographic attributes with an additional smoothing function to reduce fluctuations
- **emotion recognition**: classifies emotions into seven categories using probabilities and selects the top prediction for display.

Results are rendered on an HTML canvas using **face-api.js**'s drawing utilities. The application maintains a smoothed average of recent predictions for improved readability. The logic is modular. For example, detection parameters such as input size and threshold can be adjusted via on-screen controls. Event listeners update model settings in real time, allowing dynamic reconfiguration during execution [9-11]. The processing pipeline for facial detection and analysis is summarized in Fig. 1, illustrating how age, gender, and emotion are extracted from live webcam input.

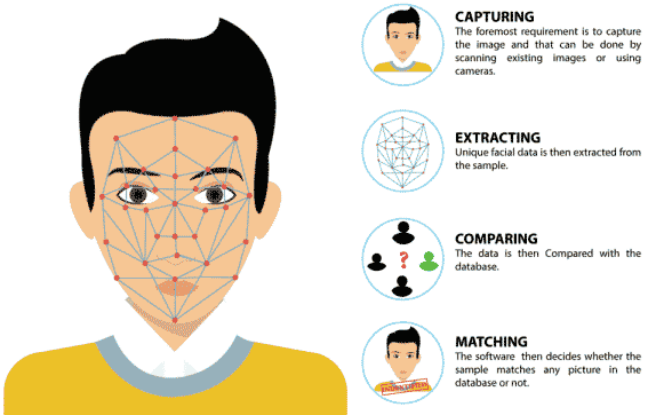


Figure 1 Pipeline for detecting face and extracting demographic and affective features [11]

2.3 Interactive Interface Design

The interface was designed to provide intuitive control and visual feedback. It includes:

- **live webcam video** with overlaid detection box and labels
- **display panel** with age, gender, and emotion data
- **buttons for toggling** dark mode, fullscreen, screenshot, and camera flip
- **fallback UI** elements for mobile users.

All user interface components are responsive and designed with accessibility in mind. Settings such as input resolution and detection thresholds can be adjusted for performance tuning. The application can function offline after initial model loading, making it suitable for classroom demos, fairs, or low-connectivity environments [2, 4]. Dark mode support is included to improve readability and visual comfort in different lighting conditions. Users can switch between front and rear cameras if multiple cameras are available. The fullscreen mode expands the webcam and overlay for immersive demonstrations [12]. An example of the application interface in operation is shown in Fig. 2 and Fig. 3. These illustrate the live detection overlay and UI elements including camera toggle, fullscreen and screenshot buttons.

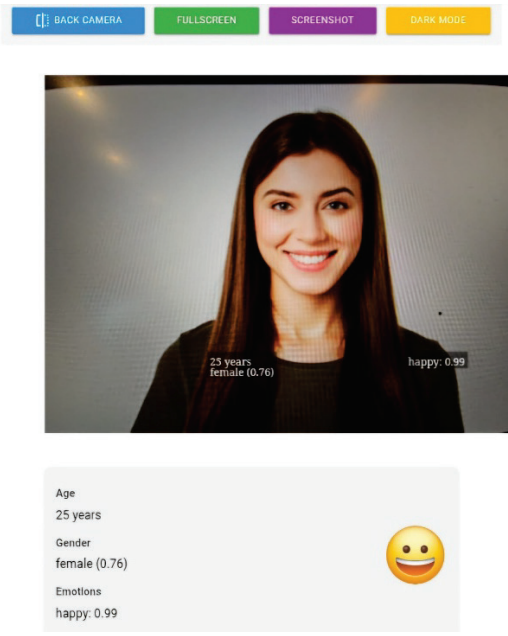


Figure 2 User interface showing live webcam detection with age, gender, and emotion overlay [12]

2.4 Software Components and File Structure

The application's structure consists of modular JavaScript files organized as follows:

- **server.js**: express server to host static files
- **index.html**: main UI containing canvas, buttons, and layout
- **faceDetectionControls.js**: control logic for adjusting model parameters
- **commons.js**: common UI rendering helpers
- **imageSelectionControls.js**: Image input handling
- **bbt.js**: Training example handlers for demo modes.

All models are stored in a **weights** folder and loaded relative to the deployment path. The layout follows a typical separation of concerns between interface, logic, and utility scripts. Additionally, the application includes a **manifest.json** and service worker (**sw.js**) for PWA support, enabling installation as a standalone offline-capable web app.

2.5 Deployment and Device Compatibility

The project has been deployed on lightweight Node.js environments and static servers. On mobile, the application functions without requiring an app store download. Tests confirmed compatibility with Android and iOS Chrome, Edge, and Safari browsers.

Its small footprint and static deployment model make it feasible for use in low-resource schools and exhibition spaces. Furthermore, since no login or user data is stored or

transmitted, it satisfies many privacy requirements by default [9].

3 DATA TESTING AND SYSTEM EVALUATION

3.1 Functional Validation

The application was tested across platforms: Windows 10, Android 12, and macOS Ventura. Browser compatibility was verified with Chrome, Firefox, and Safari. All platforms successfully detected faces, with average frame rates ranging from 15–25 FPS on modern hardware. Detection accuracy decreased under poor lighting and for occluded or rotated faces. Age estimation proved most stable in the adult age range (20–50), with increased deviation for children and elderly. Gender predictions were reliable, and emotions were generally well-classified in frontal facial orientation [10-14].

The average detection latency was measured across five test sessions using three different consumer-grade devices: a Windows 10 laptop with Intel i5 processor (8 GB RAM), an Android 12 smartphone (Snapdragon 730), and a macOS Ventura laptop (M1 chip). Each session included 3 to 5 participants who triggered detections by aligning their faces with the camera and observing on-screen prediction updates. Latency was defined as the time between face alignment and the first complete overlay of age, gender, and emotion data. Measurements were manually timed using high-frame-rate video playback (60 fps) and averaged over 10 detection cycles per device. Under optimal lighting, latency remained below 100 milliseconds on all platforms.

Tests with multiple volunteers confirmed consistent results in real-time, with very low latency (<100 ms) for detection and classification under ideal conditions. The emoji-based emotion feedback was intuitive for users of all ages. Fig. 4 presents a breakdown of the emotion classification logic, outlining how facial geometry features are mapped to probability-based emotion outputs.

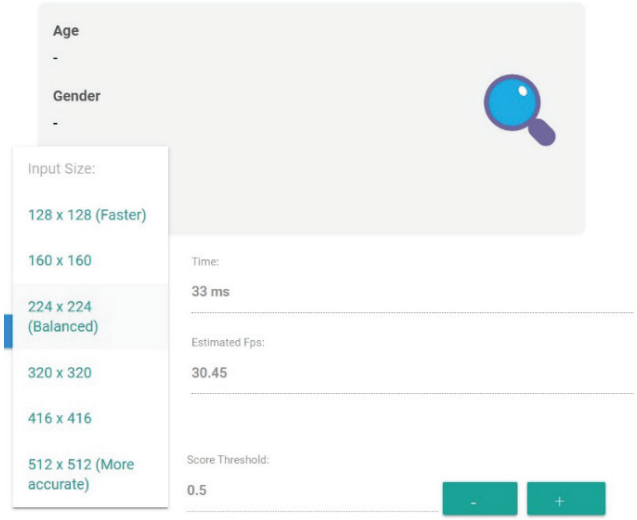


Figure 3 User interface with adjustable settings for input resolution and detection thresholds

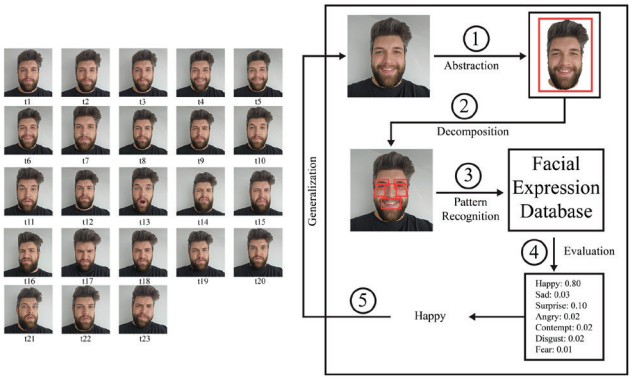


Figure 4 Process breakdown of emotion classification using facial geometry

3.2 Limitations

At the moment following limitations are being investigated:

- only **single-face detection** is implemented; multiple concurrent faces are not processed
- the application does not support **language localization**
- detection accuracy decreases with **low-resolution webcams** or under dim light
- the application relies on **pre-trained generic models** not fine-tuned on custom data.

Additionally, while no user data is sent externally, all inference is **client-side**, meaning devices with very limited computing power may experience lag. Also worth noting is that extreme facial angles, facial accessories (e.g., masks or sunglasses), and rapid motion can reduce recognition fidelity. Adding a warning or guideline for users to face the camera frontally under decent lighting may improve practical usability.

3.3 Potential Applications

With all previously presented data, following potential applications could be implemented:

- **educational settings** to demonstrate AI concepts in real time
- **user studies** in human-computer interaction
- **museums, fairs, and science festivals**
- **personal experiments** in ethical, private facial analysis
- **public kiosks** for emotion-driven art installations or engagement.

Since the application runs entirely in-browser, it can be embedded into other tools or run as a standalone PWA progressive web app (PWA). This makes it highly adaptable for institutions without cloud infrastructure [7, 10]. The tool is also being considered for integration into other research tools that require anonymized demographic and affective data collection without storing personal information. This opens new possibilities for human-computer interaction studies, behavioral research, and human-robotic integration.

3.4 Future Work

Future iterations could introduce new versions with:

- multi-face support with heatmap overlays
- model fine-tuning on local demographic datasets
- voice emotion fusion using WebRTC audio
- exportable logs for offline analysis.

Such features would not only increase its scientific applicability but also allow further customization in artistic, clinical, or assistive use cases.

4 CONCLUSION

The present research introduced Aeye, a browser-based application that use machine learning models for real-time recognition of age, gender, and emotion directly on the client side. The system is entirely based on open-source technology, including face-api.js, TensorFlow.js, and conventional web development tools, involving no backend infrastructure or external data transfer. The modular programming framework and incorporation of interactive UI features provide adaptable testing and real-time examples. Assessments conducted on several platforms demonstrated consistent performance and acceptable accuracy under ideal conditions, rendering the application appropriate for instructional, experimental, and display purposes. At the moment, application detects only one face at a time and depends on pre-trained models. However, it establishes a robust basis for possible future improvements, including multi-facial detection, model refinement, and integration with multimodal systems. These enhancements may extend its application in clinical research, emotion-sensitive user interfaces, or inclusive human-computer interaction systems. Aeye exemplifies the effective utilisation of lightweight AI tools in limited situations through the integration of real-time computer vision, privacy-centric architecture, and cross-platform accessibility. The success of this implementation facilitates further adjustment in both technical and social domains, promoting the responsible and transparent utilisation of AI in human-centered technology design.

5 REFERENCES

[1] Kim, Y. (2021). Deep Learning in Browser-Based Environments. *IEEE Access, 9*, 115024–115035. https://doi.org/10.1109/ACCESS.2021.3105270

[2] Patel, R., & Singh, V. (2020). TensorFlow.js: Machine Learning for the Web. *SoftwareX, 11*, 100465. https://doi.org/10.1016/j.softx.2020.100465

[3] GitHub (2021). face-api.js: JavaScript API for face detection and face recognition in the browser. GitHub. https://github.com/justadudewhohacks/face-api.js

[4] Zhang, L., & Müller, K. (2023). AI Applications in Human-Computer Interaction. *ACM Transactions on Interactive Systems, 12*(3), 221–240. https://doi.org/10.1145/3590345

[5] Chen, L. (2022). Privacy-Aware AI Systems for Edge Devices. *Security and Privacy in AI, 7*(1), 45–60. https://doi.org/10.1097/SPI.22.003

[6] Brown, A., & Green, H. (2023). Offline AI: Opportunities and Constraints. *Journal of Edge Computing, 4*(2), 99–113. https://doi.org/10.1016/j.edge.2023.02.007

[7] Johnson, T., & Novak, M. (2023). Client-Side Neural Networks: A Survey. *Computing Surveys, 55*(6), 1–28. https://doi.org/10.1145/3577723

[8] Nguyen, H. & Tran, Q. (2021). Real-Time Emotion Recognition Using Convolutional Neural Networks. *Neural Processing Letters, 53*(4), 2561–2575. https://doi.org/10.1007/s11063-021-10521-7

[9] Garcia, P., & Liu, J. (2020). Web-Based Face Analytics for Educational Use. *International Journal of Emerging Technologies in Learning, 15*(10), 101–110. https://doi.org/10.3991/ijet.v15i10.13745

[10] Haleem, A., Javaid, M., & Vaishya, R. (2019). Internet of Things and Cloud Computing for Smart Healthcare. *Journal of Clinical Orthopaedics, 4*(3), 204–210. https://doi.org/10.1177/2044531120922501

[11] Pipeline for detecting face and extracting demographic and affective features from https://ai.thestempedia.com (retrieved: 2025-03)

[12] Swathi, A., Kumar, S., T, V., Rani, S., Jain, A., & Kumar, R. (2022). Emotion classification using feature extraction of facial expression. *The 7th International Conference on Trends in Electronics and Informatics (ICOEI2022)*, 283-288. https://doi.org/10.1109/ICTACS56270.2022.9988544

Authors' contacts:

Mihael Holi, system integrator
(Corresponding author)
Juraj Dobrila University of Pula,
Faculty of Engineering,
Alda Negrija 6, 52100 Pula, Croatia
mholi@student.unipu.hr

Sven Maričić, PhD, Full Professor
1) Juraj Dobrila University of Pula,
Faculty of Engineering, Alda Negrija 6, 52100 Pula, Croatia
2) University of Rijeka, Faculty of Medicine, Center for Biomodeling and Innovations in Medicine, Br. Branchetta 20, 51000 Rijeka, Croatia
smaricic@unipu.hr

Irena Jurjevic, MA
King's College London,
Institute of Psychiatry, Psychology & Neuroscience,
Newcomme Street, London, SE1 1UL, The United Kingdom
irena.jurjevic@kcl.ac.uk

Adriana Prunk Drmić, MD, MSc
Thalassotherapia Opatija,
Maršala Tita 188/I, 51410 Opatija, Croatia
adriana.prunk-drmic@tto.hr

Antonija Ružić Baršić, MD, PhD, Assistant Professor
Thalassotherapia Opatija,
Maršala Tita 188/I, 51410 Opatija, Croatia
antonija.ruzic-barsic@tto.hr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250522101813

Preliminary communication
Received: 2025-05-22, Accepted: 2025-05-28

DISPO 4.0 | Simulation-Based Optimization of Stochastic Demand Forecast of Intermittent Material in the Capital Goods Industry

Alexander Schmid*, Marcel Peralt Bonell, Felix Kamhuber, Sebastian Schlund

Abstract: This research introduces a digital planning method tailored to intermittent demand, leveraging simulation-based optimization to select and parameterize forecasting techniques for individual items. Accurate demand forecasting is crucial for optimizing inventory and order quantities, especially in spare parts planning, where demand is often intermittent. Despite the recognized value of digital solutions in this area, practical implementation of optimized forecasting tools for intermittent demands remains limited. To address this gap, the proposed approach offers a practical methodology for systematically selecting and configuring forecasting techniques specifically suited to irregular demand patterns. The method combines a rule-based heuristic with a static simulation of demand time series and metaheuristic-based optimization to calibrate forecasting parameters. The outcome is an automated, item-specific forecast tailored to the characteristics of intermittent demand. The approach is validated through two practical case studies from the capital goods sector, demonstrating its effectiveness in enhancing forecast accuracy and improving intermittent item planning.

Keywords: Croston's method; demand planning; forecasting methods; intermittent demand; parameter optimization; spare parts

1 INTRODUCTION

The optimization of inventories is crucial for companies in order to minimize inventory and capital tie-up costs [1]. It is also essential to minimize scrapping costs and profit losses due to shortages [2], while at the same time ensuring a high level of delivery reliability [1]. Spare parts warehouses in particular play an important role, as they minimize downtimes, can deliver stocked spare parts quickly and avoid downtime costs [1]. High inventories in turn cause costs and account for around a third of all company values [3]. Companies therefore must manage a balancing act between security of availability and inventory costs [2].

In order to remain competitive in a volatile business environment, companies must be able to make timely decisions based on precise demand forecasts derived from existing data [4]. Data and information processing is crucial [5] and algorithms are the key to the efficient utilization of large data volumes [5]. In everyday operational planning, usually only a limited number of mathematical algorithms are applied, as there is a lack of ready-to-use methods for selection and case-specific parameterization [6].

Stochastic demand forecasts [7] - so-called time series models - are classified as quantitative forecasting methods [8]. Time series models are always based on the analysis of historical data in order to predict future demand [8]. Historical material consumption data form the basis for determining future requirements by means of suitable forecasting algorithms [9]. This is referred to as customer-anonymous sales planning [10]. The primary target of this sales planning is to predict future requirements as precisely as possible by incorporating all available information [11] and analyzing demand patterns [12]. Modern forecasting algorithms support companies to optimize planning, manage risks and estimate future material requirements [7].

This practice-oriented research describes the further evolution of a digital planning method for optimized sales forecasting in the capital goods sector, which builds on the authors' existing publications [13-15]. The further development of the method is intended to investigate the use of forecasting algorithms for intermittent material demands in the capital goods sector [14].

2 OBJECTIVES

The aim of this research is the development of an item-accurate planning for items with intermittent requirements in the capital goods industry. This encompasses expertise in forecasting challenges related to intermittent requirements, the use and selection of appropriate forecasting algorithms, their parameterization and optimization, and an evaluation based on real company data [14]. The following research questions are answered by our contribution:

1) Do stochastic time series-based forecasting algorithms specifically designed for intermittent demands achieve better forecasting results for intermittent material demand than general forecasting algorithms designed for constant, trend or seasonal demands (see Fig. 2)?
2) How can a digital, user-friendly planning method be designed to make time series models described in the literature and developed specifically for intermittent material requirements practically usable for the capital goods industry?

3 BACKGROUND STOCHASTIC DEMAND FORECAST

Intermittent demands are characterized by irregular patterns and longer phases without demand [12] as well as sudden demands and represent a material category that is difficult to forecast [16].

The literature contradicts the prevailing assumption in practice that intermittent material demands cannot be reliably predicted [16]. The *state-of-the-art* already provides specialized forecasting algorithms for intermittent material requirements [17]. Intermittent forecasting algorithms were first identified and analyzed, followed by an evaluation of

their potential applications within the operational context of the capital goods sector (see Fig. 1). The methods highlighted in blue represent the most frequently cited techniques, identified through frequency analysis, and were integrated into the developed planning approach [14]. Some of these forecast methods are already in use in IT systems. Nonetheless, a systematic approach to identifying the most appropriate forecasting techniques and their optimal parameterization is currently missing in corporate planning [14, 13].

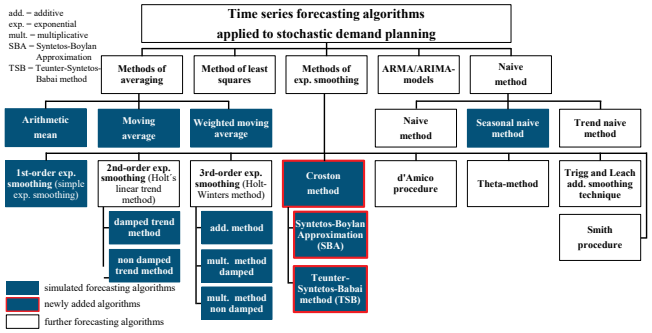


Figure 1 Overview of time series forecasting algorithms applied to stochastic demand planning - (Own representation based on [13, 14, 18, 19])

The consideration of the different types of time series patterns is initially one of the most important aspects in the selection of suitable forecasting methods [20]. When analyzing the development of demand, it is necessary to break down the time series into several components in order to produce accurate forecasts [10] (see Fig. 2).

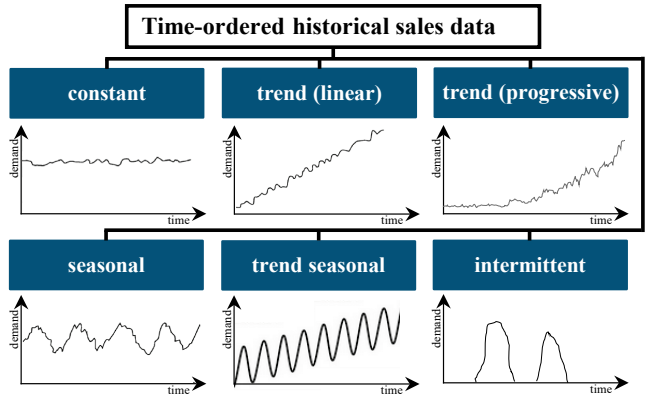


Figure 2 Overview of time-ordered historical sales data [14]

According to the literature review (see Fig. 3), there is no unified expert opinion on the best forecasting technique for specific time series types [14]. Therefore, the optimizing planning method presented here is intended to guide planners to the best possible selection and parameterization of the methods by analyzing the demand curves of all items.

Intermittent material requirements - such as for spare parts in the capital goods industry - require specific forecasts. In contrast to well-researched approaches for fast-moving items, this area has so far received little attention from the literature, which indicates a clear demand for research [21].

Forecasting technique \ Time-ordered historical sales data	constant	trend (linear)	trend (progressive)	seasonal	trend seasonal	intermittent
Saisonal naïve method				●		
Arithmetic mean	●					
Moving average	●					
Weighted moving average	●	◐				◐
1st-order exponential smoothing procedure	●					
2nd-order exponential smoothing procedure (non-damped)	◐	●				
2nd-order exponential smoothing procedure (damped)	◐	●				
3rd-order exponential smoothing procedure (add.)	◐	◐	◐		●	
3rd-order exponential smoothing procedure (mult., non-damped)	◐	◐	◐	●	●	
3rd-order exponential smoothing procedure (mult., damped)	◐	◐	◐	●	●	
Croston method						●
Syntetos-Boylan Approximation (SBA)						●
Teunter-Syntetos-Babai method (TSB)						●

Legend: ● suitable ◐ limited suitable

Figure 3 Relationship between time-ordered historical sales data and forecasting technique according to literature study (Own representation based on [14])

4 CHARACTERIZATION OF THE CASE STUDY

To assess the method's effectiveness, planning-relevant data from two Austrian capital goods manufacturers were used in the evaluation. This topic is highly relevant for the industry, as capital goods are now often ordered with shorter lead times and in smaller volumes. Consequently, manufacturers face rising demands for short delivery times, enhanced flexibility, and precise planning - all under conditions of increasing product variety [14].

	Company 1	Company 2
Sector:	Fittings and valves production	Automotive industry
Employees:	115 (full-time equivalent)	243 (full-time equivalent)
Annual sale:	22,3 m €	89 m €
Customers:	346 in 51 countries	38.800 in 22 countries
Purchasing volume:	11,5 m €	59 m €
Items (active):	12.462 items	17 517 items
Suppliers:	1.780 in 61 countries	1.258 in 28 countries
Period under review:	-36 months sales history	-36 months sales history
Forecast period:	+12 months rolling	+12 months rolling

Figure 4 Case Study – Characterization both Companies [13, 14, 15]

The term "capital goods" covers goods that are not intended for final consumption but are used in the intermediate production of other goods. In other words, goods that produce goods themselves [22].

5 EXTENDED DIGITAL PLANNING METHOD

The digital planning tool was implemented leveraging Microsoft Excel and VBA (*Visual Basic for Applications*) for digitized and (partially) automated requirements calculation for optimized sales planning for items from two Austrian capital goods manufacturers. This tool is intended to make best practices usable and facilitate the traceability of decisions [13, 14]. The tool is modeled to deliver the highest achievable forecast quality by adapting to evolving market dynamics [14].

Using static simulation and a rule-based heuristic as evaluation criteria, this planning method is used to select the best possible forecast procedures for specific items and parameterize them automatically using a *Genetic Algorithm* (GA). The benefits of using quantitative forecasting techniques for semi-automated sales planning within the capital goods sector are examined in two use cases, with detailed discussion in Chapter 6. Fig. 5 below provides an overview of the digital planning method.

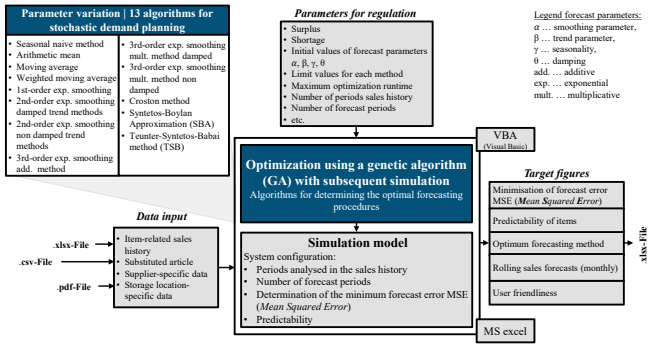


Figure 5 Schematic of the digital planning method based on simulation and optimization (Own representation based on: [23])

This planning method is intended to free the forecasting of future item-specific requirements from the experience of individuals, make best practices generally usable and facilitate the traceability of decisions down to item level based on a simulation model. The modeling aims to achieve the highest possible forecast quality [15]. The planning method involves the following seven steps (see 5.1 to 5.7).

5.1 Data Preparation and Processing

According to *Syntetos*-classification (based on the two coefficients *Average Demand Interval* (*ADI*) and the square of the *Coefficient of Variation* (CV^2)) material requirements can be categorized into different material classifications [19]:

- *Smooth*: steady demand with little fluctuation. [$CV^2 \leq 0.49$ und $ADI \leq 1.32$]
- *Intermittent:* irregular demand in which the order quantities remain largely constant. [$CV^2 \leq 0.49$ und $ADI > 1.32$]
- *Erratic*: high variability in frequency and order quantity. [$CV^2 > 0.49$ und $ADI \leq 1.32$]
- *Lumpy*: infrequent, irregular orders with strongly fluctuating quantities. [$CV^2 > 0.49$ und $ADI > 1.32$]

Table 1 Material classification according to *Syntetos* for both companies

Syntetos material classification	Company 1		Company 2	
	# Items	Percent	# Items	Percent
Lumpy	6.918	55,6%	10.798	61,6%
Erratic	623	5,0%	585	3,3%
Smooth	353	2,8%	817	4,7%
Intermittent	1	0,0%	12	0,1%
Not classifiable	4.567	36,6%	5.305	30,3%
Total:	12.462	100,0%	17.517	100,0%

The classification reveals that most items in both companies are assigned to the *Lumpy* category. This underlines the challenges associated with the high variance and irregular demand intervals in demand forecasting.

The first step is to prepare the data. This process includes a standardized reference to historical data records, the identification of missing values, the filtering of item-specific time series and other pre-processing tasks [11]. Different data sources are imported from the IT systems and processed automatically [15].

5.2 Configuration

This second step determines the parameterization and initial configuration of the individual algorithms and takes into account the setting of the parameters relevant to forecasting and optimization (surplus, shortage, initial values of forecasting parameters, limit values etc.) [15].

In the case study, 12 monthly periods were forecast using 36 monthly periods from the sales history. In a separate worksheet for each forecasting method, the monthly quantities per item are saved as the basis for the following steps [15].

5.3 Calculation

As part of the calculation or optimization of the forecast error (MSE), the sales history for all 13 methods is transferred to the calculation item by item. All forecast methods (see Fig. 1) are calculated for each item and the item-specific forecast result and the corresponding forecast error are saved in the worksheet for each method [15].

The forecast algorithms *Croston method*, *Syntetos-Boylan Approximation* and *Teunter-Syntetos-Babai method*, which were incorporated during the further development of the simulation-based optimization (see Fig. 5) for intermittent demand, are briefly described below.

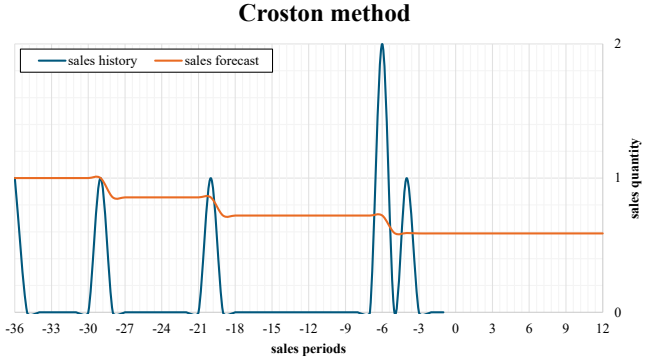


Figure 6 Sales history and sales forecast (*Croston method*: item handle)

The method named after *Croston* (see Fig. 6) is the established approach for forecasting intermittent material requirements and is still recognized as the standard method in this area today [24]. It was the first method to estimate the quantity required and the interval between the requirements separately and enabled a much more accurate forecast compared to the *Moving average* [25]. Over time, various extensions of the *Croston method* have been developed. These include the *Syntetos-Boylan-Approximation* (SBA),

which reduces systematic overestimates of the classic *Croston method*, see Fig. 7 [19]. Another further development is the *Teunter-Syntetos-Babai* (TSB) method, which combines the strengths of the SBA- and *Croston-method*, see Fig. 8 [18]. In addition to frequency and quantity, the *TSB-method* also considers the probability of future demand, thereby improving the accuracy of the forecast.

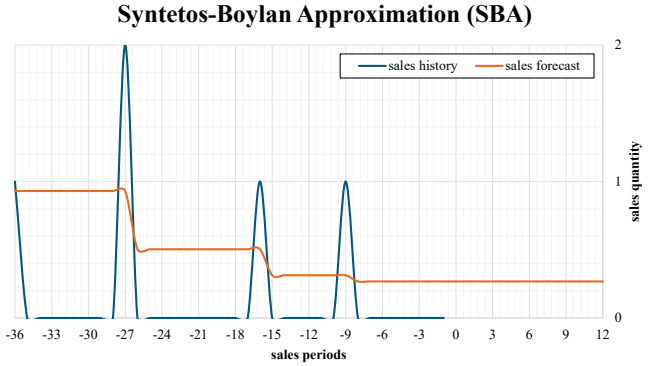


Figure 7 Sales history and sales forecast (*SBA method*: item threadbolts)

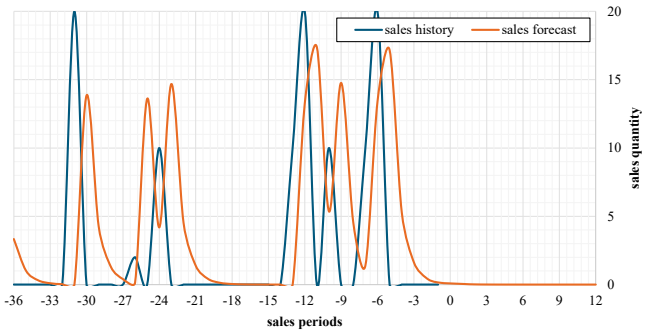


Figure 8 Sales history and sales forecast (*TSB method*: item sight glass)

Initialization of the solution uses optimized parameters: $\alpha = 0.2$ (smoothing), $\beta = 0.5$ (trend), $\gamma = 0.5$ (seasonality), and $\theta = 0.8$ (damping) [15].

5.4 Optimization

The stochastic, combinatorial optimization of the forecast parameters (α = smoothing, β = trend parameter, γ = seasonality, θ = damping) using GA is performed for the *Methods of exp. smoothing* (see Fig. 1). The other methods don't provide any forecast parameters to be optimized. MS-Excel GA is applied for parameter optimization. The GA minimizes the target function represented by the forecast error, transferring the variable cells and the constraints, including the associated settings for the algorithm, for example the runtime and termination criteria [15]. After the preparations, the forecast is calculated as follows:

1) The forecast parameter(s) is/are optimized using a *Genetic Algorithm* (GA).
2) All results (forecast values, parameters etc.) are summarized in a result sheet.
3) Identification of "non-predictable items".
4) The forecast yielding the minimal error is recorded in the results sheet.
5) The results are being exported.

The developed digital planning method enables the calculation of up to 13 forecasting procedures.

5.5 Simulation Parameterisation

The evaluation of the optimal item-specific simulation-based optimization (see Fig. 5) is based on the optimized forecast error [14]. Our approach for optimized demand planning allows one of the following ten standard evaluation methods (*MSE*, *MAD*, *MAPE*, *standard deviation*, *WAPE*, *MASE*, *BIAS*, *TS*, *rMAD* and *rMSE*) to be used to parameterize the simulation.

5.6 Simulation

In practice, the evaluation of the output of real systems is often executed using simulation [26]. The main objective is to determine the forecast quality, which compares the forecast values calculated using the corresponding method with the historical consumption values that have occurred, to be able to draw conclusions about the respective forecast quality. The *MSE* (*Mean Squared Error*) was chosen for the evaluation, as squaring gives greater weight to larger forecast deviations. The *MSE* is determined as follows [10, 12]:

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_{t+h|t})^2, \quad (1)$$

where $\hat{y}_{t+h|t}$ - forecast value for period t, y_t - real demand value, n - number of samples of the period.

The simulation compares all implemented (13) forecast methods for each item and chooses the one with the lowest *Mean Squared Error* per item (see Tab. 2).

The new methods (*Croston*-, *SBA*- and *TSB-method*) for predicting intermittent demand were identified as optimal according to *MSE* in 9.8% of the items in company 1 and 14.8% of the items in company 2. Company 2 tends to use more new methods than company 1, which is plausible as the data collected there mainly relate to spare parts with predominantly intermittent patterns. The new methods were developed for precisely this type of demand pattern. Only a small proportion of spare parts remain *non-predictable*.

The optimization potential of the respective forecast error is ~48% and ~38% on average based on an evaluation for the data set of the two case studies. However, it should be noted that the forecast parameters (α, β, γ, θ) are only optimized in nine out of 13 methods. The nine parameter-optimized methods are recommended for about ~57% or ~70% of the items.

Table 2 Results of the item-specific simulation-based optimization

Implemented algorithms for stochastic demand planning	Company 1		Company 2	
	# Item	Percent	# Item	Percent
Seasonal naive method	230	1,8%	580	3,3%
Arithmetic mean	60	0,5%	188	1,1%
Moving average	3.336	26,7%	3.139	17,9%
Weighted moving average	1.454	11,7%	1.373	7,8%
1st-order exp. smoothing procedure	226	1,8%	325	1,9%
2nd-order exp. smoothing procedure (non-damped)	1.056	8,5%	1.579	9,0%
2nd-order exp. smoothing procedure (damped)	570	4,6%	949	5,4%
3rd-order exp. smoothing procedure (add.)	517	4,2%	2.158	12,3%
3rd-order exp. smoothing procedure (mult., non-damped)	2.843	22,8%	4.198	24,0%
3rd-order exp. smoothing procedure (mult., damped)	615	4,9%	391	2,2%
Croston method	523	4,2%	1.294	7,4%
Syntetos-Boylan Approximation (SBA)	278	2,2%	545	3,1%
Teunter-Syntetos-Babai method (TSB)	424	3,4%	747	4,3%
Non-predictable items	330	2,7%	51	0,3%
Total	12.462	100%	17.517	100%

5.7 Optimization Results

The use of the digital planning method for optimized demand forecast shows that complex *Methods of exp. Smoothing* (see Fig. 1), optimally parameterized, often achieve better results (see Fig. 9) than the *Seasonal naive method* often used in practice, in other words, updating the demand of previous periods. The application of the *Genetic Algorithm* (GA) for the parameter optimization of the *Methods of exp. smoothing* enables an average reduction of ~48% or ~38% of the MSE depending on the company.

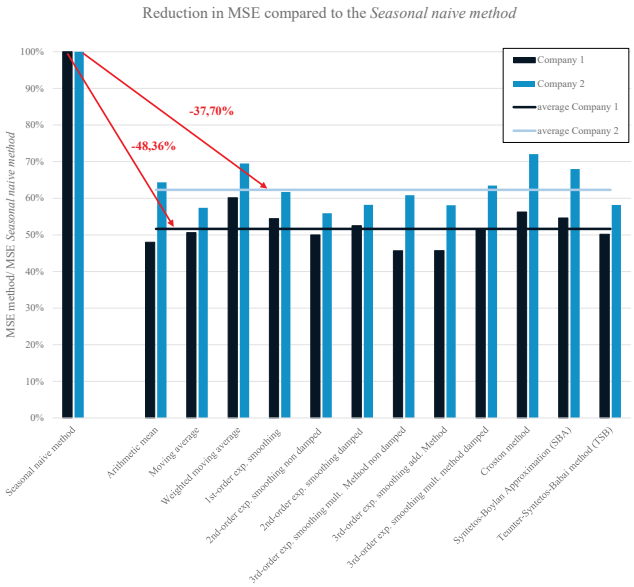


Figure 9 Reduction in MSE compared to the Seasonal naive method

Tab. 2 shows that there is no single optimal forecasting method for the capital goods industry, but that optimal algorithms are emerging for specific items. However, this only represents a snapshot of the forecast calculation for two companies. The optimal forecasting methods change over several planning horizons, as sales figures develop dynamically. Only around ~32% of the items retain the same best forecasting method over three forecasting periods (months).

At the end, an item-specific output of the optimization results, adapted to the case study, is generated.

6 CONCLUSIONS

In practice, the consumption of the past 12 months is often extrapolated as requirements for the future 12 months (*Seasonal naive method*). The digital planning method presented here was able to achieve a reduction in the MSE of ~38 % - ~48 % and thus a significant improvement in forecasting accuracy by using a mix of 13 forecasting algorithms in two case studies (see Fig. 9). The data analysis (see Chapter 5.1) shows that the majority of the items (> 55 %) have a *Lumpy* classification - rare, irregular orders with strongly fluctuating quantities - which cannot be reliably forecast using simple forecasting methods such as the *Seasonal naive method*.

The following answers can be given to the research question posed in Chapter 2:

Question 1: The newly implemented forecasting methods specially designed for intermittent demand patterns (*Croston-*, *SBA-*, *TSB-method*) were able to reduce the MSE by up to ~11 % and ~9 % on average and thus increase the forecast quality accordingly. Since the majority of the items are classified as *Lumpy* (see Chapter 5.1), no general superiority of new forecasting algorithms over classic algorithms can be determined. Nevertheless, the higher forecasting accuracy shows its advantage for intermittent requirements.

Question 2: A digital planning method for the automatic application of established forecasting methods was successfully developed. By linking simulation-based approaches, heuristic parameterization and the use of an appropriately parameterized GA, it was possible to realize an item-accurate, practical material requirements planning approach.

Overall, the digital planning approach for intermittent demand is convincing due to its high specificity (item-specific forecast), comprehensibility and empirical foundation. This is offset by an increased computational effort, a dependence on static simulations and a certain sensitivity to data quality.

As an outlook for future research, higher sophisticated forecasting methods with *Machine Learning* (ML) could be developed and their opportunities and risks evaluated.

Acknowledgements

The authors acknowledge financial support from the *Austrian Research Promotion Agency* (FFG), project 872787, and express their gratitude to the case study partners for their collaboration.

7 REFERENCES

[1] Ball, M., Barnhart, C., Dresner, M., Hansen, M., Neels, K., Odoni, A. R., Peterson, E., Sherry, L., Trani, A., & Zou, B. (2010). Total delay impact study: a comprehensive assessment of the costs and impacts of flight delay in the United States

[2] Claus, T., Herrmann, F., & Manitz, M. (2015). *Produktionsplanung und -steuerung: Forschungsansätze, Methoden und deren Anwendungen*. Springer Berlin Heidelberg. (in German) https://doi.org/10.1007/978-3-662-43542-7

[3] Gioia, D. G., Felizardo, L. K., & Brandimarte, P. (2023). Simulation-based inventory management of perishable products via linear discrete choice models. *Computers & Operations Research, 157*, 106270. https://doi.org/10.1016/j.cor.2023.106270

[4] Kolade, O. J. (2019). Demand Forecasting and Measuring Forecast Accuracy in a Pharmacy, Universität Danubius. *Oeconomica, Audoe, 15*(3), 157-169.

[5] Barton, T., Müller, C., & Seel, C. (2018). *Digitalisierung in Unternehmen, Von den theoretischen Ansätzen zur praktischen Umsetzung*, Springer, Worms, Wildau, Landshut (in German) https://doi.org/10.1007/978-3-658-22773-9

[6] Wischmann, S., & Hartmann, E. A., (2018). *Zukunft der Arbeit, Eine praxisnahe Betrachtung*, 1st edition, Springer Vieweg, Berlin. (in German) https://doi.org/10.1007/978-3-662-49266-6

[7] Rostami-Tabar, B., Greene, T., Shmueli, G., & Hyndman, R. J. (2025). Towards Socially Responsible Forecasting: Identifying and Typifying Forecasting Harms, 1-43.

[8] Gasparian, M. S., Karmanov, M. V., Kiseleva, I. A., Kuznetsov, V. I., & Sadovnikova, N. A. (2018). Modeling of the demand forecasting. *International Journal of Civil Engineering (IJCIET), 9*(12), 163-173.

[9] Pfohl, H.-C. (2018). Logistiksysteme, Betriebswirtschaftliche Grundlagen, Springer, Darmstadt. (in German)

[10] Schuh, G., & Schmidt, C. (2014). *Produktionsmanagement - Handbuch Produktion und Management 5*. 2nd edition, Springer, Aachen. (in German) https://doi.org/10.1007/978-3-642-54288-6

[11] Hyndman, R. J., & Athanasopoulos, G. (2021). Forecasting: Principles and Practice: A comprehensive introduction to the latest forecasting methods using R. Learn to improve your forecast accuracy using dozens of real data examples, 3rd edition. OTexts, Melbourne. https://otexts.com/fpp3/

[12] Meyer, J. C., Sander, U., & Wetzchewald, P. (2019). *Bestände senken, Lieferservice steigern- Ansatzpunkt Bestandsmanagement*, 3. Auflage, FIR an der RWTH Aachen. (in German)

[13] Schmid, A., Sobottka, T., Luther, S., & Sihn, W. (2022). Demand Planning Falcon – Zielgenaue, stochastische Bedarfsvorhersagen mit einer neu entwickelten digitalen Planungsmethode. *Industrie 4.0 Management, 2022*, 47-50. (in German) https://doi.org/10.30844/IM_22-6_47-50

[14] Schmid, A., Kamhuber, F., Sobottka, T., & Sihn, W. (2022). DISPO 4.0 | Simulation-based optimization of stochastic demand calculation in consumption-based material planning in the capital goods industry. *Tehički glasnik, 16*(3), 328-335. https://doi.org/10.31803/tg-20220504151004

[15] Schmid, A., Kamhuber, F., Sobottka, T., & Sihn, W. (2022). DISPO 4.0 - Simulationsgestützte Absatzprognoseoptimierung in der Investitionsgüterindustrie. *Proceedings Langbeiträge ASIM SST 2022, ARGESIM Report 20*, 73-80. (in German) https://doi.org/10.11128/arep.20.a2017

[16] Pinçe, Ç., Turrini, L., & Meissner, J. (2021). Intermittent demand forecasting for spare parts: A critical review. *Omega, 105*, 102513. https://doi.org/10.1016/j.omega.2021.102513

[17] Babaveisi, V., Teimoury, E., Gholamian, M. R., & Rostami-Tabar, B. (2023). Integrated demand forecasting and planning model for repairable spare part: An empirical investigation. *International Journal of Production Research, 61*(20), 6791-6807. https://doi.org/10.1080/00207543.2022.2137596

[18] Teunter, R. H., Syntetos, A. A., & Zied Babai, M. (2011). Intermittent demand: Linking forecasting to inventory obsolescence. *European Journal of Operational Research, 214*(3), 606-615. https://doi.org/10.1016/j.ejor.2011.05.018

[19] Syntetos, A. A., Boylan, J. E., & Croston, J. D. (2005). On the categorization of demand patterns. *Journal of the Operational Research Society, 56*(5), 495-503. https://doi.org/10.1057/palgrave.jors.2601841

[20] Hanke, J. E., & Wichern, D. W. (2013). *Business forecasting*, 9th edition. Prentice Hall/Pearson Education International.

[21] Hasan, N., Ahmed, N., & Ali, S. M., (2024). Improving sporadic demand forecasting using a modified k-nearest neighbor framework. *Engineering Applications of Artificial Intelligence 129*, 107633. https://doi.org/10.1016/j.engappai.2023.107633

[22] Hache, E., & Sander, H. (1997). *Expert-Lexikon Bilanzierung - 1333 Begriffe, systematisch und von A - Z praxisnah erläutert*. Expert-Verl., Renningen-Malmsheim. (in German)

[23] März, L., Krug, W., Rose, O., & Weigert, G. (2011). *Simulation und Optimierung in Produktion und Logistik: Praxisorientierter Leitfaden mit Fallbeispielen*. Springer, Heidelberg, Lindau, Dresden. (in German) https://doi.org/10.1007/978-3-642-14536-0

[24] Singh, A. K., Simha, J. B., & Agarwal, R. (2024). Prediction of Intermittent Demand Occurrence using Machine Learning. EAI Endorsed Trans IoT, 10. https://doi.org/10.4108/eetiot.5381

[25] Croston, J. D. (1972). Forecasting and stock control for intermittent demands. *Operational Research Quarterly (1970-1977) 23*(3), 289. https://doi.org/10.2307/3007885

[26] Weinzierl, T. (2024). *Grundlagen des parallelen wissenschaftlichen Rechnens: Ein erster Leitfaden zu numerischen Konzepten und Programmiermethoden*, 1st edition. Springer Vieweg, Cham. (in German) https://doi.org/10.1007/978-3-031-49082-8

Authors' contacts:

Alexander Schmid, Dipl.-Ing.
(Corresponding author)
FHWien der WKW,
University of Applied Sciences for Management & Communication,
Währinger Gürtel 97, 1180 Vienna, Austria
T: +43 1 476 77- 0
M: +43 690 40 476-167
alexander.schmid@fh-wien.ac.at

Marcel Peralt Bonell, Dipl.-Ing. BSc
Vienna University of Technology,
Karlsplatz 13, 1040 Vienna, Austria
marcel.peralt-bonell@tuwien.ac.at

Felix Kamhuber, Dr. techn. Dipl.-Ing.
Fraunhofer Austria Research GmbH,
Theresianumgasse 7, 1040 Vienna, Austria
felix.kamhuber@fraunhofer.at

Sebastian Schlund, Univ.-Prof. Dr.-Ing. Dipl.-Ing.
Fraunhofer Austria Research GmbH,
Theresianumgasse 7, 1040 Vienna, Austria
sebastian.schlund@fraunhofer.at

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250612184533

Preliminary communication
Received: 2025-06-12, Accepted: 2025-06-23

Systematic Literature Review about the Possibilities for Optimizing Container Loading Tasks in the Case of Intermodal Container Terminals

Péter Kováts*, Róbert Skapinyecz, Péter Tamás

Abstract: The paper examines the optimization possibilities of container loading tasks regarding intermodal terminals. The research carried out is of great importance, as the role of these terminals is expected to increase in the future based on the forecasts. The research presents the operational characteristics of intermodal terminals well as applicable loading procedures and tools. It analyses current optimization approaches, including logistics and technological developments, using a systematic literature review (SLR). The article explores additional optimization opportunities offered by artificial intelligence, 5G networks and simulation models, which can contribute to more efficient loading tasks and provide higher-quality logistics services. The presented results represent a valuable contribution to the development of railway logistics systems from both a practical and a theoretical point of view.

Keywords: container optimization; container terminals; intermodal container terminals; systematic literature review

1 INTRODUCTION

This systematic literature review examines the methods applied so far for optimizing container loading, including simulation models, mathematical programming, and heuristic algorithms. Our aim is to provide a comprehensive overview of current approaches, highlighting their strengths, limitations, and interrelationships. Through our analysis, we seek to identify and explore research gaps and potential future directions for further improving container loading efficiency. This study may contribute to the development and implementation of new optimization solutions, as well as to future research topics.

2 BASIC CHARACTERISTICS OF INTERMODAL CONTAINER TERMINALS

Intermodal container terminals are essential components of international transportation and global supply chains, as they ensure seamless transitions between various modes of transport, such as rail, road, and maritime shipping. The efficiency and optimization of container loading are becoming increasingly important for improving terminal operations, as they directly affect costs, transit times, and service quality. The goal of container loading is to ensure fast, cost-effective, and safe handling of containers between different modes of transportation. In addition to transshipment, terminals are also involved in warehousing, customs clearance, repackaging, and the documentation of these processes. The transfer of goods typically occurs between road, rail, and waterborne vehicles, but air transport can also be part of the process. Terminals play a vital role in freight forwarding, enabling the most cost-efficient transportation of goods. They are key elements of the global supply chain, facilitating the efficient movement of containers.

The role of container terminals is continuously growing due to increasing freight transport demands and stricter environmental regulations. Beyond transshipment, they offer various logistics services to enhance efficiency and provide high-quality customer service. Temporary storage and bonded warehouses offer financial security for customers while enabling faster customs clearance [1-4].

The two most common types of intermodal container terminals are rail terminals and container terminals, generally classified based on the number of available modes of transport.

The main categories are:

- Unimodal: One primary mode of transport, typically road.
- Bimodal: Two transport modes, usually road and rail.
- Multimodal: Three or more transport modes.

Beyond these types, the evaluation of terminals can also consider the following key characteristics:

- Ownership structure: Public, private, or mixed.
- Size: The facility's area (e.g., in hectares).
- Cargo handling capacity: The capacity of transshipment equipment and storage areas.
- Traffic volume: The amount of goods handled (in tons, combined transport units), the number of customers, transport routes, and cargo composition [3, 6].

3 SYSTEMATIC LITERATURE REVIEW

The research followed a structured set of steps (Fig. 1) to achieve its objectives, guaranteeing a high-quality study and preventing the loss of scientific information [7-9].

This analysis includes an extensive search and evaluation of publications in English.

The process of conducting a systematic literature analysis involves the following steps:

1) Motivation
2) Define research questions
3) Define keywords
4) Literature analysis
5) Selection of relevant literature
6) Systematic review
7) Define research gap.

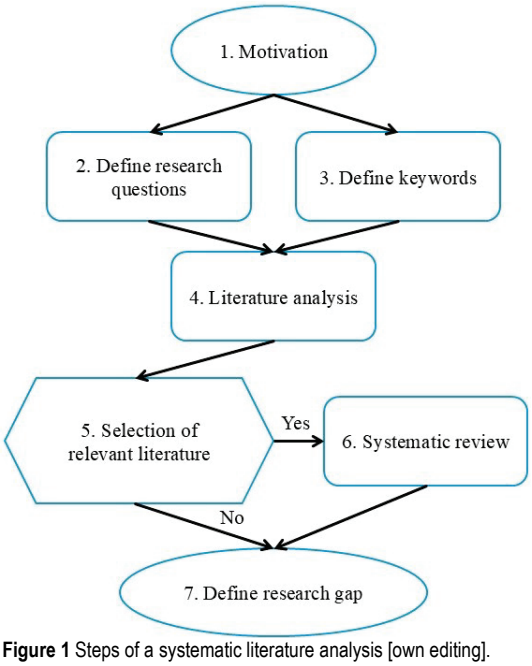


Figure 1 Steps of a systematic literature analysis [own editing].

3.1 Motivation

The purpose of this systematic literature review is to collect and analyse various studies and methods related to the optimization of container loading in intermodal container terminals by using Scopus and google scholar database.

The acquired knowledge, including methods, opportunities, and existing ideas, will support future research in the field and help identify a research gap.

3.2 Research Questions

The aim of the research is to determine what optimization methods are used to improve container loading in intermodal container terminals. The key questions posed in this systematic review are as follows:

- What mathematical models, algorithms, and simulations can be applied to optimize container loading?
- What benefits do the applied optimization methods offer, and how do they impact costs, time savings, and terminal performance?
- What scientific research has been published in this field between the 2000-2024.

3.3 Define Keywords

The primary objective of this step is to define the keywords and their respective combinations. Academic papers were selected through the use of "AND" logical combinations. The keywords are Container, container terminal, optimization, loading, unloading, method.

3.4 Literature Analysis

In the fourth step, the actual literature analysis is conducted. Two academic research databases were utilized in the current study. We used the Scopus and Google Scholar databases during the research; however, the final step was not performed on Google Scholar.

3.5 Selection of Relevant literature

In the fifth step, the relevant literature is selected based on the findings from the literature review. After reviewing the results, and selecting and reading the pertinent publications, the focus shifts to defining the primary research direction.

During a systematic literature search, the areas are narrowed down based on the topic, research area and language. In the current research, publications corresponding to the following categories were selected: research article, review article, chapter of book and minireview. The rest were not selected, and duplicates were removed.

The results of the analysis are shown in Fig. 2.

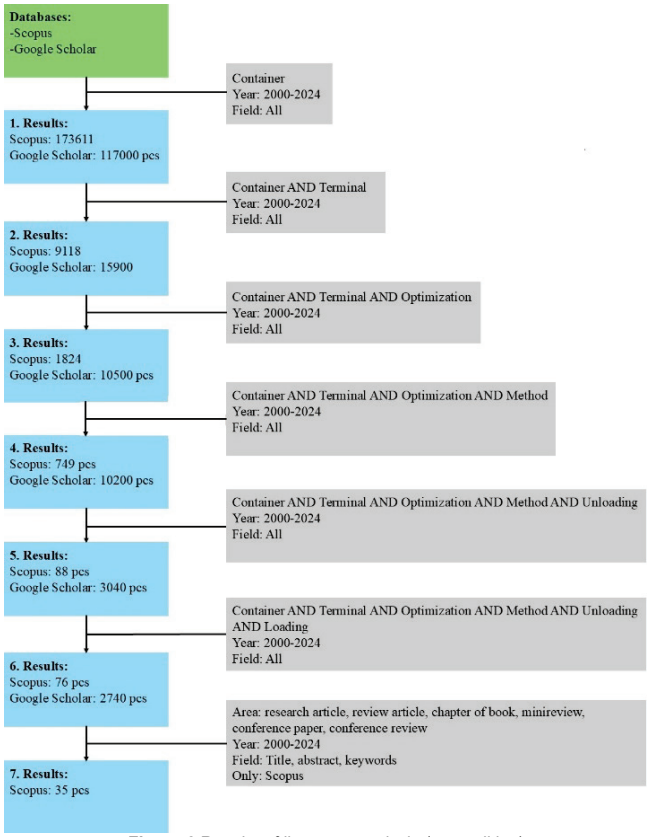


Figure 2 Results of literature analysis (own editing)

3.6 Systematic Review and Define Research Gap

Following the systematic review, the relevant literature can be summarized in the following chapter 3.2, 3.3 and 3.4.

Upon reviewing all relevant publications and abstracts from the search results, a research gap can be systematically identified, as demonstrated in the present study.

4 RESEARCHED METHODS

Research on optimizing container loading in intermodal container terminals spans a wide range of methods. Key approaches include:

Mathematical Programming: Techniques such as linear and nonlinear programming help determine optimal resource allocation. Network analysis and graph theory also support the design of efficient handling strategies.

Simulation Modelling: Simulations, including agent-based models and Monte Carlo methods, help analyze container flow dynamics and test different operational scenarios for better decision-making.

Heuristic Algorithms: Methods like genetic algorithms, ant colony optimization, and tabu search offer fast, adaptable solutions for complex terminal operations, including layout planning and resource management.

Multi-Criteria Decision Support: Optimization often balances factors such as cost, time, and environmental impact. MCDA tools help identify solutions that align with multiple objectives.

Machine Learning: Recent trends show increasing use of techniques like reinforcement learning, neural networks, and predictive analytics. These tools enhance decisions by identifying operational patterns and adapting to dynamic terminal conditions.

The next section provides an overview of integrated scheduling methods for container handling and terminal operations, including the coordination of cranes, AGVs, and other equipment. It highlights two-stage modelling techniques, priority rule-based heuristics.

4.1 Equipment Coordination Strategies

Integrated Scheduling: Coordinating the operations of quay cranes, AGVs, and yard cranes is essential to minimize total completion time (makespan) and enhance terminal performance [10]. Integrated scheduling models often employ two-stage programming and rolling horizon procedures to optimize the routing and synchronization of equipment.

Priority Rule-Based Heuristics: These heuristics, when combined with genetic algorithms, are applied in general scheduling models to efficiently handle real-world challenges, such as those encountered in the Port of Hamburg.

It is important to note that gauge-changing terminals differ from other container terminals due to the necessity of transshipment from train to train for gauge conversion. However, the methods and container loading techniques should not be restricted solely to this aspect, as processes like container loading and scheduling can be applied universally.

The following section highlights that optimizing container loading reduces energy consumption and operational time. Genetic algorithms and simulation-optimization methods shorten loading time, while mixed-integer programming and adaptive heuristics enable efficient equipment scheduling, reducing vessel berthing and dwell times.

4.2 Container Loading and Handling

Optimizing container loading and handling is crucial for reducing energy consumption and operational time. A genetic algorithm can determine the optimal loading sequence, significantly reducing handling time, rearrangements, and energy consumption in rail-road intermodal container terminals. Similarly, minimizing container rearrangement through simulation-optimization methods can reduce operational costs by up to 7% [11-15].

For efficient container handling, terminals must coordinate and schedule various equipment such as transport vehicles, storage cranes, quay cranes, and automated guided vehicles (AGVs). Mixed-integer programming (MIP) models can optimize these operations, minimizing vessel berthing times and train dwell times. For large-scale problems, adaptive heuristic algorithms, which adjust the parameters of genetic algorithms, provide near-optimal solutions, improving both berthing and computation times [16-21].

The following section presents methods for modelling and solving container loading schedules. Mixed-integer programming enforces strict constraints, while heuristic approaches such as genetic algorithms provide fast, near-optimal solutions. Machine learning further refines these strategies by leveraging historical data, enhancing both flexibility and efficiency.

4.3 Key Approaches and Models

- Mixed-Integer Programming (MIP) Models: These models mathematically formulate scheduling challenges while considering constraints related to the coordinated operation of various equipment. They are effective for small-scale problems, but heuristic methods are necessary for larger cases [21-26].
- Heuristic Algorithms: Various heuristic algorithms, such as multi-layer genetic algorithms (MLGA) and hybrid genetic-cuckoo optimization algorithms (HGCOA), help overcome computational difficulties and provide near-optimal solutions for large-scale problems. These algorithms dynamically adjust parameters to improve efficiency and solution quality [21-27].
- Simulated Annealing and Genetic Algorithms: Unique simulated annealing (SA) and genetic algorithms (GA) are used to balance solution quality and computational time, particularly in coordinated scheduling models that consider all terminal equipment collectively.
- Key optimization techniques and machine learning-driven approaches to solving the container relocation problem (CRP). Exact algorithms like Iterative Branch-and-Bound (IB&B) and heuristic methods such as Machine Learning-Driven Beam Search (MLBS). The integration of machine learning enhances decision-making efficiency, reducing computational complexity and improving optimization performance. Machine learning techniques, such as neural networks and random forests, assist in branch pruning and heuristic guidance, improving solution accuracy. These models analyse patterns from past data to refine optimization strategies dynamically [28, 29].

The unified application of methods from integrated scheduling to heuristic and machine learning based algorithms enables comprehensive optimization of container loading by reducing energy consumption, waiting times, and reshuffling operations. Mixed-integer programming ensures precise scheduling, while fast, near-optimal heuristics and metaheuristics support large-scale, real-time decision making. Machine learning–driven approaches continuously

refine these strategies, enhancing flexibility and overall terminal performance.

4.4 Detailed Reviewed Articles

Zheng et al. (2023) propose an integrated scheduling model to optimize container handling systems at container terminals, with a particular focus on green, low-carbon operations. As a first step, the study develops a Mixed-Integer Linear Programming (MILP) model that simultaneously considers the three main handling equipment types at container terminals: quay cranes (QC), yard trucks (YT), and yard cranes (YC). The model aims to minimize ship turnaround time and reduce energy consumption, ultimately improving terminal efficiency.

Since the MILP method becomes computationally intractable for larger problem sizes, the authors propose a genetic algorithm (GA). This approach utilizes a three-dimensional chromosome representation, encoding the QC loading sequence along with YT and YC assignments. The study details the algorithm's construction, including specialized order crossover, two-point crossover, and mutation operators, as well as a novel heuristic mutation procedure that facilitates finding high-quality solutions while avoiding local optima traps.

Computational experiments conducted on 25 benchmark instances validate the effectiveness of the proposed method, particularly for large-scale problems where traditional MILP solutions are either too time-consuming or entirely infeasible. Results indicate that the GA not only significantly reduces computation time but also produces competitive or even superior solutions for container terminal scheduling problems. This contributes substantially to green, energy-efficient operations at intermodal container terminals.

This study provides a valuable contribution to the literature on integrated scheduling problems in container terminals, demonstrating how traditional mathematical programming models can be adapted and combined with modern heuristic algorithms to tackle complex logistics challenges [30].

Homayouni et al. (2014) propose a genetic algorithm for optimizing the integrated scheduling of cranes, automated guided vehicles (AGVs), and split-platform automated storage/retrieval systems (SP-AS/RS) in automated container terminals. The study addresses operational challenges arising from increasing container throughput and the demand for faster vessel turnaround, emphasizing that coordinated scheduling offers significant efficiency gains over non-integrated approaches.

The paper first introduces the operating environment of automated container terminals, followed by a review of prior research, most of which focused only on partial scheduling for individual equipment types. In contrast, this study presents a fully integrated modelling approach, optimizing the sequence of crane operations, the assignment of AGVs to tasks, and the scheduling of storage platforms simultaneously.

The proposed genetic algorithm employs specialized operators, including crossover and mutation mechanisms designed to preserve the predefined sequence constraints of crane operations, ensuring solutions adhere to operational precedence conditions. Performance evaluations on various test cases demonstrate that the GA significantly outperforms traditional MIP models and simulated annealing algorithms, particularly for larger-scale problems. The study also explores the impact of a "stay-in-place" strategy for storage platforms, which reduces platform movement time and delays.

Overall, the work of Homayouni et al. makes a valuable contribution to research on integrated scheduling in automated container terminals, offering a robust metaheuristic approach for multi-equipment coordination optimization and enhancing terminal operational efficiency [31].

The next literature about Real-Time Optimization of the Container Relocation Problem (CRP)

The challenges of the Container Relocation Problem (CRP) arise when retrieving import containers in a predefined order from a container terminal. Often, blocking containers stacked above the target container need to be relocated, making the process non-productive and increasing operational costs. The time-consuming nature of CRP is further complicated by the random arrival of external trucks (ETs), making it difficult to predetermine the exact sequence of container retrievals. Additionally, CRP is classified as an NP-complete problem, which poses significant computational challenges for real-time optimization.

Several approaches have been proposed in the literature to optimize CRP, including integer programming models, exact algorithms such as A and branch-and-bound (B&B), and various heuristics. However, due to the exponential increase in problem complexity, these methods often fail to provide real-time solutions for large-scale CRP instances.

Unlike traditional approaches, Zhou and Zhang (2024) focus on optimizing individual operational rounds within the container terminal. By decomposing the problem into smaller subproblems, they develop a dynamic programming model, which derives a dynamic upper bound for the optimal solution and aims to minimize it. Their approach is built on three heuristic relocation rules:

- LL (Least-relocations-and-Lowest-precedence-gap) Rule: Minimizes future relocations by selecting a destination that preserves the container retrieval order while avoiding additional rehandling operations.
- MSS (Moving Ahead for Sequential Stack) Rule: Proactively relocates containers that may cause future blockages, leading to a more efficient stacking arrangement.
- FSS (Freeing up for Sequential Stack) Rule: Provides an alternative arrangement when only inverted stacking positions are available, reducing future operational costs.

By integrating these rules, the study introduces a new Sequential-Placement First Heuristic (SPFH) algorithm, which iteratively optimizes container retrieval sequences in a real-time operational environment, significantly reducing the number of relocations.

The study evaluates the proposed method on both small- and large-batch cases, demonstrating that SPFH achieves fast execution times and low relocation costs, making it well-suited for modern container terminals operating in real-time.

This study presents an innovative approach to the real-time optimization of the CRP. By leveraging a dynamic programming model and three heuristic rules, the method

successfully minimizes relocation operations and associated costs. The results indicate that SPFH offers a promising alternative to previous computationally intensive solutions, making a valuable contribution to the literature on CRP optimization [32].

The detailed case studies in Section 3.4 validate the effectiveness of the applied models on real terminal data, thereby substantiating the earlier general overviews with concrete metrics. By testing and comparing the novel algorithmic solutions such as the MILP-GA hybrid and the SPFH heuristic in real-world settings, this chapter clearly highlights the advantages of scalability and practical applicability.

5 SUMMARY AND RESULTS

The application of optimization methods in container loading offers several advantages:

- Cost Reduction: The use of simulation and mathematical models can lead to significant cost savings by reducing the necessary resources, such as cranes, loading personnel, and transportation time.
- Time Savings: Optimized container arrangements and reduced transit times increase terminal throughput and result in faster processes.
- Efficient Resource Allocation: The optimal allocation of resources enables terminals to operate more efficiently while ensuring the safe and rapid handling of containers.
- Reduction of Environmental Impact: Considering sustainability aspects, optimization methods can reduce carbon emissions and other environmental impacts.

Each applied optimization method contributes to improving the efficiency of container loading. However, the optimal solution depends on the specific characteristics of the terminal, such as size, the transportation modes used, and available resources. Integrated approaches that combine mathematical models, simulations, and heuristic algorithms often yield the best results, as they can adapt to dynamic environments.

Most studies confirm that optimizing container loading is a key factor in the competitiveness of terminals. However, the effectiveness and efficiency of applied methods depend on the specific characteristics of the terminal, and a combined approach is often required to achieve the best results. One of the biggest challenges is managing the dynamically changing logistics environment. Therefore, future research should focus on sustainability and the development of adaptive algorithms.

Future work should integrate digital-twin frameworks to mirror physical container yards in real time, enabling predictive optimization of stacking crane movements and reducing non-productive operations

5.1 Research Gap

Evidence derived from industry research and a systematic literature review demonstrates that although numerous examples exist for the application of the methods, scarcely any examples can be found that combine logistics simulation with machine learning in relation to the optimization of container loading tasks at container terminals.

Based on the previous, a possible research gap is the combined application of artificial intelligence combined with logistics simulation, supplemented by considering environmental conditions and traffic situations. The specified model could then provide information in a short period of time, for example based on the meteorological and traffic data, to determine optimal scheduling for loading operations, and it could prioritize more important shipments, while also prioritizing critical shipments to minimize delays.

Operators could influence the model's output with specific data to ensure the final results are as precise as possible. The rapid scheduling and optimized arrangement of container loading sequences could enhance terminal loading operations by providing up-to-date information and enabling swift recalculations in the event of unforeseen incidents, thereby reducing potential losses. Research should focus on integrating advanced AI techniques such as deep reinforcement learning to enable adaptive, data-driven loading decisions in real time. Finally, the fusion of human expertise with intelligent decision-support tools will be essential to address complex, uncertain scenarios and ensure robust, practical implementation.

Acknowledgements

Project No. 2023-1.2.4-TÉT-2023-00027 has been supported by the University Research Scholarship Program of the Ministry for Culture and Innovation from the source of the National Research, Development and Innovation Fund (EKÖP - Egyetemi Kutatói Ösztöndíj Program) and funding came from the 2023-1.2.4-TÉT funding scheme.

6 REFERENCES

[1] Woxenius, J., Andersson, E., Bärthel, F., Troche, G., Sommar, R., & Trouvè, J. (2004). *A Swedish intermodal transport service based on line-trains serving freight forwarders*. 10th WCTR '04, Istanbul, 4–8.

[2] Behrends, S., & Flodén, J. (2012). The effect of transhipment costs on the performance of intermodal linetrains. *Logistic Resources*, *4*, 127–136. https://doi.org/10.1007/s12159-012-0066-0

[3] European Court of Auditors. (2023). *Rail freight transport in the EU: Still not on the right track*. https://www.eca.europa.eu/Lists/ECADocuments/SR-2023-08/SR-2023-08_HU.pdf

[4] East West Intermodal Logistics Services. (n.d.). *Terminal*. https://eastwestil.com/terminal/

[5] Záhony-Port. (2024). *Záhony-Port EN 2024* [PDF document].

[6] Bokor, Z. (2005). Az intermodális logisztikai szolgáltatások helyzetének értékelése, fejlesztési lehetőségeinek feltárása. *Logisztika*, *10*(3), 22–64. (in Hungarian)

[7] Matyi, H., & Tamás, P. (2023). An innovative framework for quality assurance in logistics packaging. *Logistics*, *7*(4), 82. https://doi.org/10.3390/logistics7040082

[8] Kamarási, V., & Mogyorósy, G. (2015). Szisztematikus irodalmi áttekintések módszertana és jelentősége. Segítség a diagnosztikus és terápiás döntésekhez. *Orvosi Hetilap*, *156*(38), 1523–1531. (in Hungarian) https://doi.org/10.1556/650.2015.30255

[9] Rana, R. L., Adamashvili, N., & Tricase, C. (2022). The impact of blockchain technology adoption on tourism industry: A

systematic literature review. *Sustainability*, *14*(12), 7383. https://doi.org/10.3390/su14127383

[10] Zheng, Y., Xiao, Y., & Seo, Y. (2014). A tabu search algorithm for simultaneous machine/AGV scheduling problem. *International Journal of Production Research*, *52*, 5748–5763. https://doi.org/10.1080/00207543.2014.910628

[11] Wang, L., & Zhu, X. (2019). Container loading optimization in rail–truck intermodal terminals considering energy consumption. *Sustainability*. https://doi.org/10.3390/su11082383

[12] Dong, C., Boute, R., McKinnon, A., & Verelst, M. (2018). Investigating synchromodality from a supply chain perspective. *Transportation Research Part D: Transport and Environment*, *61*, 42–57. https://doi.org/10.1016/j.trd.2017.05.011

[13] Giusti, R., Iorfida, C., Li, Y., & Menerba, D. (2019). Sustainable and de-stressed international supply-chains through the SYNCHRO-NET approach. *Sustainability*, *11*, 1083. https://doi.org/10.3390/su11041083

[14] Ambra, T., Caris, A., & Macharis, C. (2018). Towards freight transport system unification: Reviewing and combining the advancements in the physical internet and synchromodal transport research. *International Journal of Production Research*, *47*, 1606–1623. https://doi.org/10.1080/00207543.2018.1494392

[15] Zeng, Q., & Yang, Z. (2009). Integrating simulation and optimization to schedule loading operations in container terminals. *Computers & Operations Research*, *36*, 1935–1944. https://doi.org/10.1016/j.cor.2008.06.010

[16] Golias, M., Portal, I., Konur, D., Kaisar, E., & Kolomvos, G. (2014). Robust berth scheduling at marine container terminals via hierarchical optimization. *Computers & Operations Research*, *41*, 412–422. https://doi.org/10.1016/j.cor.2013.07.018

[17] Golias, M., Boile, M., Theofanis, S., & Efstathiou, C. (2010). The berth-scheduling problem. *Transportation Research Record: Journal of the Transportation Research Board*, *2166*, 20–27. https://doi.org/10.3141/2166-03

[18] Bierwirth, C., & Meisel, F. (2015). A follow-up survey of berth allocation and quay crane scheduling problems in container terminals. *European Journal of Operational Research*, *244*, 675–689. https://doi.org/10.1016/j.ejor.2014.12.030

[19] Iris, C., Christensen, J., Pacino, D., & Ropke, S. (2018). Flexible ship loading problem with transfer vehicle assignment and scheduling. *Transportation Research Part B: Methodological*, *111*, 113–134. https://doi.org/10.1016/j.trb.2018.03.009

[20] Chen, L., Bostel, N., Dejax, P., Cai, J., & Xi, L. (2007). A tabu search algorithm for the integrated scheduling problem of container handling systems in a maritime terminal. *European Journal of Operational Research*, *181*(1), 40–58. https://doi.org/10.1016/j.ejor.2006.06.033

[21] Lau, H. Y. K., & Zhao, Y. (2008). Integrated scheduling of handling equipment at automated container terminals. *Annals of Operations Research*, *159*, 373–394. https://doi.org/10.1007/s10479-007-0262-5

[22] Lau, H. Y. K., & Zhao, Y. (2008). Integrated scheduling of handling equipment at automated container terminals. *Annals of Operations Research*, *159*, 373–394. https://doi.org/10.1007/s10479-007-0262-5

[23] Kim, K. H., & Bae, J. W. (1999). A dispatching method for automated guided vehicles to minimize delays of containership operations. *International Journal of Management Science*, *5*(1), 1–25.

[24] Kim, K. H., & Bae, J. W. (2004). A look-ahead dispatching method for automated guided vehicles in automated port container terminal. *Transportation Science*, *38*(2), 224–234. https://doi.org/10.1287/trsc.1030.0082

[25] Luo, J., & Wu, Y. (2020). Scheduling of container-handling equipment during the loading process at an automated container terminal. *Computers & Industrial Engineering*, *149*, 106848. https://doi.org/10.1016/j.cie.2020.106848

[26] Zhang, X., Hong, Z., Xi, H., & Li, J. (2024). Optimizing multiple equipment scheduling for U-shaped automated container terminals considering loading and unloading operations. *Computer-Aided Civil and Infrastructure Engineering*. https://doi.org/10.1111/mice.13275

[27] Ting, C. J., & Wu, K. C. (2017). Optimizing container relocation operations at container yards with beam search. *Transportation Research Part E: Logistics and Transportation Review*, *103*, 17–31. https://doi.org/10.1016/j.tre.2017.04.010

[28] Zhang, C., Guan, H., Yuan, Y., Chen, W., & Wu, T. (2020). Machine learning-driven algorithms for the container relocation problem. *Transportation Research Part B: Methodological*, *139*, 102–131. https://doi.org/10.1016/j.trb.2020.05.017

[29] Lee, J., Choi, J., Koc, G., Sayin, G., & Seo, H. (2024). Incoming container schedule-aware container rearrangement planning based on reinforcement learning in container terminal. In *2024 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)* (pp. 1417–1421). https://doi.org/10.1109/IEEM62345.2024.10857158

[30] Zheng, Y., Xu, M., Wang, Z., & Xiao, Y. (2023). A genetic algorithm for integrated scheduling of container handling systems at container terminals from a low-carbon operations perspective. *Sustainability*. https://doi.org/10.3390/su15076035

[31] Homayouni, S., Tang, S., & Motlagh, O. (2014). A genetic algorithm for optimization of integrated scheduling of cranes, vehicles, and storage platforms at automated container terminals. *Journal of Computational and Applied Mathematics*, *270*, 545–556. https://doi.org/10.1016/j.cam.2013.11.021

[32] Zhou, S., & Zhang, Q. (2024). Real-time batch optimization for the stochastic container relocation problem. *Applied Sciences*. https://doi.org/10.3390/app14062624

Authors' contacts:

Péter Kováts, PhD student and departmental engineer
(Corresponding author)
Institute of Logistics, University of Miskolc,
Egyetem út 1, 3515 Miskolc-Egyetemváros, Hungary
E-mail: peter.kovats@uni-miskolc.hu

Róbert Skapinyecz, associate professor
Institute of Logistics, University of Miskolc,
Egyetem út 1, 3515 Miskolc-Egyetemváros, Hungary
E-mail: robert.skapinyecz@uni-miskolc.hu

Tamás Péter, full professor
Institute of Logistics, University of Miskolc,
Egyetem út 1, 3515 Miskolc-Egyetemváros, Hungary
E-mail: peter.tamas@uni-miskolc.hu

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20240809124430

Subject review
Received: 2024-08-09, Accepted: 2025-09-30

Transient Applications and Insights to Whale Optimization Algorithm

Samiksha Yadav, Ankita, Shweta Jindal, Nishtha Jatana, Charu Gupta, Nitesh Singh Bhati, Dimple Tiwari*

Abstract: Whale Optimization Algorithm (WOA) is meta-heuristic nature-based optimization algorithm inspired from bubble-net hunting strategy behavior worn by humpback whales. The algorithm has been used to resolve problems in many categories. An insight to the available datasets, tools, comparison techniques and recent trends has not been elaborated comprehensively. Therefore, this work is conducted to get more updated and insightful information on how different problems are getting solved with applications, modifications, enhancement and hybridization of WOA with others. The work presents an insight to the recent publication trends, along with details of all the datasets used for testing various WOA applications. In addition, the tools used and the probable comparison techniques have all been listed along with the WOA version used to solve the problem. This would help the researchers in using the appropriate form of WOA for solving their problems. As per the findings, Whale Optimization Algorithm has been successfully employed in many areas of engineering to find optimal solutions.

Keywords: Applications of Whale Optimization Algorithm; Improved Whale Optimization Algorithm; Modified Whale Optimization Algorithm; Survey/Review; Whale Optimization Algorithm

1 INTRODUCTION

Whale Optimization Algorithm (WOA) [1] is a search-based nature-inspired optimization algorithm adopted from bubble-net chasing strategy worn by humpback whales. It was developed in 2016 by Mirjalili & Lewis [1]. WOA is basically a clustering-based clever calculation put forward for sustained optimization issues. It has demonstrated to show prevalent execution besides other meta-heuristic methods. It is simple in execution and powerful so that it persists comparable to diverse nature-inspired algorithms. The calculation requires lesser control parameters nearly, as it has been a solitary parameter (time interval) must be fine-tuned. In WOA, the populace of humpback whales appears via a multi-dimensional exploration region for nourishment as regarded within Fig. 1. The regions owned by humpback whales are explored to diverse factors, whereas the take away amongst the humpback whale entities and the nourishment compares to esteem of goal fetched.

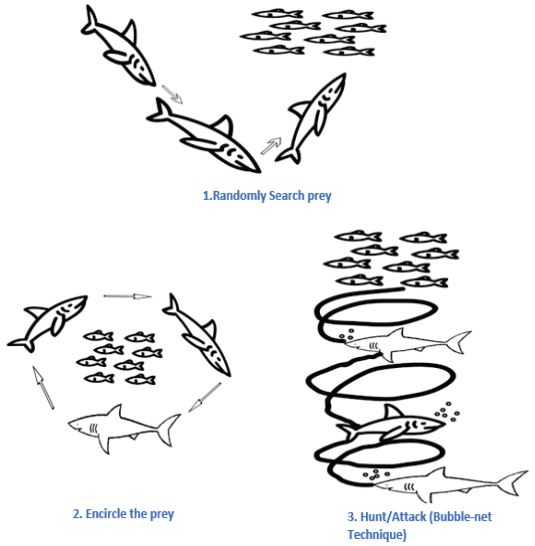


Figure 1 Visuals showing WOA

The time-based region of a whale is computed via the operational procedures: Encircling the prey, Bubble-net attacking stage and Search for the prey (Fig. 2).

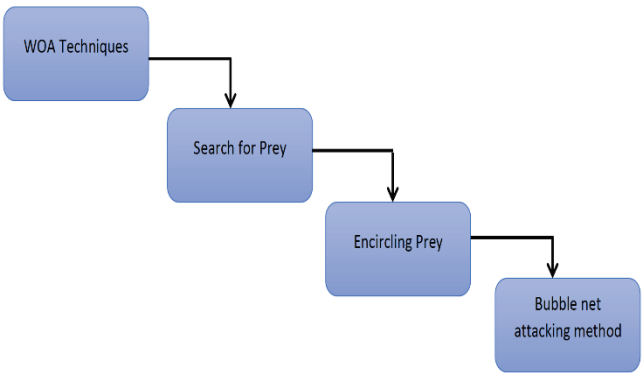


Figure 2 WOA Techniques

1.1 WOA Techniques

This section elaborates the technique used in WOA.

1.1.1 Encircling the Prey

Humpback whales are able to understand specified area of prey and then encircle those. Because location of ideal plan within the look area isn't recognized inferably, WOA supposes that present finest nominee arrangement is goal prey otherwise nearest of ideal. Then exertion is used in order of recognizing finest look operator, whereas alternative look operators will replace their positions close to the leading look operator. The conduct is communicated via the taking after conditions as expressed through equation below.

$$D = \left| \vec{C} \times \vec{X}'(t) - \vec{X}(t) \right| \quad (1)$$

$$\vec{X}(t+1) = \vec{X}'(t) - \vec{A} \times \vec{D} \quad (2)$$

$$\vec{A} = 2 \times \vec{a} \times \vec{r} - \vec{a} \quad (3)$$

$$\vec{C} = 2 \times \vec{r} \tag{4}$$

Here X is common best location X^* presenting whale location, t indicates the later cycle, a speaks to linearly decreased inside run of two to over the course of iterations, and the r may be an irregular number consistently allotted in the run of [0, 1]. The sign | | speaks to the utter value.

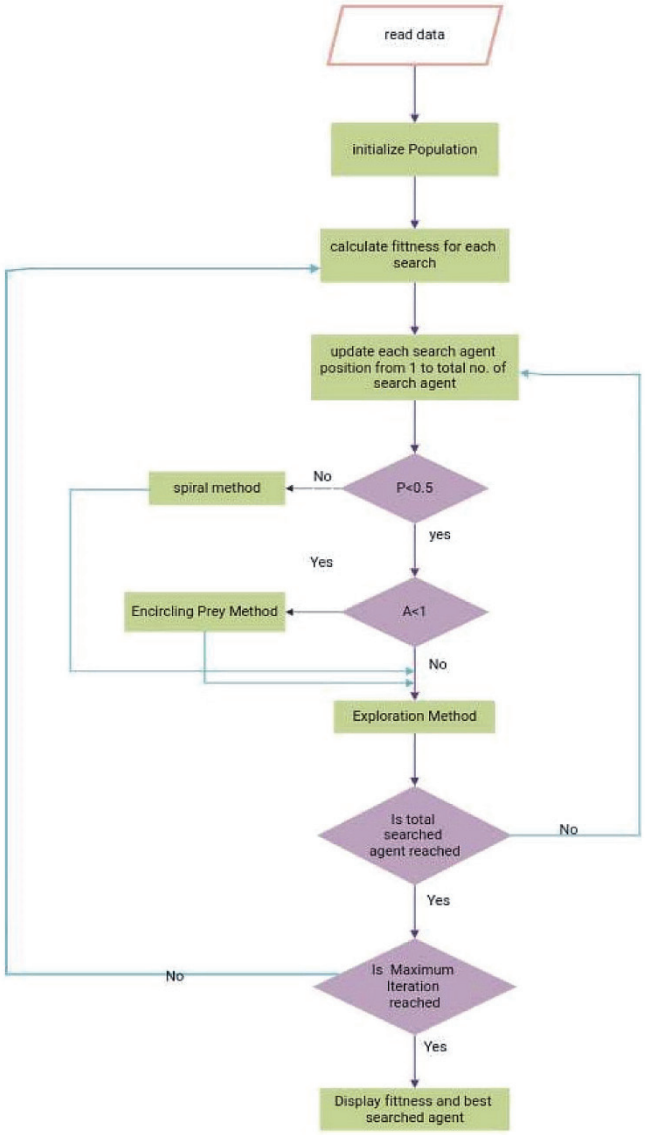


Figure 3 Flowchart for WOA

1.1.2 Bubble-net Attacking Technique (Manipulation Section)

Towards defining the bubble-net conduct about humpback whale, winding scientific definition remains connected among location of prey and whale in-order of mimicking the twirl-formed motion of humpback whale, within which p speaks to a steady for clarifying the form of the exponential winding and the k is self-assertive value uniformly distributed within extend of [–1, 1].

$$\vec{X}(t+1) = \vec{D}^n \times e^{bl} \times \cos(2\pi l) + \vec{X}' \tag{5}$$

$$\vec{X}(T+1) = \begin{cases} \vec{X}'(t) - \vec{A} \times \vec{D} \text{ if } p \geq 0.5 \\ \vec{D}'' \times e^{bl} \times \cos(2\pi l) + \vec{X}' \text{ if } p \geq 0.5 \end{cases} \tag{6}$$

1.1.3 Search for the Prey (Exploration Section)

To get worldwide optimizers, on the off chance that $A < -1$ or $A > 1$, search agent gets upgraded like expressed by a haphazardly selected search agent within the posit of leading look operator (Fig. 3). Where X is designated subjectively from whales in the current cycle.

$$\vec{D} = \left| \vec{C} \times \vec{X}_{\text{rand}} - \vec{X} \right| \tag{7}$$

$$\vec{X}(t+1) = \vec{X}_{\text{rand}} - \vec{A} \times \vec{D} \tag{8}$$

1.2 Time Complexity

The very interest in the WOA approach is driven by the fact that it is a search -based approach using random inputs. The same is observed for real whales also. Due to this, the result generated by the WOA approach may differ every time the algorithm is run. The WOA is thus a randomized-approximation algorithm used to solve NP-complete combinatorial optimization problems (problems whose optimal solution cannot be found in polynomial time due to the vastness in probable search space). Hence, the result generated is not guaranteed to be the best, yet is much better than any other random result. The justification for using an approximation approach comes from the time taken by the algorithm.

1.3 Whale Optimization – The Approach

1. Initialization – Wi digital whales are created ($1 \leq i \leq n$)
2. Compute fitness for every whale on food search
   Let W* be the best search Whale
   While (i < Total_no_of_iterations)
   For each Wi: update the parameters l, L, M, N and P
      If ($p < 0.5$)
      Then if ($|L| < 1$) then Compute Position of Wi using Eq. (2)
      else if($|L| \geq 1$)
         Choose random Wi
         Compute Position of Wi using Eq. (8)
         Else if ($p \geq 0.5$)
         Compute Position of Wi using Eq. (5)
   If any Wi goes outside the searched space, then correct its path
   Compute fitness for every Wi.
   Recompute W* for whale with best fitness (if better than before)
   $I = i + 1$
   End While loop
   Return optimum answer (W*)

There are only 2 loops in the WOA approach as given in the previous section. Hence the complexity is clearly a

product of '*N*', the number of whales, and total number of iterations (*NI*). Henceforth the complexity of WOA approach is ***O*(*N.NI*)**, which is polynomial time and very quick as compared to other approaches.

This work contributes in the following ways:

- Comprehensively providing insight to the available datasets, tools, comparison techniques and recent trends of employing WOA to various engineering problems.
- Providing a taxonomy of various kinds of applications to which WOA has been applied.
- Assessing and analyzing the results of applying WOA to various optimization problems.

Outline: Section II discusses the related work. Section III explains the transient applications of WOA. Section IV and V investigate the results and conclusion respectively.

2 RELATED WORK

WOA has been adopted for solving problems in various fields. Although, there exists some other survey papers as well. Some of them are Comprehensive survey [2], Meta-analysis survey [3] [4]. These survey papers have also, included many applications of WOA in various fields. In Addition, this paper gives more relevant information based on the type of methods used (New, Modified, Improved); Year of Publish; Method Problem; Dataset Used; Tool used; Compared with.

2.1 Research Procedure

A complete inquiry about technique was espoused for investigating articles present on WOA which had been distributed in famous logical records. The articles ensued elected by backward sifting prepare as per characterized set of measures, which navigate through four phases appeared in Fig. 7. At last, significant articles were concluded based on exclusion and consideration measures for the audit reason

2.1.1 Keyword-Based Article Search

Watch word aspect handle was exhausted two steps. In the initial step, greatest number of reports were recovered from the foremost important records utilizing the succeeding keyword: 'Whale Optimization Algorithm', 'WOA', 'Reduction encompassing prey', 'modified WOA' [2], 'enhancement of WOA', 'hybrid plus WOA', etc. Within next moment walk, a record sifting handle was executed to avoid insignificant reports. Subtle elements of categorizing prepare can be found in Orders. 4.4 and 4.5 'Assortment criteria' and 'extraction of the data' individually**.**

2.1.2 Search Strategies

The order to search to closely all spaces, anywhere WOA was performed. The article from all present trustworthy online records were looked, and after that thorough handle of confirmation for references and citations within the papers confirmed the determination as it was high characteristic peer-assessed articles. Besides, proximity of articles within diverse fields former building and computer discipline were also reserved into thought for the consider.

2.1.3 Database Selection for Research

Examination deliberates utmost of true publisher's records accessible online like on: ACM, ScienceDirect, IEEE Explorer, Springer. Tab. 1 keeps up sum of records reverted as of the over said databases. In order to supply substantial evidence of the fast development to investigate on the WOA, Fig. 4 comprises the total of articles distributed every year through the related publishers. Fig. 5 underpins the equivalent with graphic illustration of percent of articles grounded on publishers. IEEE, Elsevier and Springer were the databases in which maximum chosen articles on WOA were found to be published. Fig. 6 increases the confirmation by showing publisher-based percentage of the significant articles. Also around 30% of the papers were found to be published in databases other than the IEEE or Springer or Wiley or ACM. Surge in other publication houses other than the traditional ones could be one major reason behind this.

Table 1 WOA in general Academic Records

Online Databases	Article Returned
IEEE	151
Springer	145
Elsevier	143
Wiley	39
ACM	21
Others	208

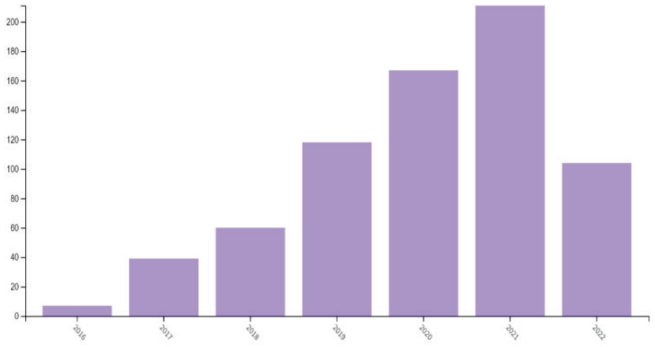


Figure 4 Number of Publications/Year

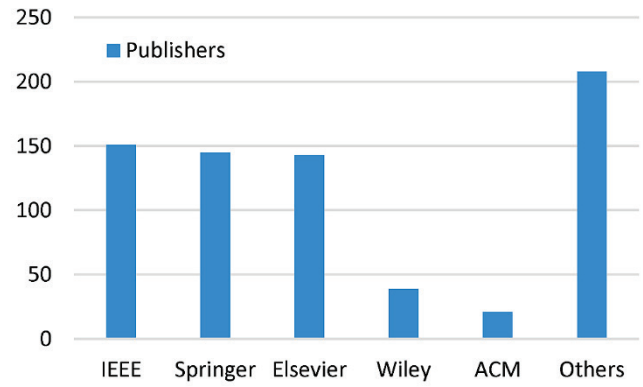


Figure 5 Sum of Articles by Publishers

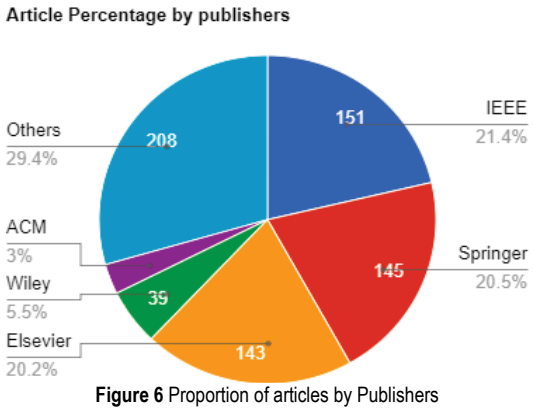


Figure 6 Proportion of articles by Publishers

2.1.4 Article Assortment Criteria

Article assortment criteria was set to integrate foremost appropriate documents and dispose of insignificant amongst them. A scrupulous way of examination was taken after to categorize related articles coordinating towards the goal were seen. At that point a careful perusing investigation was executed on titles, abstracts and conclusions in order for casting the chosen articles for additional processing (Fig. 7). In expansion, sifted articles were systematically reviewed to authenticate the suitability of articles for the purposes of this assessment. At last, foremost pertinent articles stood encompassed for this survey.

2.1.5 Mining of Data

Thorough information mining prepare was accomplished upon articles headed for dodge repetition in data. A complete record of information hence recovered from databases was sustained into a exceed expectations spreadsheet (woasurvey.xlsx). Formerly, numerous classification forms were executed upon data like application allied classification (areas and sub parts of areas), method allied classification (Alteration, Hybridizing and Multi-impartial, etc.), year allied classification.

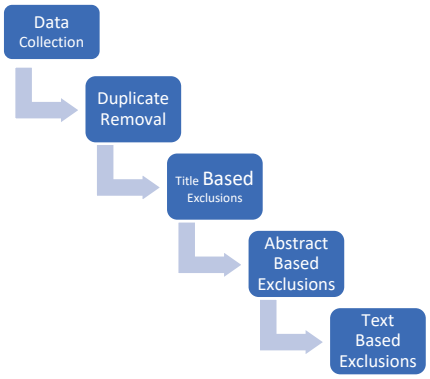


Figure 7 Inclusion Criteria

3 APPLICATIONS OF WOA

Scholarly works has appeared a few regions of the application of WOA to cope with actual-world issues. Maximum of the studies examined its execution with different standard meta-heuristic besides unique old-style optimization strategies. Also, applications of WOA are rising speedily in numerous orders owing to its proficiency and simply versatile highlights. Here, we portray the pertinency of WOA within totally diverse extents, and sub extents of building areas by means of the most purpose of this review. Mainly, applications of WOA are categorized in five key areas and numerous subfields and display scientific categorization within Fig. 8.

WOA has majorly been applied to 5 broad categories including Electrical Engineering, Computer Engineering, Applied Mathematics, Aeronautical Engineering and Construction Engineering. Out of these the major applications can be found for applying WOA in solving various computational problems.

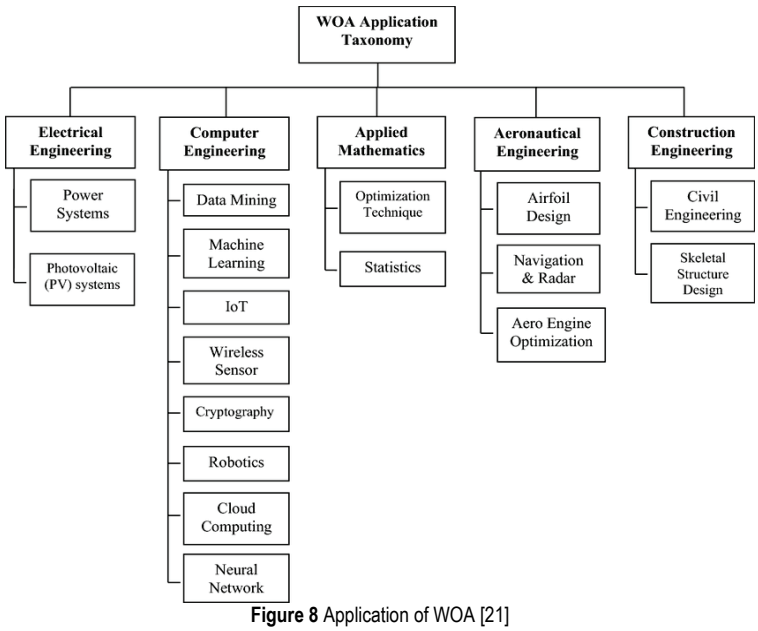


Figure 8 Application of WOA [21]

Table 2 Application Analysis of WOA

Link	Technique	YOP	Application Area	Problem	Proposed Method	Dataset Used	Tool Used	Comparison
[6]	Improved	2019	Computational Engineering	0/1 Knapsack problem	Opposition-based Whale Optimization Algorithm (OWOA)	Twenty cases of Knapsack problem,	MATLAB R2016a	OBL with WOA, (CGMA) and HS-Jaya
[7]	Proposed	2022	Cloud Computing	Privacy-protected data publishing in cloud computing	Brain storm-based Whale Optimization Algorithm	Chess T10I4D100 K, and the retail databases	-	-
[8]	Proposed	2021	Neural Networks	Enhancing Neural Networks	A New Harris Hawk Whale Optimization Algorithm	Cancer datasets and other datasets like fraud, banking note authentication.	-	-
[9]	Improved	2018	Machine Learning	Feature Selection of Network Intrusion Detection	Improved Binary Whale Optimization Algorithm	UCI database, KDD CUP 99 data set	-	Genetic Algorithm (GA), PSO and WOA,
[10]	Improved	2020	Neural Networks	multi-layer perceptron neural network training	Teaching learning-based whale optimization algorithm	Fifteen UCI benchmark datasets	-	GOA, GSO, SSO, FPA, GA and WOA
[11]	Modified	2020	Machine Learning	Feature Selection in Micro Array Cancer Dataset	Modified Whale Optimization Algorithm	Microarray datasets (Leukamia, Lymphoma, Prostate)	-	Classification models using SVM, NB and KNN with PSO algorithm
[5]	Proposed	2017	Machine Learning	Wrapper feature selection	WOA approaches for wrapper feature selection	Eighteen UCI benchmark datasets	MATLAB	Particle Swarm Optimization (PSO), Genetic Algorithm (GA), the Ant Lion Optimizer (ALO), and five standard filter feature selection methods
[12]	Proposed	2020	Computational Engineering	Enhance Software Project Effort Estimation	Exploring WOA to Enhance Software Project Effort Estimation	NASA 93	-	-
[13]	Improved	2021	Applied Mathematics	Constrained Optimization Problems	Improved Whale Optimization Algorithm	Single-modal functions and complex nonlinear multimodal functions	-	Comparision with Basic WOA
[14]	Proposed	2021	Machine Learning	Efficient high-dimensional feature selection approach	Spatial bound whale optimization algorithm	16 high-dimensional datasets gathered from Arizona State University (TOX_171, Colon, and Prostate_GE)	-	Eight state-of-the-art feature selection methods
[15]	Proposed	2018	Applied Mathematics	Maximum flow problem	WOA for solving Maximum flow Problem	-	MATLAB	Average run time and the accuracy of the results with ford-fulkerson Algorithm
[16]	Proposed	2017	Machine Learning	Feature Selection	Feature Selection approach based on WOA	Sixteen UCI benchmark datasets		particle swarm optimization (PSO) and genetic algorithm (GA)
[17]	Modified	2020	IoT	Control of switched reluctance motor	Cosine adapted modified whale optimization algorithm (CamWOA)	Unimodal benchmark functions	MATLAB/SIMULINK	WOA, WCMFOA, ABCA, GA, PSA, GS, CS, GSA, Hybrid PSOGSA, WCA, MFOA, DA
[18]	Improved	2021	Power Systems	Optimize the cascade reservoir power generation	Adaptive weights and differential mutation perturbation factors (APWOA)	Sphere, Griewank, Rastrigin, Ackley	-	WOA performance testing, APWOA, PSO, linearly decreasing weight PSO
[19]	Proposed	2019	Wireless Sensor Networks	Discover communities in networks	Whale optimization based community detection algorithm (WOCDA)	Karate, Dolphins, Polbooks, Football, Jazz, Elegans, Email, Facebook Networks	-	DPSO, iMeme-net, BA, AFSA
[20]	Proposed	2021	Computational Engineering	Enhance Apache Spark's computational performance	Spark-based parallel processing whale optimization algorithm	Alpine, Ackley, Sphere, and Schwefel Benchmark functions	AmazonWeb Services Elastic MapReduce (EMR)	MapReduce WOA (MR-WOA), and serial implementation of WOA; Nonparametric Wilcoxon's signed rank sum test

Tab. 2 summarize the gathered details about various studies analyzed for using WOA. As represented in Fig.8, many applications of WOA were used for solving computational engineering problems [5, 7-13, 15, 17, 21].

Few applied mathematics problems were also solved using WOA [14, 16].

It can also be observed that 12 different versions of WOA have been utilized (as listed in Tab. 2) based on the improving the existing WOA algorithms or utilizing a hybrid version by combining with some other technique. All of the applications have been tested on different datasets (Tab. 2) due to the difference in the very problem statement being solved using WOA. As per the tools used, majorly MATLAB has been utilized for the coding and analysis purposes as mentioned in the reviewed studies.

4 RESULT AND DISCUSSION

The current research led to the observation that WOA is already utilized within nearly every general field of building space aimed at measuring complex optimization issues. The prejudice of the search-based WOA like swarm insights optimization method is speedily expanding amid the scholarly community and businesses meanwhile its beginning. By means of the results gathered from the current research, the application of WOA is anticipated to be more developed in assorted other pertinent areas of designing further stated within this study (electrical as well as control system, computer building, aircraft building, and construction coaxing). A few adaptations of WOA grounded techniques are created via analysts throughout the globe in far along former to address various engineering optimization issues like binary-based technique, feature selection, and disapproval-based learning means.

Amid studied work, feature selection-based and disapproval-based learning strategies were foremost predominant methods attained within writing aimed at understanding advanced optimization issues.

WOA is used in the field of Neural Network to enhance masses and the prejudice of a neural network. Also, proposed algorithm is used in diverse kind of datasets like cancer, deception, investment memo authentication [8].

Proposed protection conservation demonstration was implemented particularly for cloud segment [7]. The projected prototypical intricate two fundamental phases: information decontamination then renovation through an ideal key creation. The ideal key remained augmented by means of proposed J-SSO calculation via determining a multi-intentional role amongst the constraints just as degree for adjustment, data conservation proportion. Whereas bearing in mind the level of expansion, search-based calculations for resolving privacy conservancy issue with taking care of various information, PSV-GWO faces bad neighborhood-looking capacity and moderate merging.

5 CONCLUSION

Transformation-based calculations were broadly examined search-based calculations in designing spaces to espouse optimization issues. Fast development of intelligent search-based calculations in later years enthused the analysts towards investigating pertinence within multidisciplinary areas. Within the similar vein, our survey attempts to sightsee the possibilities within recently created search-based WOA over an orderly prose review. Research is based on WOA: applications within diverse disciplines, alterations, and crossings, to confront different combinations for optimization issues. In spite of the reality that outset of this calculation is very current, its overpowering development completely diverse regions is rather self-evident. The larger part of articles checked on in this contemplate is focused on, the applications of calculation, modification-based tactic to an arrangement, cross breed strategy of a solution thru combining other strategies or calculations, and constraint enhancements. Be that as it may, there's chamber for change in calculation to unravel multi-dimensional and multimodal difficulties. Additionally, binary-based, feature selection based, choice-based require a part of effort hours within upcoming research to illuminate multifaceted optimization issues over WOA. An update of multi-intentional strategy by means of alternating the single-intentional constraint to multi-intentional constraints through same calculation may be additional exhilarating range of enhancements worth investigating.

A key parcel of works examined within this survey is towards resolving the worldwide optimization issue in excessive dimensional region. However, supplementary multifaceted perspectives, like ideal switch problem utilizing parallel mathematical strategy, disordered sequence issue, different backpack issue, multi-class SVM issue, stage stream shop issue, obligation issue, repetition assignment issue, and many additional, can be occupied into consideration for encouraging research track. This audit examines the applications of WOA within different areas and sub parts of engineering fields aimed at tackling optimization issues. This paper unlocks way for investigating additional intriguing and multi-disciplinary arenas that can be tattered by the smoothness of WOA.

6 REFERENCES

[1] Mirjalili, S., & Lewis, A. D. (2016). The whale optimization algorithm. *Advances in Engineering Software*, *82*, 51–67. https://doi.org/10.1016/j.advengsoft.2016.01.008

[2] Gharehchopogh, F. S., & Gholizadeh, H. (2019). A comprehensive survey: Whale optimization algorithm and its applications. *Swarm and Evolutionary Computation*, *48*, 1–24. https://doi.org/10.1016/j.swevo.2019.03.004

[3] Mohammed, H. M., Umar, S. U., & Rashid, T. A. (2019). A systematic and meta-analysis survey of whale optimization algorithm. *Neural and Evolutionary Computing*. https://doi.org/10.1155/2019/8718571

[4] Boopalan, S., & Preena, R. R. E. (2021). Meta-analyses in survey of whale optimization algorithm. *International Journal of Emerging Technologies and Innovative Research*, *8*(2), 803–825.

[5] Mafarja, M., & Mirjalili, S. (2018). Whale optimization approaches for wrapper feature selection. *Applied Soft Computing*, *62*, 441–453. https://doi.org/10.1016/j.asoc.2017.11.006

[6] Alamri, H. S., Zamli, K. Z., Razak, M. F. A., & Firdaus, A. (2019). Solving 0/1 knapsack problem using opposition-based whale optimization algorithm (OWOA). In *Proceedings of the*

2019 8th International Conference on Software and Computer Applications (ICSCA '19) (pp. 42–46). ACM. https://doi.org/10.1145/3316615.3316658

[7] Revathi, S. T., Ramaraj, N., & Chithra, S. (2019). Brain storm-based whale optimization algorithm for privacy-protected data publishing in cloud computing. *Cluster Computing*, *22*(2), 3521–3530. https://doi.org/10.1007/s10586-018-2200-5

[8] Agarwal, P., Mehta, S., Gupta, A., & Farooqi, N. (2021). A new Harris hawk whale optimization algorithm for enhancing neural networks. In *2021 Thirteenth International Conference on Contemporary Computing (IC3-2021)* (pp. 179–186). IEEE. https://doi.org/10.1145/3474124.3474149

[9] Xu, H., Fu, Y., Fang, C., Cao, Q., Su, J., & Wei, S. (2018). An improved binary whale optimization algorithm for feature selection of network intrusion detection. In *2018 IEEE 4th International Symposium on Wireless Systems within the International Conferences on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS)* (pp. 136–141). IEEE. https://doi.org/10.1109/IDAACS-SWS.2018.8525605

[10] Zhou, Y., Niu, Y., Luo, Q., & Jiang, M. (2020). Teaching learning-based whale optimization algorithm for multi-layer. *Mathematical Biosciences and Engineering*, *17*(5), 5987–6025. https://doi.org/10.3934/mbe.2020320

[11] Sathya, M., & Priya, S. M. (2020). Modified whale optimization algorithm for feature selection. *International Journal of Scientific & Technology Research*, *9*(3), 2277–8616.

[12] Fadhil, A. A., & Alsarraj, R. G. (2020). Exploring the whale optimization algorithm to enhance software project effort estimation. In *6th International Engineering Conference: Sustainable Technology and Development (IEC)* (pp. 146–151). https://doi.org/10.1109/IEC49899.2020.9122918

[13] Ning, G.-Y., & Cao, D.-Q. (2021). Improved whale optimization algorithm for solving constrained optimization problems. *Discrete Dynamics in Nature and Society*, *2021*, Article 6639744. https://doi.org/10.1155/2021/8832251

[14] Too, J., Mafarja, M., & Mirjalili, S. (2021). Spatial bound whale optimization algorithm: An efficient high-dimensional feature selection approach. *Neural Computing and Applications*, *33*(23), 16229–16250. https://doi.org/10.1007/s00521-021-06224-y

[15] Masadeh, R., Alzaqebah, A., & Sharieh, A. (2018). Whale optimization algorithm for solving the maximum flow problem. *Journal of Theoretical and Applied Information Technology*, *96*(8), 2208–2220.

[16] Sharawi, M., Zawbaa, H. M., & Emary, E. (2017). Feature selection approach based on WOA. In *Ninth International Conference on Advanced Computational Intelligence (ICACI)* (pp. 163–168). IEEE. https://doi.org/10.1109/ICACI.2017.7974502

[17] Saha, N., Panda, S., & Sidhartha, M. (2020). Cosine adapted modified whale optimization algorithm for control of switched reluctance motor. *Computational Intelligence*, *38*(3), 978–1017. https://doi.org/10.1111/coin.12310

[18] Lü, J.-X., Yan, L.-J., Pa, T.-S., Chu, S.-C., Pan, J.-S., He, X.-K., & Chang, K.-C. (2021). An improved whale optimization algorithm and its application to power generation in cascade reservoir. In *Advances in Intelligent Information Hiding and Multimedia Signal Processing (Smart Innovation, Systems and Technologies, Vol. 211)* (pp. 228–237). Springer. https://doi.org/10.1007/978-981-33-6420-2_28

[19] Zhang, Y., Liu, Y., Li, J., Zhu, J., Yang, C., Yang, W., & Wen, C. (2020). WOCDA: A whale optimization based community detection algorithm. *Physica A: Statistical Mechanics and Its Applications*, *539*, 122937. https://doi.org/10.1016/j.physa.2019.122937

[20] Alshayeji, M., Behbehani, B., Ahmad, A., & Imtiaz, W. (2021). Spark-based parallel processing whale optimization algorithm. *Concurrency and Computation: Practice and Experience*, *34*(4), e6192. https://doi.org/10.1002/cpe.6192

[21] Rana, N., Latiff, M. S. A., Abdulhamid, S. M., & Chiroma, H. (2020). Whale optimization algorithm: A systematic review of contemporary applications, modifications and developments. *Neural Computing and Applications*, *32*, 16245–16277. https://doi.org/10.1007/s00521-020-04849-z

Authors' contacts:

Samiksha Yadav
Indira Gandhi Delhi Technical University for Women (IGDTUW),
Kashmere Gate, New Delhi, Delhi 110006, India

Ankita
Indira Gandhi Delhi Technical University for Women (IGDTUW),
Kashmere Gate, New Delhi, Delhi 110006, India

Shweta Jindal
Indira Gandhi Delhi Technical University for Women (IGDTUW),
Kashmere Gate, New Delhi, Delhi 110006, India

Nishtha Jatana
Maharaja Surajmal Institute of Technology,
4 Janakpuri, New Delhi, Delhi 110058, India

Charu Gupta
Bhagwan Parshuram Institute of Technology,
Sector-17, Rohini, PSP-4, New Delhi, Delhi 110085, India

Nitesh Singh Bhati
Gautam Buddha University,
Opp. Yamuna Expressway, Greater NOIDA,
Gautam Budh Nagar, Uttar Pradesh-201312, India

Dimple Tiwari
(Corresponding author)
School of Engineering & Technology,
Vivekananda Institute of Professional Studies - Technical Campus,
Outer Ring Rd, AU Block, Ranikhet, Pitampura, New Delhi, Delhi, 110034, India
E-mail: dimple.tiwari88@gmail.com

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20241210161405

Subject review
Received: 2024-12-10, Accepted: 2025-09-19

Circular Logistics as Part of the Supply Chain

Dora Kolaric*, Ivan Grgačić, Matej Stoprd, Dunja Horvat

Abstract: This paper investigates circular logistics and its impact on the efficiency and sustainability of the supply chain. Key trends from 2019 to 2024 are analyzed, including the adoption of advanced technologies such as Blockchain, the Internet of Things (IoT), and artificial intelligence (AI), as well as recycling and material reuse, simplified product returns, and sustainable packaging. A qualitative analysis of secondary data, including scholarly articles and company reports, shows that the implementation of circular logistics reduces operational costs and waste, while its success depends on developed recycling infrastructure and the adaptation of business models. The study connects circular economy theory with practical examples, provides insights into the challenges and success factors in integrating circular models into supply chains, and offers guidance for future research and practical application.

Keywords: circular logistics; recycling; supply chain; sustainable packaging; sustainability

1 INTRODUCTION

Circular logistics is a comprehensive concept that has become crucial in modern supply chain management. In contrast to traditional supply chains, where products move linearly from raw materials to end consumers and often end their life cycle in landfills, circular logistics seeks to minimize waste and maximize resource use through reuse, recycling, and product regeneration [1]. This approach not only supports sustainability but also contributes to cost reduction and improved ecological efficiency, making it a strategic tool for companies seeking to optimize their operations [2]. The concept of reverse logistics flows originally emerged from the need for alternative ways of managing supply chain processes in increasingly demanding markets. The harmful effects of environmental degradation, pollution, and climate change served as tangible evidence of the inadequacy of traditional linear models. Consequently, awareness of the importance of sustainability has increased significantly in recent years. As the entire product life cycle, from raw material extraction to disposal, gained more attention, new demands arose to support the principles of sustainable development [3]. Circular logistics enables products to extend their life cycle and avoid ending up in landfills, but its definition and application have proven challenging due to its inherent stochasticity and complexity. At the same time, consumer demands have become increasingly urgent, requiring a rethinking of existing logistics networks and structures. Manufacturers and other stakeholders are forced to adopt new technologies, faster and more efficient delivery models, and greater reliance on e-commerce to remain competitive and maintain market share, particularly over the past decade [3]. Despite the growing recognition of its importance, systematic analyses of how circular logistics affect the entire supply chain remain limited, especially in specific regional contexts. The research problem therefore arises from the limitations of the traditional "take–make–use–dispose" supply chain model, which can no longer ensure long-term sustainability or competitiveness. Companies must seek new approaches that combine profitability, environmental responsibility, and responsiveness to consumer needs. Circular logistics presents one such alternative, but further research is required to determine its actual impact on supply chain efficiency and to identify the main challenges of its implementation. To address this gap, the aim of this paper is to examine the role of circular logistics within modern supply chains and to analyze its effects on business efficiency and sustainability. The specific objectives are to:

- Analyze the development of logistics in the past five years with an emphasis on circular logistics trends.
- Identify the advantages and disadvantages of circular logistics compared to traditional models.
- Examine key challenges and success factors in implementing circular logistics.
- Connect theoretical insights with practice through real-world business examples.

This study applies a qualitative-analytical method based on literature reviews, case studies, and comparisons of existing circular logistics models. The strength of this approach lies in its ability to capture the complexity of processes and highlight success factors often overlooked by purely quantitative studies. The motivation for writing this paper stems from the need to better understand how circular models can be integrated into supply chains in practice, as well as from the increasing consumer pressure that compels businesses to adapt their strategies and innovate. By answering the following research questions, the study aims to contribute to a deeper understanding of circular logistics as a driver of sustainability and competitiveness in modern supply chains:

RQ1: How does the implementation of circular logistics affect the efficiency of the supply chain in terms of cost and waste reduction?

RQ2: What are the key challenges and success factors in integrating circular logistics into a company's traditional logistics system?

H1: The implementation of circular logistics significantly reduces operational costs and the amount of waste in the supply chain.

H2: The successful application of circular logistics depends on the development of recycling infrastructure and

product return systems, as well as the adaptation of business models.

By addressing these questions and hypothesis, this research seeks to highlight the transformative role of circular logistics in creating resilient, efficient, and sustainable supply chains. The structure of the paper is organized as follows. After the introductory chapter, the second part provides the definition and theoretical framework of circular logistics, with a particular focus on its integration into supply chains, the optimization of routes in waste collection, and the challenges companies face in its implementation. The third part of the paper focuses on research into circular logistics trends as part of the supply chain and includes the research methodology, presentation of results, and discussion. The fourth part presents the conclusions of the paper, while the fifth part contains the list of references.

2 DEFINITION AND THEORETICAL FRAMEWORK OF CIRCULAR LOGISTICS

Circular logistics, as part of the supply chain, is increasingly gaining attention in the context of sustainable business and waste reduction. Circular logistics focuses on minimizing waste and maximizing resource utilization through reuse, recycling, and material regeneration. This approach not only helps preserve the environment but can also significantly improve supply chain efficiency. The implementation of circular logistics is becoming crucial for modern companies that wish to optimize their operations and align them with the principles of sustainable development [4]. Circular logistics represents an extension of the traditional logistics concept, with an emphasis on the reverse flow of products and their reuse. Unlike the linear model, circular logistics closes the loop of material movement, reducing the need for new resources and decreasing the amount of waste that ends up in landfills [5]. The connection between logistics and sustainability is also highlighted in the literature dealing with waste management and reverse logistics [6] [7]. According to Harris, sustainable development cannot be achieved without serious reforms in the ways businesses manage their supply chains [8]. Circular logistics lays the foundation for such a transformation, enabling businesses to efficiently utilize resources and respond to environmental demands and consumer pressures. The integration of this model into all stages of the supply chain, from production to the end user and the return of products after use, is essential. More efficient handling of goods in the return process is crucial due to the constant decline in product prices. The functioning of the supply chain, in addition to reducing the percentage of returned goods, also enables a faster flow of returned products. Return logistics activities include the processes a company uses to collect used, damaged, unwanted products, or those that have expired, as well as packaging from end consumers or suppliers. Reverse logistics is a relatively new concept, and services in this area are gradually developing. Once a product is returned, the company can redistribute it to maximize the product's value [9].

2.1 Integration of Circular Logistics in the Supply Chain

The implementation of circular logistics in supply chains involves the optimization of logistics systems, including product return processes. Waters and Donald emphasize the importance of logistics as a key part of the entire supply chain in their work. Traditional logistics focuses on optimizing the flow of goods and reducing costs, while circular logistics expands these goals to include sustainability and waste minimization. Modern supply chains must integrate sustainable practices such as reverse logistics, waste reduction, and recycling to meet the growing demands of the market and consumers [10]. By combining all the mentioned elements, it is possible to create sustainable supply chains that not only reduce the negative impact on the environment but also provide economic benefits through more efficient resource use and cost reduction. De Brito, Rommert, and Douwe emphasize that reverse logistics is a key aspect in supply chains when it comes to waste reduction and resource maximization [11]. In the example of urban environments, Awomeso, Taiwo, Gbadebo, and Arimoro highlight the importance of proper waste management for reducing pollution and improving the environment [12]. According to Quesada, circular logistics can also increase the value of products, as it allows materials to be recycled or reused instead of being discarded. These processes not only reduce the ecological footprint of organizations but can also bring financial benefits [13]. Rubio and Beatriz identify key challenges in managing reverse flows, including costs, the need for specialized equipment, and technical infrastructure [14]. One example of successful implementation of circular logistics is the automotive industry, which uses the return of vehicle parts for their refurbishment and reuse in the production of new cars. This reduces costs and also contributes to the reduction of CO_2 emissions and resource consumption. An integral approach to managing the municipal waste collection system involves the coordination of different sectors to ensure efficient collection, processing, and disposal of waste. According to Erdelezi, Margeta, and Knezić, such an approach allows for the optimization of operational processes and cost reduction, but it requires good infrastructure and support from local authorities [6]. By introducing modern technologies and planning methods, such as computer systems for waste tracking and analysis, it is possible to significantly improve the efficiency of the system. Ghiani, Gilbert, and Musmanno also emphasize the importance of optimizing logistics systems through planning and control, which is crucial for proper resource management in supply chains [15]. By combining these approaches, logistics systems can become more flexible and sustainable. Packaging waste represents a significant portion of total waste, and its proper management is crucial for environmental protection. Grum, in his work, emphasizes the need for developing systems that enable the recycling of packaging waste and the reduction of the amount of waste that ends up in landfills [16]. Developing legal regulations, such as the ban on single-use plastics, as well as raising awareness among consumers, are key steps toward a sustainable packaging waste management system. Marinela further emphasizes the importance of educating the population on proper waste disposal and implementing

effective recycling programs, which reduce environmental pressure and ensure the reuse of materials [17]. Packaging waste can become a valuable resource if properly managed, which is one of the fundamental principles of circular logistics. Circular logistics refers to the processes of returning products from consumers back to manufacturers so that these products can be reused, repaired, or recycled. Rogers, Lambert, Croxton, and García-Dastugue emphasize that managing the reverse flows of goods is crucial for a sustainable supply chain. This process involves complex operations related to returning products at the end of their lifecycle into the economic cycle [18]. Reverse logistics also allows companies to reduce production costs through the reuse of materials. Valle, Menezes, Reis, and Efigenio further emphasize the importance of customer service in the reverse logistics process. Customer satisfaction is a key factor for the success of recycling and reverse logistics programs, and companies must ensure efficient channels for product returns. Additionally, the development of technological solutions that enable easier tracking of products throughout their entire lifecycle helps optimize these processes [19].

2.2 Optimization of Routes in Waste Collection

Route optimization in waste collection becomes crucial in achieving an efficient waste management system, reducing costs, and minimizing the ecological footprint. In urban areas, with the increasing volume of waste and the complexity of collection, it is essential to find solutions that reduce unnecessary resources, time, and greenhouse gas emissions. One of the key elements of circular logistics is transportation optimization, especially in waste collection and recycling. In studies conducted, such as the one in the city of Trabzon (Turkey), route optimization helped reduce the total number of kilometers traveled by waste collection vehicles while improving system efficiency. Apaydin and Gonullu use mathematical models to identify optimal routes that reduce operational costs and the time needed for waste collection [20]. They also studied route optimization for solid waste collection in the city and demonstrated how proper route organization can significantly reduce costs and collection time [20]. Sulemana, Donkor, Forkuo, and Oduro-Kwarteng emphasize that route optimization is an important aspect of circular logistics, as it reduces fuel consumption and greenhouse gas emissions [21]. The integration of information and communication technology (ICT) into logistics systems also plays an important role in enhancing circular logistics. Asadi states that advanced ICT technologies can improve efficiency in managing reverse product flows and waste, contributing to better control and waste reduction in the supply chain [22]. These systems also help businesses better manage inventory, reduce costs, and increase efficiency. The digitalization of logistics processes enables better coordination between different stakeholders in the supply chain, facilitating product returns and recycling. For example, tracking systems and automatic ordering of products that are nearing the end of their lifecycle allow for quick responses to changes in demand and supply, reducing the risk of excess inventory and waste. Several algorithms are used for waste collection route optimization. The most commonly used are search algorithms. These algorithms enable the creation of optimal routes that consider variable factors such as traffic, distance between destinations, and vehicle capacity. Ghiani, Gilbert, and Musmanno emphasize that proper planning and optimization can significantly improve the efficiency of logistics systems, including those in waste collection [15]. Heuristic methods are often used when a quick, yet sufficiently good solution is needed for complex optimization problems. Heuristic approaches allow for the rapid evaluation of optimal routes and are frequently used in urban areas where there are a large number of collection points. Erdelez, Margeta, and Knezić suggest a combination of heuristic methods and traditional logistics approaches for better optimization [6]. The application of modern technologies, such as GPS systems and GIS (Geographic Information Systems), is crucial in tracking and optimizing routes. These systems enable real-time data collection, which helps adjust routes based on actual conditions on the ground, such as traffic delays or changes in waste volume. According to Marinela, GIS systems allow for precise route planning, reducing travel time and fuel consumption, which results in lower greenhouse gas emissions [17].

2.3 Challenges in the Implementation of Circular Logistics

Although circular logistics offers numerous benefits, there are challenges related to its implementation. Srivastava and Srivastava emphasize that reverse logistics often requires additional resources, such as specific infrastructure and technology for processing returned products [23]. The infrastructure for product return and recycling is often not sufficiently developed, making it difficult for companies to implement circular processes on a global scale. Kućar Dragičević, Butuči, and Kufrin state that in Croatia and other transitional countries, the main issues are related to inadequate waste management infrastructure and the lack of legal regulations that would encourage the implementation of circular models [24]. In addition, Potter emphasizes that companies must invest significant resources in adapting their business models in order to effectively implement circular logistics [25]. One of the main challenges in implementing circular logistics is the high initial cost. Investments in recycling technology, process optimization, and infrastructure adaptation can be significant. According to research by Frosch and Gallopoulos, the costs of implementing circular practices are often too high for small and medium-sized enterprises, which can slow down the spread of circular models [26]. The lack of universal standards and guidelines also presents a challenge. Currently, there are no globally accepted standards for evaluating the effectiveness of circular logistics practices, which can make comparison and implementation difficult. As Stahel points out, standardization is crucial for effective management of circular processes and reducing uncertainty in business operations [5]. Compatibility between different systems and technologies can also be problematic. Integrating existing logistics systems with new circular technologies often requires complex adjustments. According to research by the Ellen MacArthur Foundation, many organizations face challenges in adapting their IT systems and infrastructure to support circular practices [27]. The resistance of

organizations to change and the lack of support from employees also pose obstacles. Implementing circular logistics often requires a shift in culture and attitudes within the organization. Access to relevant data and information is crucial for the success of circular logistics. The lack of transparency in supply chains can complicate the tracking and optimization of circular processes. As Ghisellini, Cialani, and Ulgiati note, better access to information can significantly improve organizations' ability to implement circular models and track their performance [28].

3 RESEARCH ON CIRCULAR LOGISTICS TRENDS AS PART OF THE SUPPLY CHAIN

This research aims to support the mentioned theory with concrete examples and to answer the research questions and hypotheses. Circular logistics represents a significant shift in how organizations manage their supply chains, focusing on waste reduction and resource optimization. Over the past five years, this approach has become increasingly important due to global challenges related to sustainability and climate change. The goal of this research is to identify and analyze key trends in circular logistics that have emerged during the period from 2019 to 2024.

3.1 Research Methodology

The methods used in this study are based on a qualitative research approach, which is widely recognized in supply chain and logistics research for exploring complex, context-dependent phenomena. The research focuses on secondary data collected from scientific papers, books, industry reports, and other relevant sources published between 2019 and 2024, a period chosen to capture the most recent developments in circular logistics, including technological innovations, regulatory changes, and evolving consumer demands [1] [29]. The use of qualitative methods such as analysis, synthesis, description, induction, and deduction allows for a systematic examination of trends, patterns, and principles in circular logistics, providing insights that cannot be obtained through purely quantitative approaches [30]. This research is both descriptive and exploratory, aiming to document observed practices and identify new insights, gaps, and challenges in the application of circular logistics. The population consists of academic literature, industry reports, policy papers, and other credible sources addressing circular logistics worldwide. The sample was limited to sources published in the last five years (2019–2024) to ensure relevance and timeliness, while the universe includes peer-reviewed journal articles, books, official reports, and industry analyses relevant to circular logistics and supply chain sustainability [31]. The research was conducted in several steps. First, relevant sources were identified and collected from databases such as Scopus, Web of Science, and Google Scholar. Selection criteria included relevance to the topic, credibility, publication date, and focus on trends, innovations, and challenges in circular logistics. Finally, the findings were synthesized and linked to theoretical insights to draw conclusions and provide recommendations for practice [1]. The research instruments employed include literature review for comparison of sources, and analytical frameworks to facilitate the synthesis of insights and the connection between theory and practice. Data were analyzed thematically, using both inductive reasoning to derive general insights from specific observations and deductive reasoning to evaluate theoretical assumptions against the collected evidence. This structured approach ensures a rigorous examination of secondary data, enabling the identification of key trends and challenges in circular logistics and supporting actionable conclusions for modern supply chain management.

3.2 Research Results

In the last five years (2019–2024), circular logistics has undergone significant transformations, driven by the growth of e-commerce, increasing sustainability demands, and the adoption of advanced technologies. Recent studies highlight that reverse logistics has become a key area for reducing costs, enhancing operational efficiency, and improving customer satisfaction [36, 37, 39]. Based on an extensive literature review and supported by case examples, six key trends have been identified as shaping circular logistics globally. The selection of specific technologies was based on their proven relevance to circular logistics and their potential to address key challenges such as supply chain transparency, efficiency, resource optimization, and waste reduction. These technologies were chosen because they represent the most widely implemented innovations in reverse logistics, offering measurable improvements in operational performance and sustainability outcomes. The following are some of the key trends that have shaped reverse logistics globally:

1) Impact of Disruptive Technologies on Information Management (2019): Blockchain has increasingly been used to manage logistics, utilizing RFID chips to track the entire supply chain from manufacturers to sellers. This technology enables seamless exchange of goods and information, high security, transparency, and tracking among all participants. Customers can view the history of purchased items and track them through the supply chain [29, 30]. Blockchain can revolutionize reverse logistics management, accelerate processes, increase efficiency, and improve partner connectivity through digital platforms, thereby reducing barriers and encouraging innovation without additional costs. ShipChain, a blockchain-based solution, enables global cargo tracking with encryption and data accessibility limited to involved participants. This technology allows for better tracking of product history, generating reliable freight invoices, and reducing costs. The Internet of Things (IoT) provides clear visibility in storage and sharing of information between departments, enabling tracking, planning, and control of all processes with RFID technology and wireless sensors. The Smart Reverse Supply Chain (SRSC) allows for an intelligent combination of various systems, such as recycling and supply chain management systems, using advanced IT systems to manage product information throughout the entire lifecycle [31-33]. An example of successful Blockchain implementation in circular logistics is the introduction of this technology by De Beers in 2019. De Beers, a well-known diamond manufacturer, uses Blockchain to track diamonds throughout the entire supply chain, ensuring transparency and product

authenticity. All diamonds are tracked from the raw material stage to finished products, allowing customers to verify the history of their diamonds and ensure they are not linked to illegal activities [34, 35].

2) Recycling and Material Reuse (2020): IKEA (Ingvar Kamprad Elmtaryd Agunnaryd) is a Swedish company founded in 1943. Today, it is one of the largest global corporations in furniture retail, and in addition, it engages in real estate sales and operates restaurants within its stores. IKEA stands out as a leader in environmental conservation, implementing measures such as introducing eco-friendly bags, reducing water and electricity consumption, and using ecological materials [36]. By 2030, the company plans for all plastic in its production process to be made from renewable or recycled materials. Additionally, it aims for all wool in its products to come from responsible sources, and the cotton it uses to be sustainable, either recycled or grown with reduced resource use. In 2020, IKEA focused on recycling or incinerating 91% of the waste from its stores for energy use, while promoting more sustainable materials and LED lighting [37, 1].

3) Sustainability and Environmental Awareness (2021): In 2021, there was an increasing focus on sustainability in circular logistics. Companies began implementing strategies to reduce waste, recycle returned products, and use environmentally friendly packaging. H&M implemented textile recycling systems in 2021. H&M encourages customers to donate old clothing items at their stores, where the collected textiles are sorted and recycled into new products. H&M expanded its textile recycling program, known as "H&M Garment Collecting". Customers are invited to donate old clothes at H&M stores worldwide. The donated textiles are collected, sorted, and recycled based on their material and quality. The collected garments can be processed in various ways, including recycling into new textile fibers or converting them into industrial materials. H&M uses advanced technologies for sorting and recycling textiles. For example, the company has invested in material recognition technologies, which allow for more efficient separation and recycling. These technologies enable the recycling of fibers that would otherwise be unrecyclable, such as blends of different materials. The company aims that by 2030, all of their products will be made from recycled or sustainable materials. This includes using recycled fibers and materials from sustainable sources. H&M also plans to significantly increase the amount of textile they recycle compared to the amount they collect. H&M actively educates its customers about the importance of recycling and sustainability. The company organizes campaigns and events in its stores to raise awareness among consumers and encourage them to participate in recycling programs. H&M also provides information on how to properly dispose of old clothing and which materials are suitable for recycling [38] [39].

4) Simplified Product Returns (2022): An example is ASOS, known for its simple return policy, which offers a user-friendly interface with numerous filters and clarity to make shopping as easy as possible. The return policy allows for the return of unused and undamaged products within a specified period, with a refund. The customer simply repackages the product in its original packaging, fills out the return form, and follows the detailed instructions on the store's website. The return packaging is designed for easy reuse. ASOS also has a donation system in place, allowing customers to send clothing and footwear donations via DPD delivery service at no additional cost. This service collects donations at the time of order delivery, further simplifying the process. This efficient circular logistics system saves time and money and is focused on customer care and socially responsible activities [9, 40-42].

5) Integration of Artificial Intelligence (AI) (2023): In 2023, artificial intelligence (AI) became crucial for improving return processes in circular logistics. AI tools are used to optimize return routes, predict return patterns, analyze the causes of returns, and enhance operational efficiency. Amazon Robotics uses AI and robotics to improve its warehouse operations, including the returns process. AI algorithms in Amazon Robotics systems are used to optimize robot routes within the warehouse. These algorithms analyze real-time traffic and warehouse schedule data to determine the most efficient routes for robots transporting products, including returned items. This allows for faster and more efficient robot movement, reducing the time needed to process return shipments and increasing overall productivity. AI systems in Amazon Robotics are also used to automate the sorting of returned products. Warehouse robots are equipped with advanced sensors and cameras, allowing them to identify the type and condition of returned products. AI then analyzes this data to direct products to the appropriate locations for further processing, improving the speed and accuracy of sorting. AI also helps in predicting return volumes and planning resources. By analyzing historical return data, AI can predict return patterns and optimize warehouse capacities, robot allocation, and labor deployment. This allows Amazon to enhance its warehouse processes and reduce downtime caused by return processing. Amazon Robotics integrates its AI systems with other technologies, such as the Internet of Things (IoT) and Warehouse Management Systems (WMS). This integration enables real-time monitoring and management of return processes, as well as automatic adjustment of operations based on the current needs of the warehouse. By improving the speed and accuracy of processing returned products, AI and robotics help enhance customer experience. A faster return process reduces waiting times for customers, while more accurate sorting reduces the number of incorrect returns and exchanges, increasing customer satisfaction. Automation through AI and robotics significantly reduces the costs of circular logistics. It decreases the need for a large number of human workers, reduces errors, and increases efficiency, leading to a reduction in overall operational costs. This allows Amazon to offer more competitive prices and better service to customers [43, 44].

6) Sustainable Packaging and Circular Logistics (2024): In 2024, the trend of sustainable packaging and packaging management continues to expand within circular logistics. Companies globally are investing in eco-friendly materials and strategies to reduce waste, contributing to the sustainability and efficiency of their operations. Apple significantly advanced its sustainable packaging initiatives in 2024 through the "Environmental Progress Report" program. Apple introduced packaging using 100% recycled aluminum

for the casings of its products, including MacBook and iPhone devices. This material reduces the need for new raw materials and lowers the overall environmental footprint of production. Additionally, Apple uses recycled plastics and other materials for packaging, including recycled paper products for boxes and inserts, thus reducing the use of virgin materials. Apple has drastically reduced the size of its product packaging to cut CO_2 emissions associated with transportation. Smaller packaging allows for better space optimization in warehouses and transport vehicles, reducing the number of required shipments and greenhouse gas emissions. Apple has also optimized packaging design to use fewer materials without compromising product protection. The company employs advanced technologies, including AI and data analytics, to optimize packaging. These technologies help analyze demand and logistics patterns to improve packaging efficiency and reduce waste. For example, Apple uses AI to forecast and plan packaging needs in real time, allowing packaging to be tailored to specific market requirements and minimizing material surplus. Apple has introduced innovations in packaging design, including biodegradable materials and compostable options for packaging accessories like chargers and headphones. The company is researching new materials that can naturally decompose in the environment, further reducing the environmental footprint of packaging. As part of its "Environmental Progress Report" program, Apple regularly publishes reports on its packaging sustainability progress, including data on CO_2 emissions reductions, the amount of recycled materials used in packaging, and success in waste reduction. These reports allow customers and investors to track the company's progress toward its sustainability goals [45, 46].

Together, these trends demonstrate a shift towards more efficient, sustainable, and customer-focused reverse logistics strategies, driven by technological innovation and evolving consumer and environmental demands. By linking these observed practices with theoretical insights, this study establishes a stronger foundation for understanding circular logistics beyond individual company cases, highlighting the interplay between theory and practice in modern supply chains. Also, the study connects theoretical insights with practical applications, providing a foundation to analyze trends, challenges, and success factors in circular logistics. The methodology ensures that the selection of case studies and examples is justified, relevant, and directly linked to the research objectives, namely: to assess the impact of circular logistics on supply chain efficiency, to identify key challenges in its implementation, and to explore innovative solutions for sustainable operations.

3.3 Discussion

The research conducted on the trends of circular logistics from 2019 to 2024 confirmed everything previously supported by expert literature. The study was carried out using secondary data sources, specifically through non-financial reports from the companies mentioned and other verified sources. By combining case examples with theoretical frameworks, this research contributes to closing the gap between conceptual models of circular logistics and their real-world application. Through theoretical research, the following answers to the research questions and hypotheses were reached:

RQ1: How does the implementation of circular logistics affect the efficiency of the supply chain in terms of cost and waste reduction? The study indicates that circular logistics improves efficiency through multiple mechanisms. These include resource optimization, long-term cost savings, reduced operational management expenses, increased product durability, recycling and reuse, enhanced transparency, and compliance with regulatory frameworks [47, 42]. This confirms hypothesis H1, which posits that circular logistics reduces operational costs and waste. The theoretical basis for this finding aligns with the principles of the circular economy, emphasizing resource recirculation and the reduction of environmental impact [46]. While the initial implementation may involve higher costs, the long-term benefits, such as lower material expenditure and reduced waste management costs, generally outweigh these investments.

RQ2: What are the key challenges and success factors in integrating circular logistics into a traditional logistics system within a company? The study identified both operational and organizational considerations. Key challenges include organizational culture change, integration complexity, initial investment costs, supply chain complexity, and adherence to regulatory frameworks [39]. Success factors, supported by theoretical models of sustainable supply chain management, involve the adoption of innovative technologies, collaboration with partners, employee training and education, systematic monitoring and reporting, implementation of incentives, and the use of financial instruments to support transitions [44, 29]. These findings confirm hypothesis H2, emphasizing that successful circular logistics requires developed recycling infrastructure, efficient product return systems, and adaptive business models. Well-established infrastructure ensures effective material and product processing, while adaptive business models facilitate the integration of circular practices into daily operations, supporting waste reduction, resource optimization, and long-term sustainability goals [1].

H1: The implementation of circular logistics significantly reduces operational costs and the amount of waste in the supply chain. This hypothesis is confirmed both in theory and practice. According to [46], sustainable business models based on the principles of the circular economy enable cost reduction through recycling, reuse, and product life-cycle extension. This leads to lower demand for new resources and reduced waste management costs. Govindan and Hasanagic [47] further emphasize that the implementation of circular logistics not only decreases waste but also enhances supply chain efficiency by optimizing resources and reducing operational risks. Thus, it is evident that circular logistics can generate long-term financial and environmental benefits.

H2: The successful application of circular logistics depends on the development of recycling infrastructure and product return systems, as well as the adaptation of business models. The theoretical framework also confirms this hypothesis. Kirchherr, Reike, and Hekkert [48] stress that the circular economy cannot function without supportive

infrastructure that enables efficient collection, sorting, and recycling of products. In addition to infrastructure, the adaptation of business models is crucial for sustainability, as it allows the integration of reverse product flows into everyday operations [36]. Companies that have adapted their models, such as those adopting the "product-as-a-service" system, have demonstrated greater resilience and competitiveness in the market. This confirms that both infrastructure and business model adaptation are essential prerequisites for the successful implementation of circular logistics.

Overall, the results corroborate both hypotheses and provide theoretical support for the benefits and requirements of circular logistics. At the same time, they contribute to the literature by addressing the gap between theory and practice. Specifically, the study illustrates that the implementation of circular logistics is not merely an operational adjustment but a strategic approach requiring alignment between technology, infrastructure, organizational culture, and business processes to achieve sustainable supply chain performance. In doing so, this paper contributes to filling the gap in existing research by providing empirical evidence on how circular logistics trends are operationalized in global corporations, thereby enhancing our understanding of the transformative role of circular logistics in modern supply chains.

4 CONCLUSION

Circular logistics offers significant potential for improving ecological sustainability and business efficiency in supply chains. The integration of circular logistics flows, recycling, and material reuse are key to reducing the ecological footprint and increasing economic benefits. Although there are challenges in implementation, advanced technologies and examples of best practices show that sustainability can be achieved through circular logistics. Developing infrastructure, applying new technologies, and aligning legal regulations with the principles of a circular economy are essential steps that companies and governments must take to encourage the further adoption of circular logistics. Through this approach, it is possible to create sustainable and profitable supply chains that will meet the challenges of the 21st century. Route optimization in waste collection is crucial for achieving efficiency, reducing costs, and minimizing the ecological footprint in waste management systems. The use of advanced algorithms, heuristic approaches, and technologies such as GIS enables significant improvements in system efficiency. However, for successful implementation, operational and social challenges must be overcome, and circular logistics practices must be integrated into waste management systems. The future of circular logistics within the supply chain represents a key step towards more sustainable and ecologically friendly business operations. As global pressures to reduce waste and CO_2 emissions increase, more and more organizations are adopting the principles of the circular economy to improve the efficiency of their operations, reduce their ecological footprint, and create added value. These trends include optimizing recycling, reusing materials and products, and integrating closed loops into logistics operations. Given global trends and the need to reduce ecological footprints, circular logistics will become an increasingly important component of successful business models in the future. Implementation can help organizations adapt to these changes and achieve long-term benefits for their operations and the environment. The research questions and hypotheses highlight the importance of understanding these elements to achieve success in circular logistics and allow for further research in this field. However, this study has several limitations. First, the research relies primarily on secondary data and non-financial company reports, which may limit the comprehensiveness and generalizability of the findings. Second, the focus on selected case examples may not fully capture the diversity of circular logistics practices across industries and regions. Third, the analysis concentrates on the period from 2019 to 2024, so trends outside this timeframe were not explored. Future research could address these limitations by conducting primary data collection through surveys, interviews, or field studies to capture a broader range of organizational experiences with circular logistics. Comparative studies across different industries and countries could provide deeper insights into contextual factors affecting implementation. Additionally, longitudinal studies could examine the long-term impacts of circular logistics adoption on both operational efficiency and sustainability outcomes. Research could also explore the integration of emerging technologies, such as artificial intelligence, blockchain, and IoT, in greater depth to evaluate their potential for optimizing circular supply chains. Circular logistics represents a critical step toward more sustainable and ecologically responsible business operations. Its implementation can help organizations adapt to increasing environmental pressures, reduce waste and CO_2 emissions, and create long-term value. By addressing the research questions and testing the hypotheses, this study contributes to a better understanding of the drivers, challenges, and success factors in circular logistics, while highlighting avenues for further investigation to support sustainable and profitable supply chains in the future.

5 REFERENCES

[1] Geissdoerfer, M., Savaget, P., Bocken, N. M. P., & Hultink, E. J. (2017). The Circular Economy – A new sustainability paradigm? *Journal of Cleaner Production, 143*, 757–768. https://doi.org/10.1016/j.jclepro.2016.12.048

[2] Genovese, A., Acquaye, A. A., Figueroa, A., & Koh, S. C. L. (2017). Sustainable supply chain management and the transition towards a circular economy: Evidence and some applications. *Omega, 66*, 344–357. https://doi.org/10.1016/j.omega.2015.05.015

[3] Alkandi, I., Alhajri, N., & Alnajim, A. (2025). Green Supply Chain Management, Business Performance, and Future Challenges: Evidence from Emerging Industrial Sector. *Sustainability, 17*(1), 1–36. https://doi.org/10.3390/su17010029

[4] Rosen, M. A. (2017). The Future of Sustainable Development: Welcome to the European Journal of Sustainable Development Research. *European Journal of Sustainable Development Research, 1*(1), 1–2. https://doi.org/10.20897/ejosdr.201701

[5] Stahel, W. R. (2019). *The Circular Economy*. London: Routledge. https://doi.org/10.4324/9780429259203

[6] Erdelez, A., Margeta, J., & Knezić, S. (2007). Integralni pristup upravljanju sustavom prikupljanja komunalnog otpada. *Građevinar, 59*(6), 505–516. (in Croatian)
[7] Wright, R. E., Richey, R. G., & Tokman, M. (2011). Recycling and reverse logistics. *Journal of Applied Business and Economics, 12*(5), 9–20.
[8] Harris, J. M. (2003). Sustainability and Sustainable Development. *International Society of Ecological Economics*, 1–12.
[9] Krpan, L., Furjan, M., & Maršanić, R. (2014). Potencijali logistike povrata u maloprodaji. *Tehnički glasnik, 8*(2), 182–191. (in Croatian)
[10] Waters, D. (2003). *Logistics: An Introduction to Supply Chain Management*. Palgrave Macmillan.
[11] De Brito, M. P., Rommert, D., & Simme Douwe, P. F. (2005). Reverse logistics: A review of case studies. In *Distribution logistics* (pp. 243–281). Springer. https://doi.org/10.1007/978-3-642-17020-1_13
[12] Awomeso, J. A., Taiwo, A. M., Gbadebo, A. M., & Arimoro, A. (2010). Waste disposal and pollution management in urban areas: a workable remedy for the environment in developing countries. *American Journal of Environmental Sciences, 6*(1), 26–32. https://doi.org/10.3844/ajessp.2010.26.32
[13] Quesada, I. F. (2003). The concept of reverse logistics. A review of literature. In *Annual Conference for Nordic Researchers in Logistics*, 1–15.
[14] Rubio, S., & Beatriz, J.-P. (2014). Reverse Logistics: Overview and challenges for supply chain management. *International Journal of Engineering Business Management, 6*, 6–12. https://doi.org/10.5772/58826
[15] Ghiani, G., Gilbert, L., & Musmanno, R. (2004). *Introduction to Logistics Systems Planning and Control*. John Wiley & Sons. https://doi.org/10.1002/0470014040
[16] Grum, Đ. (2013). Gospodarenje ambalažnim otpadom i zaštita okoliša. *Sigurnost, 55*(1), 37–44. (in Croatian)
[17] Marinela, B. (2009). The Concept of Waste Management. *Analele Universităţii din Oradea, Fascicula: Protecţia Mediului, 14*, 669–673.
[18] Rogers, D. L. D. M., Kelly, L. C., & Garcia-Dastugue, S. J. (2002). The returns management process. *The International Journal of Logistics Management, 13*(2), 1–18. https://doi.org/10.1108/09574090210806397
[19] Valle, P. O., Menezes, J., Reis, E., & Efigenio, R. (2009). Reverse logistics for recycling: The customer service determinants. *International Journal of Business Science and Applied Management, 4*(2), 1–17. https://doi.org/10.69864/ijbsam.4-1.35
[20] Apaydin, Ö., & Gonullu, M. T. (2007). Route optimization for solid waste collection: Trabzon (Turkey) case study. *Global NEST Journal, 9*(1), 6–11. https://doi.org/10.30955/gnj.000388
[21] Sulemana, A., Donkor, E. A., Forkuo, E. K., & Oduro-Kwarteng, S. (2018). Optimal Routing of Solid Waste Collection Trucks: A Review of Methods. *Journal of Engineering, 2018*, 1–8. https://doi.org/10.1155/2018/4586376
[22] Asadi, S. (2011). Logistics system: Information and communication technology. In *Logistics Operations and Management: Concepts and Models* (pp. 221–245). https://doi.org/10.1016/B978-0-12-385202-1.00012-8
[23] Srivastava, S. K., & Srivastava, R. K. (2006). Managing product returns for reverse logistics. *Journal of Physical Distribution & Logistics Management, 36*(7), 524–546. https://doi.org/10.1108/09600030610684962
[24] Kučar Dragičević, S., Butuči, J., & Kufrin, J. (2006). Zbrinjavanje otpada u Republici Hrvatskoj - postojeće stanje. *Arhiv za higijenu rada i toksikologiju, 57*(3), 263–266. (in Croatian)
[25] Potter, C. (2022). *The Circular Economy: The Next Step in Sustainable Living*. Laurence King Publishing.
[26] Frosch, R. A., & Gallopoulos, N. E. (1989). Strategies for manufacturing. *Scientific American, 261*(3), 144–153. https://doi.org/10.1038/scientificamerican0989-144
[27] Ellen, F. M. (2015). *Towards a Circular Economy: Business for an Accelerated Transition*. Ellen MacArthur Foundation.
[28] Ghisellini, P., Cialani, C., & Ulgiati, S. (2016). A review on circular economy: the expected transition to a balanced interplay of environmental and economic systems. *Journal of Cleaner Production, 114*, 11–32. https://doi.org/10.1016/j.jclepro.2015.09.007
[29] Kshetri, N. (2018). Blockchain's roles in meeting key supply chain management objectives. *International Journal of Information Management*, 39, 80–89. https://doi.org/10.1016/j.ijinfomgt.2017.12.005
[30] Queiroz, M. M., Telles, R., & Bonilla, S. H. (2019). Blockchain and supply chain management integration: A systematic review of the literature. *Supply Chain Management: An International Journal, 25*(2), 241–254. https://doi.org/10.1108/SCM-03-2018-0143
[31] Farouk, M., & Darwish, S. M. (2020). Reverse Logistics Solution in e-Supply Chain Management by Blockchain Technology. *Egyptian Computer Science Journal, 44*(1), 22–34.
[32] Shah, S., Dikgang, G., & Menon, S. (2019). The Global Perception of Industry 4.0 for Reverse Logistics. *International Journal of Economics and Management, 4*, 103–107.
[33] Shah, S., Rutherford, R., & Menon, S. (2020). Emerging Technologies of IoT Usage in Global Logistics. In *International Conference on Computation, Automation and Knowledge Management (ICCAKM)*, Dubai, United Arab Emirates. https://doi.org/10.1109/ICCAKM46823.2020.9051530
[34] De Beers Group. (2022). *De Beers Group Introduces World's First Blockchain-Backed Diamond Source Platform at Scale*. Retrieved from https://www.debeersgroup.com/media/company-news/2022/de-beers-group-introduces-worlds-first-blockchain-backed-diamond-source-platform-at-scale (Accessed 18 September 2024).
[35] Saberi, S., Kouhizadeh, M., Sarkis, J., & Shen, L. (2018). Blockchain technology and its relationships to sustainable supply chain management. *International Journal of Production Research, 57*(7), 2117–2135. https://doi.org/10.1080/00207543.2018.1533261
[36] Lacy, P., & Rutqvist, J. (2016). *Waste to Wealth: The Circular Economy Advantage*. Springer. https://doi.org/10.1057/9781137530707
[37] IKEA. (2020). *IKEA*. Retrieved from https://www.ikea.com/us/en/this-is-ikea/ (Accessed 18 September 2024).
[38] H&M Group. (2022). *H&M Group Sustainability Disclosure 2021*. Retrieved from https://hmgroup.com/wp-content/uploads/2022/03/HM-Group-Sustainability-Disclosure-2021.pdf (Accessed 18 September 2024).
[39] Seuring, S., & Muller, M. (2008). Core issues in sustainable supply chain management – A Delphi study. *Business Strategy and the Environment, 17*(8), 455–466. https://doi.org/10.1002/bse.607
[40] ASOS. (2022). *ASOS Annual Report 2022*. Retrieved from https://asos-12954-s3.s3.eu-west-2.amazonaws.com/files/1816/6807/7015/ASOS_AR2022_Interactive_3November.pdf (Accessed 18 September 2024).

[41] ASOS. (2024). *ASOS Returns Policy*. Retrieved from https://www.asos.com/payments-and-deliveries/returns/?country=hr (Accessed 18 September 2024).
[42] Rogers, D. S., & Tibben-Lembke, R. (2001). An examination of reverse logistics practices. *Journal of Business Logistics, 22*(2), 129–148. https://doi.org/10.1002/j.2158-1592.2001.tb00007.x
[43] Amazon. (2023). *Amazon Introduces New Robotics Solutions*. Retrieved from https://www.aboutamazon.com/news/operations/amazon-introduces-new-robotics-solutions (Accessed 18 September 2024).
[44] Waller, M. A., & Fawcett, S. E. (2013). Data Science, Predictive Analytics, and Big Data: A Revolution That Will Transform Supply Chain Design and Management. *Journal of Business Logistics, 34*(2), 77–84. https://doi.org/10.1111/jbl.12010
[45] Apple. (2024). *Apple Environmental Progress Report 2024*. Retrieved from https://www.apple.com/environment/pdf/Apple_Environmental_Progress_Report_2024.pdf (Accessed 18 September 2024).
[46] Bocken, N. M. P., Short, S. W., Rana, P., & Evans, S. (2014). A literature and practice review to develop sustainable business model archetypes. *Journal of Cleaner Production, 65*, 42–56. https://doi.org/10.1016/j.jclepro.2013.11.039
[47] Govidan, K., Soleimani, H., & Kannan, D. (2015). Reverse Logistics and Closed-Loop Supply Chain: A Comprehensive Review to Explore the Future. *European Journal of Operational Research, 240*(3), 603–626. https://doi.org/10.1016/j.ejor.2014.07.012
[48] Kirchherr, J., Reike, D., & Hekkert, M. (2017). Conceptualizing the circular economy: An analysis of 114 definitions. *Resources, Conservation and Recycling, 127*, 221–232. https://doi.org/10.1016/j.resconrec.2017.09.005

Authors' contacts:

Dora Kolaric
(Corresponding author)
Sveučilište Sjever,
Trg dr. Žarka Dolinara 1, 48000 Koprivnica, Croatia
dokolaric@unin.hr

Ivan Grgačić
Sveučilište Sjever,
Trg dr. Žarka Dolinara 1, 48000 Koprivnica, Croatia
ivgrgacic@unin.hr

Matej Stoprd
Sveučilište Sjever,
Trg dr. Žarka Dolinara 1, 48000 Koprivnica, Croatia
mastoprd@unin.hr

Dunja Horvat
Sveučilište Sjever,
Trg dr. Žarka Dolinara 1, 48000 Koprivnica, Croatia
dunjhorvat@unin.hr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250101230414

Subject review
Received: 2025-01-01, Accepted: 2025-04-29

Comparative Evaluation of PLA and PETG's Mechanical Characteristics for 3D-Printed Bone Implant Applications

Reem Sabah Khzaal*, Wisam Khdim Hamdan, Marwan Nafaa Ali

Abstract: This study examines the mechanical properties of polylactic acid (PLA) and glycol-modified polyethylene terephthalate (PETG) in relation to 3D-printed cranial implants. Using fused deposition modelling (FDM) technology, the materials were tested for tensile, bending, and impact toughness under various printing conditions (layer thickness, infill density, and infill pattern). The findings indicated that PETG exhibited superior impact toughness and flexural strength, making it ideal for load-bearing applications, while PLA demonstrated lower tensile strength, making it suitable for temporary implants. The study highlights the literature background on PETG's resilience and processing efficiency along with PLA's biodegradability and bio absorption to satisfy specific medicinal needs. These findings demonstrate how 3D printing can be utilized to produce patient-specific, biocompatible implants, increasing precision and lowering costs in biomedical applications. The improvement of material properties and printing parameters provides the foundation for improving the functionality of implants tailored to particular requirements.

Keywords: Bone Implant; Glycol-modified Polyethylene Terephthalate (PETG); Mechanical Properties; Polylactic-Acid (PLA); 3D-Printer

1 INTRODUCTION

Three distinct materials (polyethylene terephthalate glycol - PETG, white resin, and Bone STN) were used to create 3D-printed skull samples using three distinct 3D print techniques (fused filament fabrication, stereo lithography [SLA], and material jetting). This is the first study to assess the safety of 3D-printed materials for burr hole surgical drilling, as far as the authors are aware. While resin-based materials will need safety controls to deal with airborne particles, PETG is recommended as a safe material [1]. polylactic acid (PLA) and polyethylene terephthalate glycol (PETG) specimens with Raster angles of ±45° were used for the infill in order to achieve a 100% infill density. A layer thickness of 0.15 mm was considered for printing. There were four thicknesses for each material. According to the ISO standard, a pendulum impact system was used to test the specimens. All specimens, regardless of thickness, showed brittle fracture in addition to comparable behavior in the elastic zone of deflection for the same thickness and material type. If the PETG specimens allow for a larger deflection, the PLA specimens exhibit a higher impact force; however, the IZOD impact strength values for the two materials are comparable [2]. Polyethylene terephthalate glycol (PETG) is a transparent, miscible thermoplastic have strong tensile qualities, high ductility, and resistance to heat and chemical exposure. There are several methods for producing PETG, but 3D printing is one of the most prominent. Although polylactic acid (PLA) is more biodegradable than PETG, because of its high tensile strength and high degradation temperature, PETG is a unique thermoplastic that is both economical and efficient. Tissue engineering, dentistry, optometry, vascular health, cardiology, orthopedics, neurology, gynecology, and surgery are just a few of the medical specialties that use it because of its advantageous qualities. Furthermore, PETG contributes significantly to biomedical engineering and research, advancing cell studies, it is expected that more medical research and development will solidify PETG's position as an affordable, adaptable biomaterial with a broad range of applications [3]. The use of polylactic acid (PLA) implants customized through 3-dimensional (3D) printing as an affordable and precise cosmetic option for cranioplasty was performed on two cases of skull defects. A 20-year-old woman with a history of severe traumatic brain injury (TBI) and epidural hematoma was the subject of the first case. The second case present an elderly man, 46, whose had previously experienced spontaneous intracerebral hemorrhage associated with arteriovenous malformation (AVM). A 3D model of the skull defect was created by reconstructing both of the CT scan data. A 3D printer was used to create the prosthesis using PLA. In both cases, cosmetic satisfaction was achieved and no postoperative complications were reported. outcome with no complications reported after the cranioplasty procedure, the use of PLA implants customized through 3D printing was shown to be economical and to provide good cosmetic satisfaction [4]. Fourteen distinct infill patterns were used to 3D print materials made of polylactic acid (PLA), enhanced polylactic acid (PLA+), and polyethylene terephthalate glycol (PETG). The printed specimens' mechanical characteristics and surface characteristics were evaluated and discussed. The mechanical properties of the produced specimens were determined to be ultimate tensile strengths, Young's modulus, and strain at break percentage, while the average, maximum, and total height of protuberances (*Ra*, *Rz*, and *Rt*) were measured as surface characteristics. On the other hand, the best *Ra*, *Rz*, and *Rt* were obtained from cross, quarter cubic, and concentric patterns. The cubic, gyroid, and concentric patterns were found to be the best overall patterns for the mechanical properties of PLA, PLA+, and PETG materials [5].

The specimen with 100% filling showed more deformation in the compressive strength test than the specimen with 50% filling. This is explained by the PETG polymer's highly ductile mechanical properties. The PLA polymers with the highest filling percentages (100 and 50 percent) demonstrated less deformation until failure in the mechanical flexural strength tests, indicating that they were more ductile materials. However, specimens that had 30% filling showed more deformation than those that had 100%

filling. Because the internal spacing absorbs the impact of loads, PLA exhibits greater flexibility and tenacity for fillings of low percentages, when compared to the PETG polymer, the PLA polymer demonstrated superior mechanical qualities, including Young's modulus, ductility, and more satisfactory resistance. Because PLA is a biopolymer and has exceptional processability, it also describes the synthesis of PLA in a more sustainable manner [6]. Use PETG in 3D printing to create intricate, patient-specific implants that are impossible to make using conventional techniques, highlighting the significance of biocompatible materials for upcoming surgical applications [7, 22]. Several 3D-printable polymers, including polyetheretherketone (PEEK), polylactic acid (PLA), polymethyl methacrylate (PMMA), glycol-modified polyethylene terephthalate (PETG), polyvinylidene fluoride (PVDF), and polypropylene (PP), were subjected to tensile tests at varying crosshead speeds in order to examine the application temperature range. It was discovered that variations in body temperature had little effect on the mechanical characteristics of PETG, PEEK, PLA, and PMMA. The materials were further ranked based on their filling density, which is a gauge of their processability, in addition to their mechanical integrity (strength, stiffness, and its strain-rate and temperature dependency inside the body). As a result, helpful information is offered for the design process of 3D-printed implants as well as for choosing potential medical uses for each material [8]. Flexural testing of the 3D-printed bone pieces shows how the mechanical behavior of the parts is affected by sample geometry and printing orientations. The results were not significantly impacted by the printing orientation due to the high infill percentage, to find out how the sectioning technique affected the overall mechanical performance. Tests for lateral compression and frontal penetration were conducted on the dummy skulls. The findings show that PLA-based surrogates may accurately replicate the behavior of the cranium [9]. Characterizing the polylactic acid (PLA) and polyethylene terephthalate glycol (PETG) materials of FDM under four loading conditions (tension, compression, bending, and thermal deformation) using a range of printing temperatures and speeds. According to the findings, there is a clear tensile and compression asymmetry in PLA and PETG materials. Higher printing temperatures were found to improve the mechanical characteristics (tension, compression, and bending) of PLA and PETG, and speed had varying effects on the two materials. Furthermore, PLA has better mechanical qualities than PETG, for thermal deformation, however, the opposite is true. These findings will be very helpful to researchers who employ polymers and FDM processes to reach sustainability [10], studied maximize the mechanical properties of printed PETG by optimizing the printing parameters (nozzle temperature, printing speed, layer height, and filling). The optimal bending characteristics for PETG were found to be achieved at extrusion temperatures of 265 °C, printing speeds of 20 mm/s, layer heights of 0.4 mm, and an internal density of 100% [11]. The successful fabrication of simple implants for defect reconstruction was the aim of the project. Depending on the material choice, digitally created three-dimensional (3D) models addressing two representative mandibular bone defects were printed. Glycol-modified PET (PETG) were ultimately chosen. This polymer offered excellent printability, good mechanical qualities, and high cell acceptance. The following manufacturing procedure produced implant strategies: the PETG standard implant because the print time for the standard implant is less than three hours, even for larger skull defects, it is intended to be applied intraoperatively. Standard implants were found to be cleanly printed, well-fitting, and mechanically stable. This bio functional material is appropriate for pre-operative manufacturing based on radiological image scans of the defect and may be utilized in certain reconstructive patient cases [12]. In order to determine the optimal mechanical properties of the PLA material, polylactic acid was used through a fused deposition process to study combinations of process parameters (layer thickness, build direction, filling speed, filler density, etc.). Among these, fill rate is the most important component in compressive property, whereas build direction significantly affects tensile, bending, and surface roughness properties [13], how Charpy impact strength is affected by print speed, layer height, raster angle, and infill density. The most important factor is infill density, which is followed by raster angle, layer height, and print speed, according to analysis using S/N ratios and ANOVA. According to ANOVA, print speed and infill density have respective influences of 36.51% and 38.93%. Impact strength is increased by 1.39% to 38.54 kJ/m² using the Taguchi method with the optimized parameters of 100% infill density, 45/−45° raster angle, 0.25 mm layer height, and 75 mm/s print speed [14]. The mechanical properties of polylactic acid (PLA) components are examined in this study, along with the impacts of layer height, fill width, fill density, layer orientation, printing speed, and fill pattern on defect resistance. Through testing on various geometries using rotating and four-point bending machines, the optimal parameter set for maximal fiber resistance was discovered. According to the results, fill pattern and density have the least impact, whereas layer orientation has the biggest impact, followed by layer height, fill density, and printing speed [15]. This study investigates the need for bone implants to be composed of safe, long-lasting, and biocompatible materials. Despite being easily accessible for 3D printing, PLA and PETG may not be suitable for medical purposes due to their unique mechanical properties. This study is significant because:

1) Material Suitability: Comparing mechanical properties including tensile strength, impact, bending, and flexibility to determine which material best suits the needs of a bone implant.
2) Biocompatibility: Assessing PLA's ability to absorb biomaterials and PETG's resilience for secure uses. This research could greatly enhance the availability of patient-specific, accessible bone implant treatments.

2 MATERIALS AND METHODS

Because of their suitability for creating cranial implants and their biocompatibility with the human body, two types of materials polylactic acid (PLA) and polyethylene terephthalate glycol (PETG) were chosen. These materials were obtained from Uruk Tech Co., based in Baghdad, Iraq,

and manufactured by Torwell Co., Ltd., China. The tensile test was conducted by a WDW-200E tensile testing machine measuring control software manufactured by People's Republic of China Time Group Inc. in the strength laboratory of the College of Production and Metallurgical Engineering, University of Technology, Baghdad, Iraq. The impact test was also performed by a Chinese-made XJU-22 impact tester to compare their mechanical properties. A FDM 3D printer was used to create filament out of these materials.

2.1 Polylactic Acid (PLA)

For bone implants, polylactic acid (PLA) is a biodegradable polymer that has been thoroughly studied due to its mechanical properties, biocompatibility, and bio absorption. Self-reinforced PLA implants demonstrate sufficient mechanical strength for bone fixation and gradual bio absorption, with a low inflammatory response and effective osteotomy healing over a 48-week period [16], [20], [21]. Because PLA improves toughness and bioactivity, it works incredibly well for bone tissue engineering [17]. This material was used in this study, and Tab. 1 shows the material parameters as provided by the manufacturer. Additionally, Fig. 1 depicts the PLA and FDM Creality K1 Max 3D printer filament.

Table 1 Description Material parameters of PLA

Material parameters	Limitation
Density	1.24 g/cm^3, ASTM D1505
Melt Flow Index	6 g/10 min, 190 °C/2.16 kg
Heat Distortion Temp	50 °C
Tensile Strength	6.08012 MPa
Tensile Elongation	4.4 %
Flexural Strength	6.47239 MPa
Flexural Modulus	2 745.862 MPa
Impact Strength	4.2 kJ/m²
Diameter Accuracy	1.75 mm : 1.70 - 1.78 mm

Table 2 Description Material parameters of PETG

Material parameters	Limitation
Density	1.27 g/cm^3, 21.5 °C
Melt Flow Index	20 g/10 min, 200 °C/5 kg
Heat Distortion Temp	70 °C
Tensile Strength	4.90332 MPa
Tensile Elongation	120 %
Flexural Strength	70.60788 MPa
Flexural Modulus	2 147.65635 MPa
Impact Strength	8.7 kJ/m²
Diameter Accuracy	1.75 mm : 1.70 - 1.78 mm

2.2 Polyethylene Terephthalate Glycol (PETG)

PETG exhibits strong mechanical qualities, biocompatibility, and adaptability for bone implants, making it a viable option for load-bearing and bioactive applications. Its use in high-stress or long-term applications is still subject to considerations regarding stress response and degradation in specific environments. When compared to other polymers, PETG scaffolds made for bone tissue engineering have greater mechanical strength and improved cell attachment and proliferation, suggesting that they are suitable for load-bearing applications [3]. Tab. 2 will show some parameters of the material produced by the manufacturer, and Fig. 2 shows the filament of PETG and the FDM Sanlu S9 Plus 3D printer machine.

2.3 3D-printing Machine

Two types of printers were used in this study, as shown in Fig. 1a. 3D printer, the CREALITY K1 MAX, was used to print PLA material Fig. 1b. The second printer, the SANLU S9 PLUS, was employed for printing PETG. Tab. 3 presents the specifications of the 3D printers used for both materials.

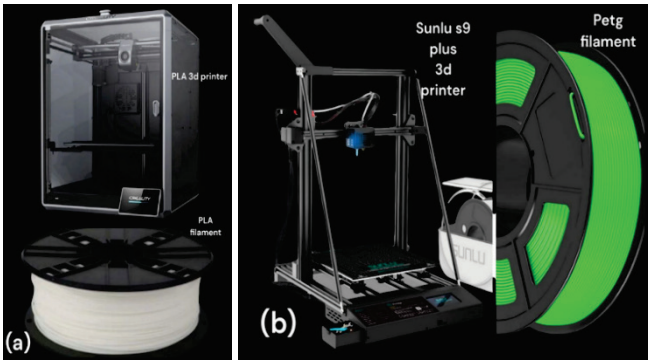


Figure 1 (a) k1 max creality 3d printer for PLA material, (b) sanlu s9 plus 3d printer for PETG material

Table 3 The specifications of the 3D printer

Specification	PLA	PETG
Brand	Creality	SANLU
Printing Technology	FDM	FDM
Product Dimensions	435×462×526 mm	309.88×309.88×400mm
Filament diameter	1.75 mm	1.75 mm
Printing Speed	≤600 mm/s	250 mm/s
printing temperature	0 - 40 °C	260 °C
Heatbed Temperature	≤100 °C	100 °C
Nozzle Diameter	0.4 mm (compatible with 0.6/0.8 mm)	0.4 mm (compatible specifications: 0.4 - 0.8 mm)

This study looked into how printing parameters, such as (infill density, infill pattern, and layer thickness) affected mechanical performance. In order to compare the performance of PLA and PETG, the parameters were assessed using mechanical tests like the Charpy impact test, bending test, and tensile test.

Table 4 Process parameters of the printing material

Printing parameters	Low level	Mid-level	High level
Layer thickness	0.15	0.225	0.3
Infill density	20 %	50 %	80 %
Infill pattern	grit	concentric	Tri-hexagonal

First, a 3D virtual geometry of the (impact, bending, tensile) test specimens was produced using SolidWorks software as shown in Fig. 2. After the virtual test specimens were created, the SolidWorks file was exported into the slicing program in.(stl) format. The same printing factors were established for both materials at three levels, as indicated in Tab. 4.

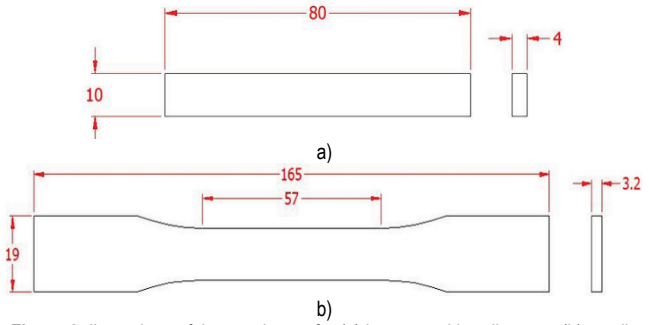


Figure 2 dimensions of the specimens for (a) impact and bending test, (b) tensile test.

2.4 The Impact, Bending, and Tensile Test

To assess the impact, bending and tensile properties of the specimens, the impact test or so-called Charpy tester with ISO179 standard, was performed using a shock device, and the specimen is placed between the two columns at the bottom of the holder. To release the paddle, it is taken and placed at a specified height, depending on the load capacity. The specimen is violently struck by the paddle as it is released down for testing, which causes it to split into two parts with an energy of 5.5 joules. The impact toughness was measured by the law in Eq. (1) [18] to find out the value of the energy absorbed for each specimen and for both materials.

$$Impact\ Toughness = \\ = \frac{Absorbed\ Energy\ (\text{Joules})}{Cross\text{-}sectional\ Area\ of\ the\ Test\ Specimen\ (\text{mm}^2)}. \quad (1)$$

The bending test was conducted on the same samples according to ISO 178 standard, where the bending test was carried out using a 3-point contact bending test device. The test sample was placed on the support and a vertical force was applied to the sample at a rate of (40 N) and at a speed of (2 mm/min) to measure the maximum force that the sample can withstand before failure to measure the value of the flexural strength according to Eq. (2) [19] shown below.

$$\sigma = \frac{3PL}{2bd^2}, \quad (2)$$

where L is the length of the support span, b is the specimen's width, d is its thickness, and P is the load at a specific location on the load-deflection curve.

Lastly, tensile samples with the dimensions displayed in Fig. 2b were made in compliance with ASTM 638. The maximum tensile strength and ductility values of the two materials were ascertained by testing these samples with a tensile testing device at a speed of 1 mm/min and a load of 40 N.

3 EXPERIMENTAL DESIGN USING THE BOX-BEHNKEN DESIGN BY RESPONSE SURFACE METHODOLGY

The Box-Behnken design (BBD) is a response surface method for optimizing experimental designs with many variables. In this study, three factors—layer thickness, infill density, and infill pattern—that significantly affect the mechanical properties and functionality of 3D-printed components are examined. BBD is ideal for examining the interactions between these variables because it requires fewer experimental runs while maintaining efficiency compared to a full factorial design. The design generates a set of combinations by systematically varying the levels of these parameters in order to investigate their effects both individually and in combination. With this method, the printing process can be precisely optimized for desired results, like strength, durability, or material efficiency. Finally, based on three levels for each factor, the Box-Behnken design designed 15 experiments shown in Tab. 5.

Table 5 Design of experimental of process parameters

Run	Coded matrix			Un coded matrix		
	A	***B***	***C***	Layer thickness	Infill density	Infill pattern
1	–1	–1	0	0.15	20	Concentric
2	1	–1	0	0.3	20	Concentric
3	–1	1	0	0.15	80	Concentric
4	1	1	0	0.3	80	Concentric
5	–1	0	–1	0.15	50	Grid
6	1	0	–1	0.3	50	Grid
7	–1	0	1	0.15	50	Tri-hexagonal
8	1	0	1	0.3	50	Tri-hexagonal
9	0	–1	–1	0.225	20	Grid
10	0	1	–1	0.225	80	Grid
11	0	–1	1	0.255	20	Tri-hexagonal
12	0	1	1	0.225	80	Tri-hexagonal
13	0	0	0	0.225	50	Concentric
14	0	0	0	0.225	50	Concentric
15	0	0	0	0.225	50	Concentric

4 RESULTS AND DISCUSSION

In this study, 15 experiments were conducted using identical and controlled printing conditions (layer thickness, fill density, and fill shape) to compare the PLA and PETG's mechanical characteristics. The findings show clear variations in the two materials' performance attributes, emphasizing both their advantages and disadvantages when manufactured under the same conditions.

For both materials PLA and PETG used in orthopaedic implant analysis, impact test images were obtained for 15 specimens under the same test conditions and are shown in Fig. 3. The purpose of these tests was to ensure that the results for both materials were comparable and consistent.

Experimental findings for PLA and PETG materials' impact toughness under various 3D printing scenarios are shown in Tab. 6. An analysis of the findings is provided below. Eq. (3) and Eq. (4) represent the mathematical model for the response variable, impact toughness for PLA and PETG respectively.

$$impact\ toghness = 0.276 - 0.79\ layer\ thickness + 0.00356\ infill\ density - 0.0396\ infill\ pattern + 3.19\ layer\ thickness * layer\ thickness + 0.000006\ infill\ density * infill\ density - 0.0071\ infill\ pattern * infill\ pattern - 0.0111\ layer\ thickness * infill\ density + 0.167\ layer\ thickness * infill\ pattern - 0.000833\ infill\ density * infill\ pattern \quad (3)$$

$$\textit{impact toughness} = 0.0345 - 0.241\ \textit{layer thickness} + 0.000152\ \textit{infill density} + 0.01911\ \textit{infill pattern} + 0.635\ \textit{layer thickness} * \textit{layer thickness} - 0.000002\ \textit{infill density} * \textit{infill density} + 0.00782\ \textit{infill pattern} * \textit{infill pattern} + 0.00033\ \textit{layer thickness} * \textit{infill density} - 0.0792\ \textit{layer thickness} * \textit{infill pattern} + 0.000083\ \textit{infill density} * \textit{infill pattern} \quad (4)$$

Table 6 printing parameters and impact toughness of PLA and PETG material.

Runs	Parameters			Impact toughness of PLA (J/mm^2)	Impact toughness of PETG (J/mm^2)
	Layer thickness	Infill density	Infill pattern		
1	0.15	20	Concentric	0.00625	0.015
2	0.3	20	Concentric	0.00875	0.0225
3	0.15	80	Concentric	0.01	0.0175
4	0.3	80	Concentric	0.01	0.028
5	0.15	50	Grid	0.01	0.01625
6	0.3	50	Grid	0.00875	0.0375
7	0.15	50	Tri-hexagonal	0.0075	0.035
8	0.3	50	Tri-hexagonal	0.0076	0.0325
9	0.225	20	Grid	0.0078	0.02
10	0.225	80	Grid	0.01125	0.015
11	0.255	20	Tri-hexagonal	0.00625	0.03
12	0.225	80	Tri-hexagonal	0.0075	0.035
13	0.225	50	Concentric	0.0076	0.02375
14	0.225	50	Concentric	0.0071	0.0125
15	0.225	50	Concentric	0.01	0.0205

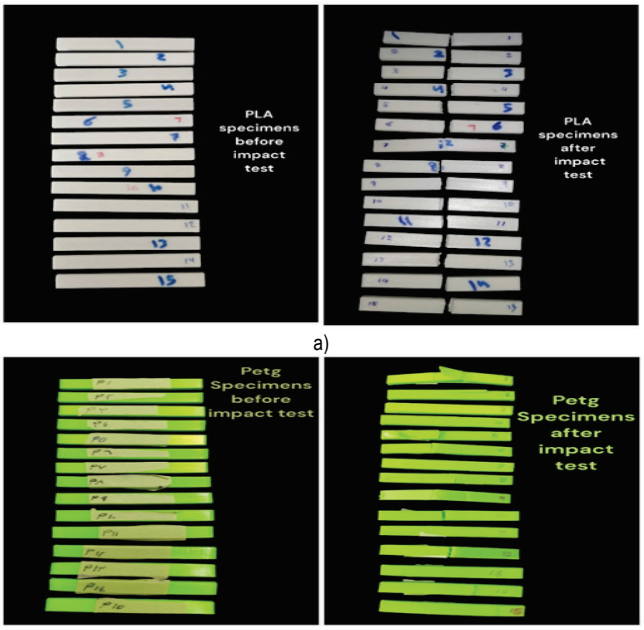


Figure 3 impact test for (a) polylactic acid specimens, (b) polyethylene terephthalate glycol specimens

When compared to PLA, PETG continuously shows greater impact toughness in the same circumstances. For instance: Run 1: PETG (0.015) versus PLA (0.00625) PLA (0.00875) versus PETG (0.0375) in Run 6 This demonstrates the superior energy absorption capabilities of PETG. Differences Particular to Patterns: PETG reacts favorably to the tri-hexagonal pattern, while PLA appears to gain more from the Concentric and Grid patterns. Material Selection for applications needing greater impact toughness, PETG is typically a better option. Parameter Optimization, Impact resistance can be increased for both materials by carefully choosing infill patterns and optimizing infill density. PLA benefits from thinner layers, while PETG can withstand or even benefit from thicker layers. A spider chart (radar chart) comparing the impact toughness of PETG (orange) and PLA (blue) is shown in Fig. 4.

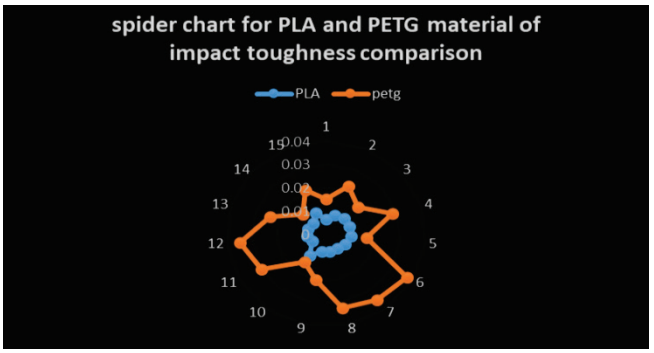


Figure 4 comparing the impact toughness of PETG and PLA

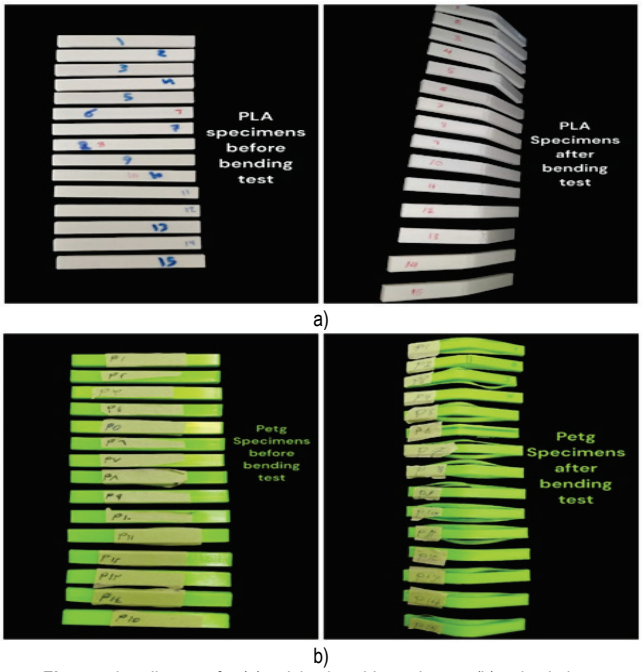


Figure 5 bending test for (a) polylactic acid specimens, (b) polyethylene terephthalate glycol specimens

In most regions, the orange polygon (PETG) is farther from the center than the blue polygon (PLA). This suggests that PETG has greater impact toughness across a range of parameters. Features of PLA: The blue polygon is more centralized and smaller. This illustrates PLA's limited capacity to absorb impact energy and brittle nature. PETG Features: The orange polygon is more flexible and energy absorption due to its larger coverage area. Compared to PLA, PETG is less prone to fracture under dynamic loads. This has been observed during laboratory experiments because it has greater ductility than PLA. Finally, the spider diagram effectively depicts the superior performance of PETG over

PLA in terms of impact strength. The ductility and toughness of PETG make it a better choice for applications that require resistance to mechanical stress, such as orthopedic implants, while PLA is better suited for applications where stiffness is more important than toughness.

The PLA and PETG bending samples are shown in Fig. 5 following testing under the same printing conditions listed in Tab. 7.

Table 7 Flexural strength of PLA and PETG material specimens.

Runs	Parameters			Flexural strength of PLA (MPa)	Flexural strength of PETG (MPa)
	Layer thickness	Infill density	Infill pattern		
1	0.15	20	Concentric	50	93
2	0.3	20	Concentric	54	72
3	0.15	80	Concentric	72	63
4	0.3	80	Concentric	59	64
5	0.15	50	Grid	48	45
6	0.3	50	Grid	45	47
7	0.15	50	Tri-hexagonal	54	73
8	0.3	50	Tri-hexagonal	63	54
9	0.225	20	Grid	59	36
10	0.225	80	Grid	54	55
11	0.255	20	Tri-hexagonal	53	50
12	0.225	80	Tri-hexagonal	50	70
13	0.225	50	Concentric	55	48
14	0.225	50	Concentric	59	65
15	0.225	50	Concentric	53	68

The PLA and PETG materials' bending properties (flexural strength) were measured and contrasted over 15 different runs, as seen in Fig. 6.

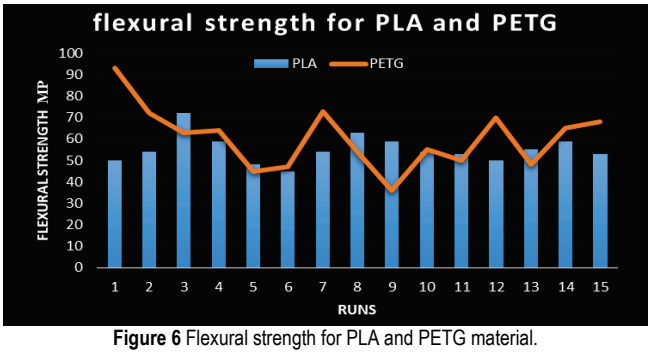


Figure 6 Flexural strength for PLA and PETG material.

The PLA's flexural strength values vary slightly but are generally constant throughout the runs. The PETG's flexural strength fluctuates more. Beginning at a high value, it declines over a number of runs before rising once more in the later runs. In most runs, PETG exhibits greater flexural strength than PLA, suggesting better bending properties under the same circumstances. Nonetheless, PETG's performance fluctuation may indicate sensitivity to printing or testing circumstances as in run 9 and others shown in Tab. 7. A Box–Behnken design was used to create a mathematical model that better explains how print parameters affect flexural strength. The second-order regression model that is produced in Eq. (5) and Eq. (6) for PLA and PETG material, respectively

$$\textit{flexural strength}\ (\text{MPa}) = 45.1 + 26\ \textit{layer thickness} + 0.250\ \textit{infill density} - 8.1\ \textit{infill pattern} + 141\ \textit{layer thickness} * \textit{layer thickness} + 0.00255\ \textit{infill density} * \textit{infill density} - 3.96\ \textit{infill pattern} * \textit{infill pattern} - 1.89\ \textit{layer thickness} * \textit{infill density} + 40.0\ \textit{layer thickness} * \textit{infill pattern} + 0.017\ \textit{infill density} * \textit{infill pattern} \quad (5)$$

$$\textit{flexural strength}\ (\text{MP}) = 182.3 - 771\ \textit{layer thickness} - 1.14\ \textit{infill density} + 23.3\ \textit{infill pattern} + 1304\ \textit{layer thickness} * \textit{layer thickness} + 0.00593\ \textit{infill density} * \textit{infill density} - 12.92\ \textit{infill pattern} * \textit{infill pattern} + 2.44\ \textit{layer thickness} * \textit{infill density} - 70.0\ \textit{layer thickness} * \textit{infill pattern} + 0.008\ \textit{infill density} * \textit{infill pattern} \quad (6)$$

PETG is better suited for applications needing greater mechanical strength, like bone implants, due to its higher flexural strength. PLA is appropriate for other areas of medical use where moderate mechanical properties are acceptable, but its consistent but lower flexural strength may make it less suitable for load-bearing applications.

Fig. 7 displays pictures of the PLA and PETG samples following the tensile test. Both materials' ultimate tensile strength was assessed using the same printing parameters as the other tests. Tab. 8 provides a summary of the related test results.

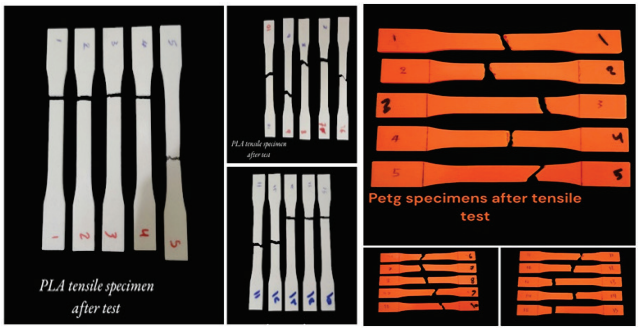


Figure 7 Tensile specimens after testing for PLA and PETG material.

Table 8 Ultimate tensile strength values of PLA and PETG material specimens.

Runs	Parameters			Ultimate tensile strength for PLA (MPa)	Ultimate tensile strength for PETG (MPa)
	Layer thickness	Infill density	Infill pattern		
1	0.15	20	Concentric	28.07	28.26
2	0.3	20	Concentric	32.69	30.38
3	0.15	80	Concentric	45.57	46.34
4	0.3	80	Concentric	39.61	39.8
5	0.15	50	Grid	29.03	27.69
6	0.3	50	Grid	32.11	29.23
7	0.15	50	Tri-hexagonal	32.12	33.65
8	0.3	50	Tri-hexagonal	32.11	32.3
9	0.225	20	Grid	25.76	25.76
10	0.225	80	Grid	33.26	34.03
11	0.255	20	Tri-hexagonal	28.65	27.88
12	0.225	80	Tri-hexagonal	37.11	41.34
13	0.225	50	Concentric	34.61	35
14	0.225	50	Concentric	34.42	36.53
15	0.225	50	Concentric	35	36.15

A bar graph comparing PLA and PETG's Ultimate Tensile Strength (UTS) over 15 distinct runs or samples is shown in Fig. 8. Variability in UTS across the various runs is seen in both PLA and PETG. A number of variables, including testing conditions, material composition, and manufacturing procedures, and interference with printing agents, as shown in the quadratic regression equation Eq. (7) for PLA and Eq. (8) for PETG, may be responsible for this variation. may be responsible for this variation. Generally speaking, PETG slightly higher UTS than PLA. PLA and PETG comparison PETG seems to be a stronger material than PLA, according to the UTS data. In some bone implant applications where mechanical stability and load-bearing capacity are essential, this increased strength might be beneficial.

$$UTS = 23.50 - 16.6\ layer\ thickness + 0.340\ infill\ density + 3.15\ infill\ pattern + 174\ layer\ thickness * layer\ thickness + 0.000923\ infill\ density * infill\ density - 4.312\ infill\ pattern * infill\ pattern - 1.176\ layer\ thickness * infill\ density - 10.3\ layer\ thickness * infill\ pattern + 0.0080\ infill\ density * infill\ pattern \quad (7)$$

$$UTS = 13.40 + 90.4\ layer\ thickness + 0.3196\ infill\ density + 2.31\ infill\ pattern - 110\ layer\ thickness * layer\ thickness + 0.001020\ infill\ density * infill\ density - 4.559\ infill\ pattern * infill\ pattern - 0.962\ layer\ thickness * infill\ density - 9.63\ layer\ thickness * infill\ pattern + 0.0432\ infill\ density * infill\ pattern \quad (8)$$

PETG consistently exhibits high tensile strength (UTS) in the majority of experiments across various printing parameters, indicating its strong potential for orthopaedic implant applications where mechanical strength is of primary concern, according to the data shown in the Tab. 8. However, concerns about biocompatibility and biodegradability require further study as the addition of other materials improves the material and increases its biodegradability. Since PLA is biocompatible and biodegradable, it is still a good choice for temporary implants, even though it is weaker.

5 CONCLUSIONS

The study focuses on the suitability of PLA and PETG for orthopedic implant applications by highlighting the mechanical performance differences between the two materials under identical 3D printing conditions. When compared to PLA, PETG continuously exhibits higher impact toughness, flexural strength, and slightly higher ultimate tensile strength (UTS). Because of these benefits, PETG is a viable option for applications like bone implants that call for load-bearing capacity and resistance to mechanical stress. PETG is less likely to crack under dynamic loads due to its superior energy absorption, flexibility, and toughness, as demonstrated by the spider chart. The ability of PETG to withstand thicker layers and its ideal infill patterns particularly the tri-hexagonal design further enhances its impact resistance.

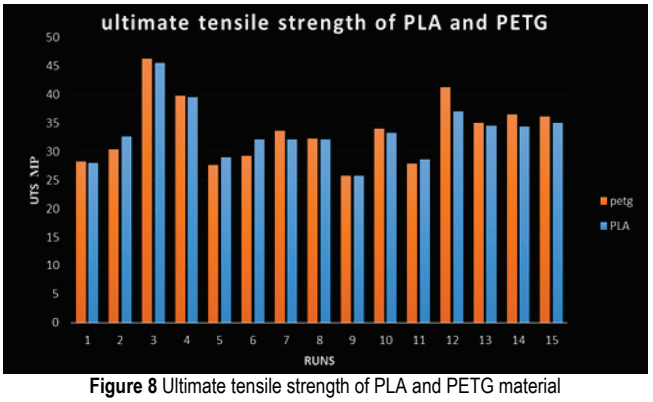


Figure 8 Ultimate tensile strength of PLA and PETG material

Despite its worse mechanical performance, PLA has consistent flexural strength, thinner layers, and specific infill patterns like Concentric and Grid. Its low energy absorption capacity and brittleness are two major disadvantages. However, due to its biocompatibility and biodegradability, PLA is a good choice for temporary implants where stiffness is more crucial than toughness. Changes in PETG's flexural strength reveal how sensitive the material is to testing and printing circumstances. This suggests that precise parameter control is necessary when using PETG for important applications. PLA's lower strength and impact resistance limit its use in demanding environments, despite its generally consistent performance.

When all is said and done, PETG turns out to be a more durable and stronger material that performs admirably in load-bearing and mechanical stress applications, such as orthopedic implants. Its biocompatibility and biodegradability need to be further investigated in order to increase its suitability for medical applications. For short-term medical applications, PLA is still a sustainable and biocompatible option. The performance of both materials can be optimized by modifying printing parameters, ensuring their effective use in a range of applications.

6 REFERENCES

[1] Dissanayaka, N., Alexander, H., Carluccio, D., Redmond, M., Vandi, L.-J., & Novak, J. I. (2024). How safe are 3D-printed skull models for neurosurgical simulation? Measurement of airborne particles and VOCs while burr hole drilling. *Rapid Prototyping Journal, 30*(5), 1046-1054. https://doi.org/10.1108/RPJ-09-2023-0318

[2] Popa, C.-F., Mărghitaș, M.-P., Galațanua, S.-V., & Marșavina, L. (2022). Influence of thickness on the IZOD impact strength of FDM printed specimens from PLA and PETG. *Procedia Struct. Integr., 41*(6), 557-563. https://doi.org/10.1016/j.prostr.2022.05.064

[3] Hassan, M. H., Omar, A. M., Daskalakis, E., Hou, Y., Huang, B., Strashnov, I., Grieve, B. D., & Bártolo, P. (2020). The potential of polyethylene terephthalate glycol as biomaterial for bone tissue engineering. *Polymers, 12*(12), p. 3045. https://doi.org/10.3390/polym12123045

[4] Apriawan, T., Permana, K. R., Darlan, D., Arifianto, M. R., Fitra, F., Alfauzi, A., & Bajamal, A. H. (2020). Polylactic acid implant for cranioplasty with 3-dimensional printing customization: A case report. *Open Access Macedonian Journal of Medical Sciences*, *8*(T1), 151-155.

https://doi.org/10.3889/oamjms.2020.5156
[5] Kadhum, A. H., Al-Zubaidi, S., & Abdulkareem, S. S. (2023). Effect of the infill patterns on the mechanical and surface characteristics of 3D printing of PLA, PLA+ and PETG materials. *ChemEngineering*, *7*(46), 2-35. https://doi.org/10.3390/chemengineering7030046
[6] da M. Dantas, G., & dos Santos, D. F. (2023). Effect of the degree of filling on mechanical properties of polymeric specimens from polyethylene terephthalate glycol and polylactic acid produced by 3D printing. *Key Engineering Materials*, *7*, 7-18. https://doi.org/10.4028/p-0Vk0vf
[7] Neamaha, Z. H., Al-Kindia, L. A., & Al-Kindib, G. (2023). Design and manufacturing of custom 3D printed bone implants. *Engineering and Technology Journal*, *6*, 665-676. https://doi.org/10.30684/etj.2023.143788.1609
[8] Petersmann, S., Spoerk, M., Van De Steene, W., Üçal, M., Wiener, J., Pinter, G., & Arbeiter, F. (2020). Mechanical properties of polymeric implant materials produced by extrusion-based additive manufacturing. *Journal of the Mechanical Behavior of Biomedical Materials*, *104*, 103611. https://doi.org/10.1016/j.jmbbm.2019.103611
[9] Mantecón, R., Marco, M., Muñoz-Sanchez, A., Youssef, G., Díaz-Álvarez, J., & Miguélez, H. (2023). Additive manufacturing and mechanical characterization of PLA-based skull surrogates. *Polymers*, *15*(1), 58. https://doi.org/10.3390/polym15010058
[10] Hsueh, M.-H., Lai, C.-J., Wang, S.-H., Zeng, Y.-S., Hsieh, C.-H., Pan, C.-Y., & Huang, W.-C. (2021). Effect of printing parameters on the thermal and mechanical properties of 3D-printed PLA and PETG, using fused deposition modeling. *Polymers*, *13*, 1758. https://doi.org/10.3390/polym13111758
[11] Valvez, S., Silva, A. P., & Reis, P. N. B. (2022). Optimization of printing parameters to maximize the mechanical properties of 3D-printed PETG-based parts. *Polymers*, *14*, 2564. https://doi.org/10.3390/polym14132564
[12] Katschnig, M., Wallner, J., Janics, T., Burgstaller, C., Zemann, W., & Holzer, C. (2020). Biofunctional glycol-modified polyethylene terephthalate and thermoplastic polyurethane implants by extrusion-based additive manufacturing for medical 3D maxillofacial defect reconstruction. *Polymers*, *12*, 1751. https://doi.org/10.3390/polym12081751
[13] Xu, J., Xu, F., & Gao, G. (2021). The effect of 3D printing process parameters on the mechanical properties of PLA parts. *Journal of Physics: Conference Series*, *2133*, 012026. https://doi.org/10.1088/1742-6596/2133/1/012026
[14] Tunçel, O. (2024). Optimization of Charpy impact strength of tough PLA samples produced by 3D printing using the Taguchi method. *Polymers*, *16*(4), 459. https://doi.org/10.3390/polym16040459
[15] Travieso-Rodriguez, J. A., Jerez-Mesa, R., Llumà, J., Traver-Ramos, O., Gomez-Gras, G., & Roa Rovira, J. J. (2020). Mechanical properties of 3D-printing polylactic acid parts subjected to bending stress and fatigue testing. *Polymers*, *12*, 3859. https://doi.org/10.3390/ma12233859
[16] Govsa, F., Celik, S., Turhan, T., Sahin, V., Celik, M., Sahin, K., Ozer, M. A., & Kazak, Z. (2018). The first step of patient-specific design calvarial implant: A quantitative analysis of fresh parietal bones. *European Journal of Plastic Surgery*, *41*, 511-520. https://doi.org/10.1007/s00238-018-1411-6
[17] Shi, H., Gan, Q., Liu, X., Ma, Y., Hu, J., Yuan, Y., & Liu, C. (2015). Poly (glycerol sebacate)-modified polylactic acid scaffolds with improved hydrophilicity, mechanical strength and bioactivity for bone tissue regeneration. *RSC Advances*, *5*, 79703-79714. https://doi.org/10.1039/C5RA13334C
[18] Sunarno, Z., & Zainuddin. (2020). Impact test analysis on steel metal materials and aluminum. *Journal of Social Research*, *2*(7), 2378-2392. https://doi.org/10.55324/josr.v2i7.1198
[19] Hasan, M. R., Davies, I. J., Pramanik, A., et al. (2024). Impact of process parameters on improving the performance of 3D printed recycled polylactic acid (rPLA) components. *International Journal of Advanced Manufacturing Technology*, *131*, 3751-3779. https://doi.org/10.1007/s00170-024-13140-7
[20] Athanasiou, K. A., Niederauer, G. G., & Agrawal, C. M. (1996). Sterilization, toxicity, biocompatibility and clinical applications of polylactic acid/polyglycolic acid copolymers. *Biomaterials*, *17*(2), 93-102. https://doi.org/10.1016/0142-9612(96)85754-1
[21] Miettinen, H., Mäkelä, E. A., Vainio, J., Rokkanen, P., & Törmälä, P. (1992). The effect of an intramedullary self-reinforced poly-L-lactide (SR-PLLA) implant on growing bone with special reference to fixation properties. An experimental study on growing rabbits. *Journal of Biomaterials Science, Polymer Edition*, *3*(6), 443-450. https://doi.org/10.1163/156856292X00420
[22] Jasiūnaitė, Ž. M., Šešok, A., Stonkus, R., & Šešok, N. (2021). Research of materials properties for 3D printing of orthoses. *Mokslas – Lietuvos Ateitis / Science and Future of Lithuania*, *13*. https://doi.org/10.3846/mla.2021.14168

Authors' contacts:

Reem Sabah Khzzal
(Corresponding author)
Department of Production Engineering and Metallurgy, University of Technology,
Al-Sina'a Street, 10066 Baghdad, Iraq
sabahreem18@gmail.com

Wisam Khdim Hamdan, Professor
Biomedical Engineering Department, University of Technology,
Al-Sina'a Street, 10066 Baghdad, Iraq
wisam.k.hamdan@uotechnology.edu.iq

Marwan Nafaa Ali, Assist Professor
Biomedical Engineering Department, University of Technology,
Al-Sina'a Street, 10066 Baghdad, Iraq
Marwan.n.arbilei@uotechnology.edu.iq

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250502094727

Subject review
Received: 2025-05-02, Accepted: 2025-06-02

Global Logistics in a Changing Environment and Best Practices Implementation

Stefan Schmidt

Abstract: The international economic interdependence and transnational organization of goods and services with global supply chains determine the economy, environment and working conditions in many countries around the world in numerous ways. The main risks in the supply chain are of an economic, ecological, political and ethical nature. The EU Supply Chain Act obliges companies to operate fairly and sustainably and to respect human rights and ecological protection in their supplier network. This contribution describes the effects of geopolitical hazards on the tactical level of the supplier network. Challenges include costs, disruption and regulatory and compliance issues. A case study of an automotive manufacturer shows supply security and resilient supply chains; risk management in purchasing, e.g. internal manufacturing, Innovations as a strategical competitive edge; digitalisation; social and environmental responsibility and carbon emissions.

Keywords: best practices; disruption; environment; global logistics; risk management; supplier

1 INTRODUCTION

For the first time since the end of the Cold War, Germany and Europe must again engage intensively with geopolitical risks and potential threats of war. Supply chains continue to be affected globally by unprecedented disruptions. Geopolitical instability and uncertain political climates increase the likelihood of sudden disruptions in supply. Supply chain resilience and sustainability help stakeholders to succeed throughout and after the interruption. A resilient supplier network provides to the sustainable competitive edge of the whole value stream [1, 2].

- Tactical risks: relate mainly to external issues with suppliers or customers: they include geopolitical supply chain risks, fires, quality problems, late deliveries, compliance issues, extreme weather phenomena caused by climate disasters and more.
- Operational Risks: involve mainly internal processes that the company has control over and that align with day-to-day front-line issues.

Geopolitical risks of the supply chain in focus in 2024 were, for example, the Russia-Ukraine war and Russia-Nato tensions and their global impact, the Israel-Hamas war, Red Sea crises, the US election, US-China tension, China-Taiwan tension, the advancement of climate change and environmental policies, the rise in global health challenges, protectionism and trade policy shifts, energy security and transition and demographic change. Allianz's Country Risk Atlas provides extensive overviews of the general policy, the macroeconomic and commercial situation and of the ecological aspects that influence default risk for companies in 84 countries [3]. How theoretical approaches from other fields can be used and practical research is presented in a variety of contexts in [4].

Logistics, as an accessory to and victim of the climate crisis, is among the biggest emitters of greenhouse gases, which demonstrates the need for environmental protection in logistics. Companies are not isolated in the supplay chain and relying on distant suppliers can be disastrous.

What significant role does managing resilience in operational and supply chains play in the transition to a more sustainable future [5]? To this end, the OECD policy toolkit provides pointers with regard to increasing supply chain resilience [6]. Limited sourcing options lengthen lead times and reduce flexibility, thus causing operational delays. As a result, companies are now adopting a ‘local for local’ concept with their suppliers and production networks. They collaborate with local suppliers to acquire parts near their manufacturing locations. The shortened transport distances reduce the carbon footprint of the parts. Furthermore, in times of extensive unstable supply chains, this localisation also allows for a greater level of flexibility and robustness in maintaining the supply of manufacturing facilities.

Embedding human rights due diligence in the procurement process is a key element for effectively enforcing defined social and environmental standards. Manufacturers' requirements are a significant factor for suppliers in implementing measures that help to ensure respect for human rights and are also considered by rating agencies. The "Act on Corporate Due Diligence in Supply Chains" came into effect on January 1, 2023. The legislation governs the accountability of German companies for compliance with human rights in worldwide supplier networks [7]. For the customer, the consideration of sustainability aspects and human rights is now playing an increasingly important role in purchasing decisions in addition to quality.

The aim of this article is to show how the network of a resilient supply chain is secured in the face of geostrategic aspects and that to establish a more resilient and flexible supply chain, it is necessary prioritize risk management in purchasing, strive to achieve in-house production as a means of gaining a strategic competitive advantage, seek collaboration with technology partners to implement innovations and develop a digitalization strategy. This must also include the execution of social and environmental responsibilities and acting with due diligence, for example, regarding carbon emissions in the supplier network.

2 CASE STUDY: SUPPLY CHAINS & SUPPLIER NETWORK

This part describes a case study of a premium automobile manufacturer, whose Procurement and Supply Organization is in charge worldwide for the sourcing and quality control of

manufacturing and raw materials, parts, assets and services as well as for the internal manufacturing of car parts.
The most important topics are, see [8]:

- Securing supply for the manufacturing site
- Extending robust supply channels
- Sourcing high-quality parts at favourable conditions
- Realisation of potential applications at an early phase
- Digitising the sourcing processes
- Integration of both social and environmental norms
- Integrating & improving the production of internal parts.

The case study is based on the author's many years of experience in the company in the fields of logistics, quality, risk and health management, environmental impact studies, with intensive access to internal documents and meetings, and in his activities as a consultant. More information was obtained and analysed through audits, expert interviews, questionnaires and data updated from the 2024 Company Report [8].

2.1 Supply Security and Resilient Supply Chains

Supply security: The year 2023 was shaped by enhanced supply robustness and by increasing commodity costs caused by raging global inflation. This led to higher expenses for the manufacturer, especially for commodities, and for energy and freight. The general supply conditions also continued to be difficult for specific technologies 2024 [8].

Natural hazards represent a further risk. As in 2023, there were bottlenecks in 2024 due to supply shortfalls resulting from the increase in severe weather conditions such as flooding and severe storms. The outlook for more severe climate impacts and increasing occurrences of severe weather conditions are also expected to complicate future planning. The company can reduce the impact of these unpredictable conditions thanks to its agile supplier ecosystem and thereby almost completely secure its capability to provide cars.

In order to reduce the effects of the war in Ukraine on its European sites, the manufacturer has taken extensive measures since 2022 to enable it to continue sourcing parts from western Ukraine [8]. Simultaneously, its partners have set up further manufacturing facilities outside Ukraine and supplies to the sites have largely returned to normal.

Expanding resilient supply chains: In 2023, the manufacturer was largely certain of the robustness of its supplier network due to the challenges imposed by geopolitical disasters and the unsecure availability of crucial commodities. The rise in costs on the world markets has also worsened to an even greater extent the tense procurement conditions. The manufacturer's RiskHub is a crucial part of improving the robustness of its worldwide supply network.

The RiskHub studies data from external, publicly available data sources concerning issues such as natural disasters and financial risks. It enables risks to be recognized at an early phase using artificial intelligence (AI) and data analytics [8]. The Catena-X Automotive Network (Catena-X) is a platform for the uniform transmission of information across its entire supply chain from the manufacturer to the SME suppliers and on to recycling businesses. Catena-X creates a collaborative, distributed information and service ecosystem. The Catena-X group consists of car manufacturers and suppliers, distributor societies and plant manufacturers [9].

The manufacturer applies a policy of sourcing car parts near to its manufacturing sites wherever possible. Local sourcing teams are therefore present in all major markets. Local risks and chances are therefore recognised fast and the company can respond quickly to the relevant market circumstances (Tab. 1). A robust supply chain strategy is becoming more important and the manufacturer is increasingly taking this into account when selecting its partners.

Table 1 Regional distribution, companies' purchased volumes [8]

Germany	33.1 %
Eastern Europe	19.6 %
North America	15.3 %
Other Western Europe	10.7 %
China	17.3 %
Other	4.0 %
Total:	96.1 billion € (2023)

2.2 Risk Management

Risk management in purchasing: A cross-linked risk management system is essential for the supply chain. It proactively steps in when appointing suppliers and responds quickly in recognising and analysing potential supply shortages. The manufacturer's RiskHub uses external, freely accessible data to analyse information on potential risks such as natural disasters or financial risks. State-of-the-art methods such as big data analytics and AI are also utilised [8]. Risk Data Hub is a platform that enhances the access and exchange of EU-wide maintained disaster risk data for Disaster Risk Management (DRM) related activities and forms the foundation of science-based information for DRM policies [10]. For chosen parts, the sites of sub-suppliers are also considered when evaluating the total risk [11].

To prevent cyber threats and efficiently manage the ever-growing number of supplier-related security threats, the manufacturer remains committed to certifying suppliers according to the vehicle standard Trusted Information Security Assessment Exchange (TISAX®). TISAX is an across-company assessment and exchange procedure for information relating to security in the automotive industry. It is concerned with the protection of data and its integrity as well as ensuring availability in the manufacturing process and in the operation of vehicles [12]. TISAX is an embedded feature of the sourcing process and an integral requirement of agreements with suppliers that fulfil the specified relevant criteria. The joint IT security analyses and the frequent exchange of information via a B2B portal makes the supplier network more resilient [8].

More and more trainings and information sessions are held both in-house and outside the company to increase awareness of the fast-increasing number of cyber threats.

2.3 Internal Manufacturing and Innovations

The strategic benefits of internal manufacturing: The manufacturer periodically reviews its product program regarding profitability and strategic significance. The priorities of this assessment relates the areas of electromobility, digitalization and sustainability. In these central transformation

fields, the manufacturer is extending its internal know-how in relation to strategy-critical parts in a focused and lasting manner. Simultaneously, the capability to innovate must be increased and supply secured through internal manufacturing. The advantages of this solution are evident, for example, in the controller unit of the high-voltage battery and in the electrical powertrain system. The manufacturer also wants to reinforce its process expertise in the technological areas crucial to quality assurance and in the qualification of the supplier ecosystem. A commitment to quality for the customer provides a decisive competitive advantage and is a key factor in the manufacturer's long-term success.

Innovations: By collaborating with technological associates and establishing innovative networks, the manufacturer can apply innovations to its cars in the shortest possible time. The company works closely with the Startup Garage, which has assessed a total of 4,700 startups to date, of which 30 are now established as suppliers and service providers of the manufacturer and form an integral part of its network [8]. A collaboration of the company with iVentures in Silicon Valley reveals that investments have been made in startups in the fields of automation, manufacturing, robotics, logistics, and quality assurance [8].

The suppliers' results are regularly evaluated within the technology networks [8]. Suppliers contribute directly to the success of the manufacturer and, together with their employees, advance all their tasks via the partner portal, a worldwide platform for collaboration characterized by a mutual awareness of quality, security of supply, innovative capacity and the commitment to sustainability [13].

2.4 Digitisation of the Supply Network

A digitised transfer of data across the entire supply network is indispensable for the company. This is an absolute precondition for creating a more robust and agile supply network infrastructure.

Catena-X [9] is the car industry's frontrunner project for the digitisation of the supply network. This takes place between car producers, sub-suppliers, suppliers and, upcoming, recycling enterprises. The digital basis allows all parties in the supply network to overcome the obstacles within the sectors. This applies, for example, to improving robustness and fulfilling sustainability targets and legal regulations by means of digital cooperation. A beta test phase was effectively implemented in 2023 with the substantial involvement of the manufacturer as consortium leader. In cooperation with its component production, pilot projects have been implemented. This has generated useful knowledge that will help to encourage innovative IT processes in a manufacturing setting (Catena-X rollout phase since 10/2023) [9, 36].

The manufacturer concentrates primarily on the tracking of parts in the supply network, conducting collaborative root cause analyses to eliminate possible quality issues and gathering data of CO_2 emissions in the supply network. 2024 "use cases" with suppliers have been implemented and spread throughout the entire industrial sector. An important progress is the delivery of the Digital Product Passport (DDP) [14], which includes product-specific information for parts such as cells or complete cars. Data are available through Catena-X on the basis of information from n-tier suppliers. This contains information on the source and composition of the materials and on recycling.

In future, products must be aligned with the principles of the circular economy. Therefore, a knowledge of the preceding steps of the supply chain and the effect of the manufacturer's activities upon down-stream steps is essential. Here, DPP is a key enabler as it delivers the required data and networking opportunities. DPP's goal is to provide more traceability throughout the supply chain by digitally saving product-specific data (Ecodesign for Sustainable Products Regulation in force since 18.08.2024 [15]).

2.5 Ecological and Social Accountability

The manufacturer sees sustainable supplier management as an essential element of solid business management. The demands and anticipations it imposes on its vendors are defined in the manufacturer's "Supplier Code of Conduct" [11] and contractually binding as an essential element of the manufacturer's purchasing requirements [16].

In compliance with human rights, the social and environmental regulations of the complete supplier network form the basis for the permanent success of the manufacturer. This is defined in the manufacturer's "Supplier Code of Conduct". It obliges its Tier 1 suppliers to impose their demands on their upstream sub-suppliers [11].

A corporation-wide "Compliance Management System (CMS)" strengthens the behaviour of strict adherence and probity and minimises threats. Globally accredited standards for the safeguarding of human rights form the foundation for the company's entire supplier network. It considers these as to be special requirements for itself, its supplier's network and its other business partners [17], e.g. the Global Reporting Initiative (GRI) Content Index [18]. GRI is an independent, non-profit organisation that provides a comprehensive framework for sustainability reporting.

2.6 Due Diligence within the Supplier Network

The manufacturer's entrepreneurial accountability for social and environmental rights in the supplier network is institutionalized within the framework of a multiple-step due diligence procedure [19]. It primarily applies standards-based monitoring processes to fulfil the due diligence responsibilities anchored in its corporate procedures. The manufacturer works with campaigns and federations such as the "Responsible Business Alliance (RBA)" [20] and "Drive Sustainability" [21]. RBA represents the worldwide biggest industrial association. It is committed to sustainable corporate governance in worldwide supplier networks. Drive Sustainability is an automotive Partnership between 16 car manufacturers, facilitated by CSR Europe, which provides a Self-Assessment Questionnaire on Corporate Sustainability and Responsibility (CSR) for automotive sector suppliers. The manufacturer follows the principles created by the "German Automotive Industry Dialogue" for the five key aspects of human rights due diligence (risk analysis, policy statement, measures, grievance and reporting procedure) [22]. The manufacturer advocates the implementation of a cross-indus-

try complaints system [17]. For guidance on organising, implementing and evaluating complaint procedures, see the German Supply Chain Due Diligence Act [23].

Risk evaluation: The company reviews and evaluates the relevance of the risks of its suppliers at both prospective and already existing supplier locations. Different in-house and outside information channels, such as national and product-specific key figures and media reviews, are utilized to detect and evaluate human rights and ecological risks. The manufacturer employs standard ecological questions (see "Sustainability Code Database" [23]) and engages third-party on-spot audits to perform extensive risk evaluations of its Tier 1 suppliers. With the Sustainability Codex, the "German Council for Sustainable Development (RNE)" intends to encourage enduring corporate policies [24]. Illustrating the supply chains serves as the foundation for identifying the risks in supplier chains and the manufacturer is constantly striving to improve visibility in the whole supplier network [18].

Prevention and remediation: If risks are identified, suppliers are obliged to carry out the required prevention and monitoring tasks by a specified date before finalising the contract with the manufacturer. Suppliers are also required to demand these measures from their sub-suppliers, depending on the specific risks they pose. As a part of the procurement procedure, these activities are validated and assessed via the sector-wide Drive Sustainability Survey [21]. This is intended to overcome existing shortcomings and reduce possible risks. For GRI disclosures on the environmental assessment of applicable supplier sites, see also [18] and see SAQ Toolbox [25]. The on-line evaluation is carried out for providers of manufacturing and non-manufacturing products and services with a minimum order volume, based on market and product-specific hazards. Different modules are used and the deployment of each must be successfully finished before the launch of manufacturing. The extent of the preventive activities depends on the type and scale of the commercial operations, the possible threats and the scope of the supplier. Trainings on the topic of sustainability in the supplier network are offered by the company for internal insourcers. This is also applied to suppliers and process partners on an optional basis.

The manufacturer has set up further monitoring procedures for suppliers in critical locations or high-risk product categories. Tools consist for the audits of ecological and social requirements at supplier sites with the help of cross-sector assessment programs supplied by the "Responsible Business Alliance (RBA)" [20] and the "Responsible Supply Chain Initiative (RSCI)" [26, 27] of the VDA (Verband der Automobil Industrie). The RSCI association was launched in October 2021 by the VDA together with 14 other founding members, including both automotive manufacturers and suppliers.

The manufacturer's target is to fulfil local legal demands for both sustainability and international human rights standards for all assessed Tier 1 supplier locations.

2.7 Carbon Emissions in the Supply Chain

Logistics and transport accounts for approximately 24% of worldwide CO_2 emissions. Climate-neutral logistics is the future. The most important pollutants in terms of Greenhouse gases found in the global atmosphere are CO_2, nitrogen dioxide (N_2O), water vapour, methane (CH_4), and ozone. By 2050, the European Environment Agency expects that, without effective measures, global logistics will cause 40% of global carbon emissions [28]. Transport is responsible for around a fourth of all EU greenhouse gas emissions and is the cause of air and noise pollutants and for the fragmentation of living areas [29].

The manufacturer is committed to supporting the decarbonization of its supplier network and reducing carbon emissions from the supplier network, e.g. by 40% per vehicle by 2030 compared to 2019 [30, 31, 32]. It is therefore one of its procurement criteria that suppliers pledge to take decarbonisation measures. Measures include reducing carbon emissions by, for example, using green electricity, using secondary commodities, operating newer production processes for commodities and using innovative products and materials such as organic materials. Therefore, offsetting activities not covered. They specifically support their suppliers by offering online training, face-to-face partner workshops [18] via manufacturer's Partner Academy [33].

The figure for contracts with suppliers that defined decarbonization activities increased significantly in 2023. This increase includes activities that require the usage of recycled materials, bio-materials and de-carbonized steel. According to the manufacturer, the usage of green electricity has the biggest effect on decarbonisation of the supplier network, (internal production, directly procurement or "Energy Attribute Certificates" [34], for example proof of origin). The manufacturer has determined that the utilization of green electricity is a prerequisite for the allocation of orders for the procurement of carbon-intensive parts and commodities. This applies to suppliers (tier 1) and for energy-demanding processes of sub-suppliers (n-tier). The manufacturer has signed many call-off contracts with its largest suppliers for the utilization of green electricity since mid-2021. An initiative was created in 2023 to evaluate suppliers' schemes for the utilization of green electricity before the start of manufacturing.

In addition, the manufacturer continues to appoint an independent third-party contractor to perform an annual assessment of decarbonization efforts in mass manufacturing. The carbon dioxide emitted in the manufacturer's logistics and supply network is measured globally in in tons of CO_2 equivalent (tCO_2e) per vehicle produced. The result in terms of reducing the negative impact on the environment was a 78% reduction in carbon dioxide emissions per car manufactured from 2006 to 2020. The target is the reduction of average carbon dioxide emissions per car by a projected 80 % by 2030 in comparison to the base year 2019 [35]. The number of suppliers with contracts to reduce CO_2 emissions increased from 468 (2022) to 707 companies in 2023. The result was a saving of 1.7 million tonnes of carbon dioxide emissions [36]. The company further announced that it will avoid emissions of over 200 million tons of CO_2 by 2030 [30].

Scope 1 to Scope 3 emissions are founded on emissions from different sources [37]:

- Scope 1: immediate emissions resulting from the incineration of fossil fuels within a firm.
- Scope 2: indirect emissions generated using electricity and heating from outside energy supplies

- Scope 3: indirect emissions resulting from the downstream and upstream stages of the processes: downstream - utilisation of goods and services; upstream - in the supplier network

The manufacturer is committed to the research-based objective of lowering its Scope 1 and Scope 2 emissions by 80 percent and its upstream Scope 3 emissions by 22 percent per car by 2030 [37]. As expected, emissions from the fleet's supply chain will increase. It expects that the implemented CO_2-reduction activities will overturn the present rising tendency by 2026 at the very latest. The manufacturer still uses the "Carbon Disclosure Project (CDP) Supply Chain Programme" [38] for evaluation of the efficiency of its supplier network regarding its decarbonisation.

3 CONCLUSION

Can we provide an outlook and conclusion for the near future? For the Bavaria 2025 Raw Materials Report, the Bavarian Chambers of Industry and Commerce (BIHK) surveyed ca. 7,000 member companies in 2024 [39]. The results:

- Drastic price increases & supply bottlenecks predicted
- Continuing uncertainties on global markets expected
- 60% expect that the supply situation will worsen in the next 5 years
- It is expected that a tighter supply of local raw materials will lead to longer transport routes, higher transport costs & CO2 emissions
- Due to the political environment, 40% welcome the establishment and expansion of strategic partnerships in the raw materials sector
- Measures necessary to combat impending supply bottlenecks: optimise warehousing and manage inventories
- Diversification of raw material sources and long-term supply contracts is required.

Company adjustments to be made: optimise transport (86%), warehousing (67%), packaging (49%).

Expectations for the next five years of the companies surveyed:

- 60% predict a worsening of the supply situation
- 49% expect a deterioration due to the war in Ukraine
- 44% expect bottlenecks in international maritime trade
- 26% expect an impact from the inception of the "European Supply Chain Due Diligence Act".
- 25% expect an impact from China's actions regarding critical raw materials
- 39% expect an impact from the entry into force of the EU Critical Raw Materials Act.

Global supply chains will continue to be challenging and subject to increasing pressure from higher prices and costs, raw material shortages, natural disasters, geopolitical crises and extreme weather events. Logistics and transport accounts for approximately 24% of worldwide CO_2 emissions. Climate-neutral logistics and production are the future.

Respect for human rights, compliance with social standards and the reduction of CO_2 emissions are becoming more of a focus and are being demanded by the public, politicians, rating agencies, laws and regulations and especially by customers in their purchasing decisions.

Reducing CO_2e emissions from logistics challenges is a priority that goes beyond traditional risk management. Companies that proactively address these risks secure competitive advantages in a volatile global economy. A well-structured, resilient supply chain transforms potential challenges into opportunities for sustainable growth and risk reduction and it respects human rights, adheres to social standards and supports the supply chain in its decarbonisation.

4 REFERENCES

[1] Rausch-Phan, M. T., & Siegfried, P. (2022, January 29). *Sustainable supply chain management*. In *Sustainable Supply Chain Management* (pp. 17–46). Springer. https://doi.org/10.1007/978-3-030-92156-9_3

[2] Kummer, S. (Ed.), Wakolbinger, T. (Ed.), Novoszel, L. (Ed.), & Geske, A. M. (Ed.). (2022, September 25). *Supply chain resilience: Insights from theory and practice* (Springer Series in Supply Chain Management, Vol. 17). Springer. https://doi.org/10.1007/978-3-030-95401-7

[3] Allianz SE. (2024, December 31). *Allianz country risk atlas 2024*. https://www.allianz.com/de/mediencen–ter/news/studien/240131-allianz-country-risk-atlas-2024.html

[4] Elias, A. A., Pepper, M., Gurumurthy, A., & Shukla, A. K. (2024, July 25). *Sustainable and resilient supply chain - Environmental accounting and management focus*. Emerald Publishing Ltd.

[5] Holgado, M., Blome, C., Schleper, M. C., & Subramanian, N. (2024, February 25). Brilliance in resilience: Operations and supply chain management's role in achieving a sustainable future. *International Journal of Operations & Production Management, 44*(5), 877–899. https://doi.org/10.1108/IJOPM-12-2023-0953

[6] OECD. (2024). *Keys to resilient supply chains - Increasing supply chain resilience: OECD policy toolkit*. https://issuu.com/oecd.publishing/docs/resilient-supply-chains_brochure-2024

[7] Federal Ministry of Labour and Social Affairs. (2021, July 16). *Act on corporate due diligence obligations in supply chains of July 16, 2021*. https://www.csr-in-deutschland.de/SharedDocs/Downloads/EN/act-corporate-due-diligence-obligations-supply-chains.pdf?__blob=publicationFile

[8] BMW Group. (2025, March 11). *BMW Group report 2024*. https://www.bmwgroup.com/en/report/2024/downloads/BMW-Group-Report-2024-en.pdf

[9] Taba, R., & Weiß, M. (n.d.). *Digitalisation of the automotive value chain - Catena-X*. German Aerospace Center (DLR). https://www.dlr.de/en/ki/research-transfer/projects/catena-x

[10] European Union UCP Knowledge Network. (2023, September 6). *Risk Data Hub analytical reports*. https://civil-protection-knowledge-network.europa.eu/news/risk-data-hub-analytical-reports

[11] BMW Group. (2022, December 6). *BMW Group supplier code of conduct (BMW Group Supplier Sustainability Policy), Version 3.0*. https://www.bmwgroup.com/content/dam/loads/en/2022/BMW-Group-Supplier-Code-of-Conduct-V.3.0_englisch_20221206.pdf

[12] ENX Association. (2024, April 8). *TISAX participant handbook 2.7.2*. https://www.enx.com/handbook/TISAX%20Participant%20Handbook.pdf

[13] BMW Group. (2025). *BMW Group partner portal*. https://b2b.bmw.com/en/web/b2b/home

[14] Neligan, A., Schleicher, C., Engels, B., & Kroke, T. (2023, September 22). *Digital product passport as enabler of the circular economy – Relevance and practicability for companies*. Institut der deutschen Wirtschaft Köln e. V. https://circulareconomy.europa.eu/platform/sites/default/files/2024-02/IW-Report_2023-Digitaler-Produktpass-englisch.pdf
[15] European Union. (2025). *Ecodesign for sustainable products regulation – Making sustainable products the norm in the EU*. https://commission.europa.eu/energy-climate-change-environment/standards-tools-and-labels/products-labelling-rules-and-requirements/ecodesign-sustainable-products-regulation_en
[16] BMW Group. (2022, December 1). *BMW Group international terms and conditions for the purchase of production materials and automotive components (IPC), version December 01, 2022*. https://b2b.bmw.com/documents/14402/7501963/20221201_IPC+2022_EN+%282%29.pdf/00f588f4-6296-0b7d-8b92-c0be291fc4c3
[17] BMW Group. (2025). *Compliance and human rights at the BMW Group*. https://www.bmwgroup.com/en/company/compliance.html
[18] BMW Group. (2025). *Global Reporting Initiative (GRI) content index 2023*. https://www.bmwgroup.com/en/report/2023/downloads/BMW-Group-GRI-Index-2023-en.pdf
[19] BMW Group. (2025). *Due diligence in the supply chain*. https://www.bmwgroup.com/en/sustainability/supply-chain.html
[20] Responsible Business Alliance (RBA). (2025). *The Responsible Business Alliance (RBA) advancing sustainability globally*. https://www.responsiblebusiness.org
[21] CSR Europe. (2019, August 31). *Self-assessment questionnaire on CSR/Sustainability for automotive sector suppliers* (SAQ version 4.0). https://www.drivesustainability.org/wp-content/uploads/2020/06/DS_SAQ-FORM2019_17012020-EN.pdf
[22] Federal Ministry of Labour and Social Affairs (BMAS). (2022, June 29). *Guidelines*. https://www.csr-in-deutschland.de/EN/Business-Human-Rights/Implementation-support/Sector-dialogues/Automotive-Industry/Guidelines/guidelines-art.html
[23] Federal Office for Economic Affairs and Export Control (BAFA). (2022, October). *Organising, implementing and evaluating complaints procedures* (1st Ed.). https://www.bafa.de/SharedDocs/Downloads/EN/Supply_Chain_Act/guidance_complaints_procedure.pdf?__blob=publicationFile&v=6
[24] Deutscher Nachhaltigkeits Kodex. (2025). *BETA-Version (German version) of the Sustainability Code Database is online*. https://www.deutscher-nachhaltigkeitskodex.de/en/
[25] Drive Sustainability / CSR Europe. (2023). *SAQ Sustainability Assessment Questionnaire - Toolbox for Suppliers* (SAQ version 4.0). https://www.drivesustainability.org/saq-toolbox/
[26] RSCI Association. (2024). *The Responsible Supply Chain Initiative (RSCI)*. https://www.rsci.online/en
[27] RSCI Association. (2024, August 20). *Detailed assessment information 2.0*. https://www.rsci.online/dam/jcr:66010966-91a6-4abd-92ee-b53e046f7537/RSCI-Assessment-Information.pdf?mode=view
[28] Carbon-Care Switzerland. (2025). *Responsibility of transport and logistics*. https://www.carboncare.org/en/climate-change
[29] Bakker, S. (2025, February 10). *Transport and mobility*. European Environment Agency. https://www.eea.europa.eu/en/topics/in-depth/transport-and-mobility
[30] Krivevski, B. (2021, May 17). Over 200 million tonnes: BMW Group sets ambitious goal to reduce CO_2 emissions by 2030. *ElectricCarsReport.com*. https://electriccarsreport.com/ 2021/05/over-200-million-tonnes-bmw-group-sets-ambitious-goal-to-reduce-co2-emissions-by-2030/
[31] Schneider, D., & Bene, R. (2024, October 8). Die Neue Klasse von BMW wird die E-Mobilität beschleunigen. *Go Green Media GmbH CH*. https://gogreen.ch/de/die-neue-klasse-von-bmw-wird-die-e-mobilitaet-beschleunigen/
[32] Neste Germany GmbH. (2025). *Wie die BMW Group in Pilotprojekten mit Neste MY Renewable Diesel den CO2-Fußabdruck in der Transportlogistik senkt*. https://www.neste.com/de-de/nachrichten-und-einblicke/transport/bmw-group-pilotprojekten-mit-neste-my
[33] BMW Group. (2025). *Partner Academy - Better together*. https://bmw-partner.academy
[34] United States Environmental Protection Agency (EPA). (2025, January 14). *Energy Attribute Certificates (EACs)*. https://www.epa.gov/green-power-markets/energy-attribute-certificates-eacs
[35] BMW AG. (2023, August 1). *Welcome to your CDP Climate Change Questionnaire 2023* (pp. 1–359). https://www.bmwgroup.com/content/dam/grpw/websites/bmwgroup_com/ir/downloads/en/2024/beicht/BMW_Group_CDP_Climate_Change_Questionnaire_2023.pdf
[36] BMW Group. (2024, June 25). *Driving decarbonisation in the supply chain*. https://www.bmwgroup.com/en/news/general/2024/decarbonisation.html
[37] World Benchmarking Alliance. (2024). *Automotive and transportation manufacturers benchmark – BMW*. https://www.worldbenchmarkingalliance.org/publication/automotive-transport/companies/bmw-6/
[38] CDP Worldwide. (2025). *Supplier engagement assessment*. https://www.cdp.net/en/supply-chain/supplier-engagement-assessment
[39] CoC Bavaria BIHK. (2025, January). *Rohstoffreport Bayern 2025*. https://www.bihk.de/fileadmin/eigene_dateien/bihk/PDF/Rohstoffreport_2025_UA_final.pdf

Author's contacts:

Stefan Schmidt, Dipl.-Ing., Dipl.-Wirtschafting.
Technology Management and Management Training/VWI/VDI,
Fritzstrasse 41, 82140 Olching, Germany
Stefan.Schmidt-1@t-online.de

CC BY

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250504123321

Subject review
Received: 2025-05-04, Accepted: 2025-05-16

Advancing Aircraft Maintenance through Predictive Technologies and Artificial Intelligence

Belma Atapek-Yagan*, Jiri Tupa, František Steiner

Abstract: This paper explores the transition in aircraft maintenance from conventional preventative methods to predictive maintenance driven by artificial intelligence (AI). Predictive maintenance utilizes data-driven methodologies to predict failures, reduce downtime and enhance safety. Artificial intelligence plays a pivotal role by processing substantial data to identify patterns and predict maintenance requirements with a high degree of accuracy. The present study addresses the fundamental challenges associated with integrating AI into aircraft maintenance, including robust data collection, algorithm transparency and cybersecurity risks. By addressing these issues, the paper provides actionable insights and solutions to effectively utilize AI while minimizing the associated risks. This study offers a novel perspective on the use of AI to revolutionize aircraft maintenance and improve efficiency and reliability, building on existing research and identifying opportunities to advance AI-driven maintenance strategies in aviation.

Keywords: artificial intelligence; Industry 4.0; passenger aircraft; predictive maintenance; quality process

1 INTRODUCTION

The increasing complexity of modern aircrafts, particularly in the area of electronic systems, has sparked a research interest in exploring innovative solutions for aircraft maintenance [1]. This area encompasses not only technical requirements but also critical factors to ensure passenger safety, operational efficiency, and cost effectiveness in the aviation industry [2]. The impetus for this research stems from the realization that traditional preventive maintenance approaches, while effective to a certain extent, are no longer sufficient to meet the demands of today's highly interconnected and technology-driven systems [3].

The emergence of predictive maintenance supported by artificial intelligence (AI) has proven to be a transformative opportunity in this area [4]. The potential of this approach lies in its ability to anticipate failures before they occur, leading to a significant reduction in unplanned downtime, optimized resource allocation and improved safety standards [5-7]. Despite the promising possibilities of this approach, it is important to recognize the challenges associated with it, such as the need for reliable data, transparent algorithms, and robust cybersecurity measures [8-10]. These challenges underscore the necessity for concerted research and innovation in this field. With this in mind, a systematic literature review providing a basic understanding of the challenges in aircraft maintenance associated with the shift from preventative to AI-powered predictive maintenance is provided in this paper.

Prior to presenting the main literature review findings, it is first necessary to give sufficient background information about the key concepts that undergird this study. Some of these concepts also form the basis for the keywords for the literature search. This is therefore done throughout the course of the subsequent chapter.

2 OVERVIEW OF KEY THEORETICAL CONCEPTS

2.1 Predictive Maintenance

The term "predictive maintenance" refers to condition-based maintenance that is conducted based on forecasts from analytics and known qualities of the characteristics of the state of an asset's degradation. It is considered to be an advancement from traditional maintenance practices, which typically only take place based upon information concerning the current condition of assets. It frequently draws upon data provided via artificial intelligence and the Internet of Things, which facilitate real-time access to insights into the asset's availability, condition, and location. This enables the augmentation of human decision-making via the prediction of asset health and allows responses to be made pre-emptively as opposed to reactively [11].

2.2 Artificial Intelligence

Although there is no ubiquitously defined definition of AI, it can be broadly conceptualized as digital technology that seeks to make machines operate in an intelligent manner dependent on workings that resembles those of the human mind [12, 13]. Within this context, the term "intelligent" refers to possessing the ability to find a solution to complex problems. It typically evolves data analysis, with the deployment of computer-generated rules based upon this data for the purpose of problem-solving being known as "machine learning" [12, 13].

AI has undergone numerous different advancements since the 1950s and recently experienced a period of dramatic progression. This has been heralded by various different developments including more sophisticated algorithms coming into being, improvements in the leveraging of big data, and an increase in computation speed [12]. Artificial intelligence is regarded as being especially important in the field of maintenance, where it can be regarded as a particularly disruptive innovation [14].

2.3 Industry 4.0

The term "Industry 4.0" refers to an industrial paradigm characterized by the integration of communication technologies, intelligence and digitalization into manufacturing practices [4]. This concept has been included in this section on account of the fact that both AI and

preventative management are considered to be closely tied to it [4, 15, 16] Although there are various different perspectives on what Industry 4.0 entails, it is generally held to be characterized by the following 9 pillars:

- Big data – high volumes of data are collected, prepared and analysed in order to facilitate advanced decision-making.
- Autonomous robots – intelligent machines that are able to carry out tasks with minimum human involvement.
- Simulation – digital imitations of systems and their characteristics are used for testing purposes.
- Additive manufacturing – the production of physical objects based upon 3D models by joining together successive material-layers.
- Vertical and horizontal integration – the former entails an interaction at various different levels within the hierarchical management-structure of an organization. The latter assumes that all internal and external entities and department are involved in creating value chains.
- The Internet of Things – physical objects are equipped with sensors that store, interchange and analyze data in order to enable adaptive decision-making and improve efficiency.
- Cloud computing – the storage of data and usage of resources via remote online platforms.
- Cybersecurity – practices, processes and technologies aimed at defending interconnected systems from cyber-attacks and data leaks.
- Augmented reality – the use of enhanced replicas of the physical domain using sound, computer graphics and other sensory data [17].

The impact of Industry 4.0 on maintenance has been identified as one of the main ways in which it can drive value for an organization [18]. The use of Industry 4.0 technologies for industrial-maintenance purposes is known as Maintenance 4.0 [19]. Now that the key concepts that underpin this study have been defined and described, the methodology can be presented.

3 METHODOLOGY

Before carrying out a systematic literature review, it is necessary to create a search protocol. The Systematic Review Model, as detailed by Vasquez and Castrillon, formed the basis for this protocol [20]. This model holds that the protocol consists of three different elements: the search question, keywords, and exclusion and inclusion criteria. The search question is the question that the review is geared towards answering [20]. It is "What are the main challenges associated with these advancements?"

The keywords that were used in this review are "artificial intelligence", "predictive maintenance", "quality process", "passenger aircraft" and "Industry 4.0". Predictive maintenance and artificial intelligence are both considered to be products of Industry 4.0, hence the selection of this concept as a search term [15, 16].

The inclusion criteria for the papers included in this study are that they need to have been published in the last 5 years and appear in a peer-reviewed academic journal, book published by an academic publisher, or proceedings of a reputable academic conference. They also need to be directly relevant to the topic of study. The exclusion criteria are publication over 5 years ago, appearance in texts other than those listed in the inclusion criteria, and obvious sources of bias or error. It was determined that the Scopus database would be used for the literature search, as it is considered to be amongst the most substantial curated databases, containing over 1.7 billion sources [21]. This means that it is likely to provide a suitably broad view of the current base of theory and research about the subject of study.

Furthermore, it is notable that Zaaba et al. have previously described Scopus as a "valuable resource for obtaining a global overview of scientific research output" [21, p. 533]. They express the view that it can facilitate significant insights into specific topics and trends [22]. This highlights the utility of searching within this database and its suitability for the present review. The Web of Science database was also used to further expand the volume of results.

The data included in the review was extracted by first screening the abstracts of articles for relevance and to see if they pass the inclusion criteria then thoroughly reading to full text articles to further check eligibility. The relevant portions of each text were then compiled in a document along with the accompanying bibliographical data ready for review. The data was analysed by identifying key themes relevant to the subject of study.

In addition to synthesizing, analysing and discussing the literature that is revealed by this search within the context of the research topic, the visualization tool VOSviewer was also deployed to provide a visualization of the connections and key concept clusters in the research area. It was selected for this purpose on account of it being easy and efficient to use and the fact that Ma has emphasized It is also useful for providing a rich, insightful display of bibliometric data and linkages [22]. The findings of the literature review are now outlined.

4 LITERATURE REVIEW

Fig. 1 shows the visualization of the connections and key concept clusters. As can be seen, digital twins, artificial neural networks, prediction accuracy, prediction models, maintenance costs, and maintenance accuracy emerged as prominent concepts within the literature in addition to topics such as predictive maintenance and artificial intelligence that featured as keywords. Various other related paradigms were also touched upon in the literature including AI techniques, data quality, defect detection and process monitoring. A total of 19 articles are included in the review.

It is now necessary to examine the main insights from the review. Kabashkin, Federov and Perekrestov note that it needs to be possible to integrate the requisite programs for AI-drive predictive-maintenance with the organization’s existing maintenance systems [23]. Furthermore, Kabashin et al. and Kabashkin and Susanin put forward the perspective that organizational resistance can act as an impediment to the

effective utilization of AI-based predictive maintenance in aircrafts [2, 23]. The latter point out that concerns about jobs being replaced could potentially act as a catalyst for discontent about the deployment of AI for this purpose [2]. Kabashin et al. and Kabashkin and Susanin posit that change management can go some way towards rectifying this issue [2, 23].

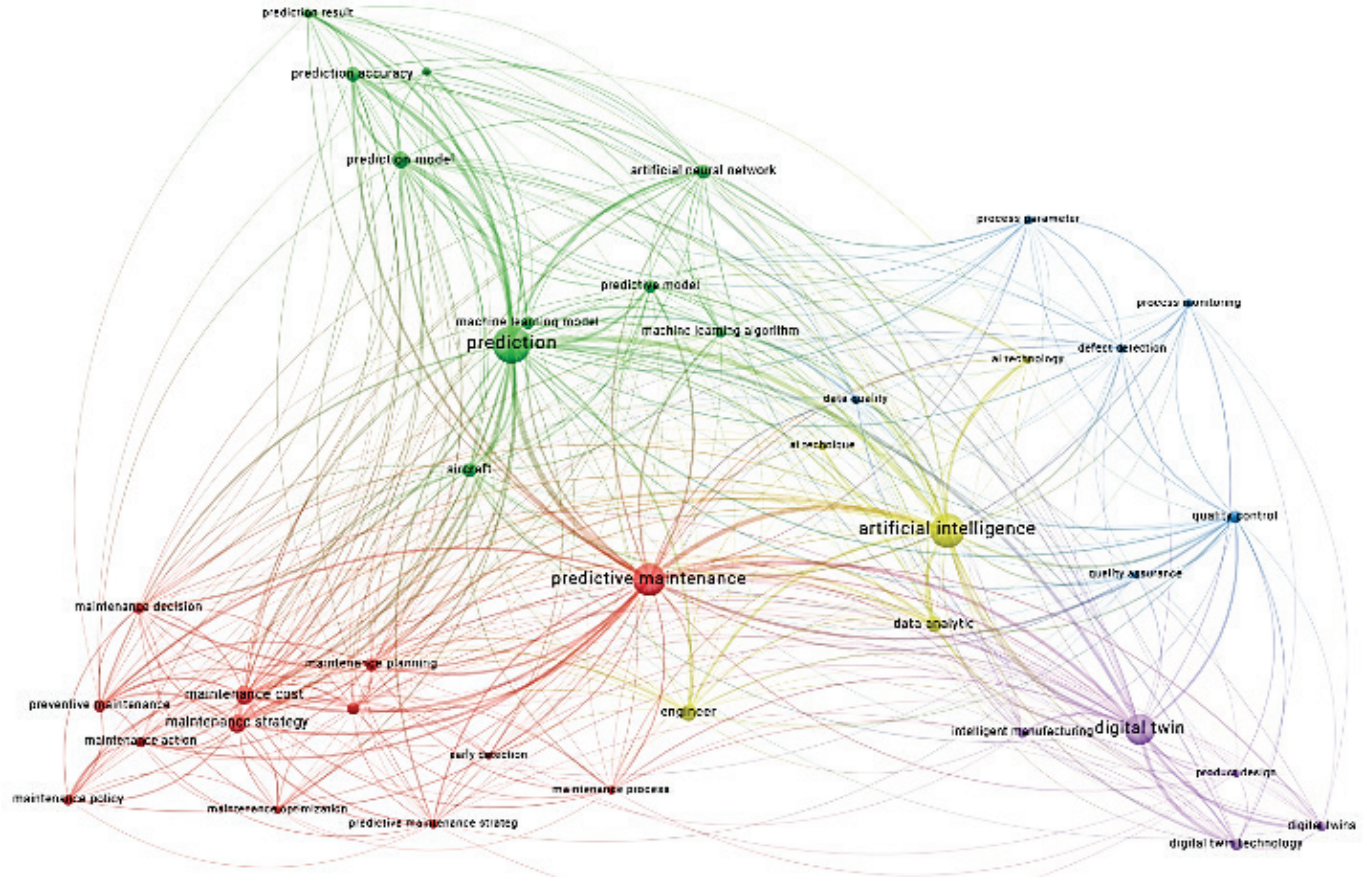


Figure 1 Visualization of Connections and Key Concept Clusters

Furthermore, Kabashkin and Susanin point out that there are challenges associated with the new skill sets required for employees to deal with AI-driven predictive-maintenance systems. They indicate that additional training is required in order to accommodate the new requirements. The workforce essentially needs to be upskilled in order to enable it to interact effectively with the new approaches [2].

Kabashkin et al. claim that compliance and regulatory requirements create additional complications, as they mean that predictive-maintenance strategies need to gain approval before being implemented [23]. Kabashkin points out that it can be time consuming for aviation authorities to approve comprehensive changes to maintenance of the types required to implement predictive-maintenance systems involving AI. He pinpoints this as a potential barrier to adoption. The literature also indicates that a considerable number of resources frequently typically need to be put towards attaining approval and thorough documentation and validation are needed to demonstrate compliance [24]. This could potentially reduce the usefulness of machine-learning-oriented predictive-maintenance systems, given that it means that they are relatively resource-intensive.

Furthermore, the aviation industry's global nature adds extra complexity to the situation, as different nations have different regulations that need to be adhered to. They also have different data-protection requirements and available technologies. These factors make it more complicated to meet the compliance-based and technical requirements of AI-based predictive-maintenance systems [2].

Kabashkin points out that transparency is needed in order to meet regulatory requirements [25]. AI decisions need to be adequately explained to ensure that regulations are met. This is especially the case with regards to safety-critical scenarios [2, 25]. Transparency is also necessary for building trust amongst stakeholders [2].

According to Kabashkin and Susanin, the key to additional transparency lies in conducting further research into interpretable models for machine learning and explanation frameworks intended for blackbox models. Furthermore, Pelt and Stamoulis highlight the additional legal and compliance-based issues that are brought up by the utilisation of AI in predictive maintenance [26]. Kabashkin and Susanin point out that it raises questions regarding certification and liability. This could potentially make it more complicated to navigate the regulatory environment. It can

also make it more expensive and resource-intensive to do so [2].

Kabashkin and Susanin also highlight the ethical challenges posed by the use of AI-based predictive-maintenance systems in aircrafts [2]. They state that "Questions arise about the level of autonomy these systems should have and where human oversight is necessary" [2]. Difficult human decisions clearly need to be made in order to accommodate the changes to new approaches to maintenance.

According to Kabashkin, there are also questions regarding who should be accountable for safety issues stemming from inaccurate predictions made by AI-based predictive-maintenance systems. This is likely to be more of an issue throughout the years to come as AI develops the possibility of acting more and more autonomously [25].

Further to this, Korba et al. have identified privacy-related challenges associated with the use of AI-based predictive-maintenance systems [27]. Korba et al. point out that sensitive operational data should be kept safe and Kabashkin and Susanin note that difficulties may arise regarding the responsible collection of data and the avoidance of impinging upon privacy rights of crew members and passengers [2, 37]. This could potentially necessitate additional effort and the deployment of extra resources.

This is especially the case when it comes to sharing training data for machine learning between airlines. When sharing data for this purpose, the issue of information being shared that about an aviation company that could potentially provide a rival with a competitive advantage also arises [2; 25]. Kabashkin and Susanin note that federated learning can be deployed in order to mitigate this issue. This entails multiple different airlines collaboratively training models without sharing any raw data [2]. Aircrafts learn from their own data and only share model updates with the central server. The server then aggregates the updates [25].

However, it is notable that federated learning is not without its limitations, as it creates additional computational complexity. Furthermore, it can be hampered by differences in the models used by airlines. It also requires careful coordination, meaning that it could potentially be demanding in terms of the staff competencies that are required in order to effectively facilitate it [2].

In addition to these points, federated learning involves the additional challenge of ensuring that the model generalises effectively across different operating conditions and aircrafts. Furthermore, cybersecurity is paramount and it is important to protect against vulnerabilities throughout the learning process. It is also notable that the airplanes generate substantial highly heterogenous datasets that necessitate efficient processing in order to facilitate real-time decision-making. However, the extensive data aggregation involved in federated learning can cause delays that can compromise safety in some cases [25]. This suggests that this is not always an ideal solution for privacy concerns.

Kabashkin and Susanin indicate that data privacy issues can also be mitigated by encrypting and anonymising sensitive information. They indicate that privacy constraints should be identified from the outset, as should data-sharing requirements. Comprehensive data-governance policies can further safeguard against the improper sharing or usage of data [2].

Stanton et al. highlight the expenses involved in AI-based predictive maintenance. They point out that it can be costly to attain the necessary resources and train staff to effectively utilise the requisite new technologies. Furthermore, they point out that the fact that aircrafts are complex, engineered objects contributes substantially to the costs that are involved. They highlight the fact that installing advanced predictive-maintenance systems in these vehicles requires specialist knowledge and especially intensive sensor-networks. According to them, the required expenses can create psychological barriers to aviation companies implementing AI-based preventative maintenance [28].

Kang et al. note that HetGPO, the current best-performing machine-learning-based method for resource scheduling in aircrafts based on ongoing feedback, lacks sufficient competency at modelling more complex multi-level and multi-dimensional relationships. This limits the usefulness of predictive maintenance based upon it, as the relevant models are unable to comprehensively capture the required depth of information from the available data [29].

Lee, Hussain and Oh also highlight the unsuitability of some AI-based approaches to predictive learning that are deployed in aircrafts for dealing with situations involving high complexity-levels. They point out that logistic regression is ineffective at dealing with complex, non-linear patterns [2]. This suggests that some of the current techniques for predictive maintenance are sometimes unable to cope with the complicated requirements of passenger aircrafts. However, Kang et al. proposed and tested a solution involving algorithms that leverage heterogeneous hypergraphs and found that they are more capable of handling the necessary levels of complexity [29]. This indicates that as the underlying models to continue to become more and more advanced, issues of this nature are likely to become less and less prominent throughout the years to come.

Klekowicki et al. indicate that insufficient access to real-world data can act as an impediment to the use of predictive-maintenance systems based on large language models. A large language model is a machine-learning model that is capable of designing and processing human-like text based upon a vast quantity of training data. Whilst synthetic, non-real-world data can be leveraged in order to mitigate issues relating to data access, it does not typically accurately replicate the anomalies and noise that naturally take place during the actual functioning of aircraft equipment. According to Klekowicki et al., these real-world irregularities often come about as the result of environmental factors, wear and tear and unpredictable environmental conditions and are essential for comprehensive diagnostics. This can limit the potential for accurate predictions and decrease the robustness of models for diagnosis when they are applied to live circumstances [31].

Berghout et al. claim that predictive-maintenance systems for aircrafts based on deep-learning models can often struggle in situations in which there is high levels of noise,

outliers, large numbers of required measurements, and high levels of distortion. They indicate that this makes it hard for the models to accurately capture data points and the relationships that exist between them. However, they indicate that this challenge can be overcome by implementing sufficient data-processing to mitigate the impact [32].

It is also notable that Berghout et al. highlight the need for additional research into outlier removal algorithms, denoising and data-complexity reduction methods. They indicate that this could potentially reveal insights that go some way towards remedying the aforementioned challenges [32]. This therefore appears to constitute a means of uncovering information that can help to circumnavigate this limitation.

The literature suggests that support-vector-machine-based and K-nearest-neighbour-based predictive-maintenance systems for aircrafts in particular tend to be sensitive to noisy datasets, outliers and missing data points. Support-vector-machine-based also tend to underperform in situations in which the number of features is greater than the number of dimensions. K-nearest-neighbour-based systems are unable to effectively deal with large numbers of dimensions and necessitate feature-scaling [28, 33]. Research conducted by Farina et al. involving the use of AI-driven predictive maintenance in aviation suggests that k-means clustering algorithms can sometimes produce false positives as a result of overlapping clusters [34]. This suggests that this is a challenge associated with this type of model.

According to Berghout et al., drift is another significant issue associated with predictive-maintenance systems based on deep learning [32]. They define this as "continuous change in statistical properties over time" [32]. It can occur as a result of variations in conditions and changes to the system's undergirding dynamics.

Berghout et al. [32] also draw attention to the need for data quality. They express the view that there is a pressing need for additional techniques to be developed that centre upon assessing it [32]. Research in this area could therefore potentially help to overcome this challenge.

Apostolidis, Bouriquet and Stamoulis highlight the fact that required data-quality cannot be considered to be fixed within the context of AI-based predictive-maintenance systems for aircrafts. They point out that it depends upon the desired outcome and application [35]. This indicates that there are also challenges involved in defining what actually constitutes a high standard of data in different situations.

Furthermore, it is notable that Korba et al. also highlight the fact that a high quality of data is necessary for implementing artificial-intelligence-driven predictive-maintenance systems in aircrafts while also acknowledging that effective data-management systems need to be in place [27]. This suggests that it is insufficient to merely have the necessary data; procedures also need to be in place for ensuring that it is adequately stored and managed. It highlights the role of systemic, organisational factors in the effective usage of this technology.

Kabashkin expresses concern that there is a danger that overconfidence in machine-learning-driven predictions could potentially result in human oversight [24]. He points out that for predictive maintenance driven by AI to be successful within the context of the aviation industry, it is essential to maintain "a balance between technological assistance and human expertise in decision-making" [24]. This indicates that there are challenges associated with obtaining this balance and that failure to do so could have life-threatening consequences.

The findings of research carried out by Lee at al. suggest that recurrent neural networks, which are an AI-based approach that is sometimes used in aircrafts, have high computational costs associated with them [29]. Kabashkin and Susanin also note that AI-based predictive-maintenance systems can be computationally demanding. Furthermore, they can necessitate extensive infrastructures, which could be challenging to smaller airlines. Optimisation techniques and lightweight AI models can help to mitigate this issue [2].

The literature also indicates that random-forest-based and particle-filler-based predictive-maintenance systems used in aviation are computationally intensive. The former also suffers from long training times and struggles to work out the significance of parameters. Convolutional neural networks also struggle with slow training-times and are unable to model long-term dependencies [28].

Lee et al.'s research indicates that recurrent neural networks suffer from a vanishing gradient problem, as do approaches involving Long Short-Term Memory. This reduces the effectiveness of these methods [30]. The former of these techniques also requires a considerable number of resources. Furthermore it, is incapable of modelling long-term dependencies. The latter suffers from high computational requirements and slow training times [28]. Zhou et al. note that Long Short-Term Memory models also have poor generalisation ability when under multiple different operating conditions or when multiple different faults are present [36]. Given the possibility of a diverse range of faults taking place in aviation, this arguably substantially reduces their suitability for usage within the context of aircrafts.

It is also notable that Lee et al. found that approaches involving Support Vector Machines can suffer from high processing-times when attempting to deal with high-dimensional data. The researchers point out that whilst convolutional neural networks appear to constitute the most effective approach for usage in predictive maintenance, they have not yet been validated for all aircraft components [30]. This arguably limits their utility.

Kabashin and Perekrestov note that it is both costly and technically challenging to integrate the new security measures that are required for AI-driven preventative maintenance with the existing aviation infrastructures. This could also be problematic in many cases and might necessitate expensive upgrades or even tailor-made solutions that fail to provide the same degree of protection as before. Furthermore, security threats need to be constantly evaluated in order to identify and address emerging threats to the systems [14].

According to Kabashin and Perekrestov, the solution to this lies in collaboration throughout the industry to create

standardised security-protocols. They claim that this can help to mitigate the additional security risks that AI brings with it. However, these protocols nevertheless still need to be regularly revised [14].

The findings of research carried out by Kumar et al. and Khalifeh and AlMeqdadi indicate that there is a trade-off between the time taken to complete the necessary computations and the accuracy of the results when it comes to the underlying mechanisms behind machine-learning-powered predictive-maintenance systems that are suitable for the aviation industry [35; 36]. Finding the optimum balance for each use case consequently constitutes an additional challenge [37]. Kumar et al. also highlight the fact that models can be complex to train [36]. This could potentially necessitate a high level of human expertise.

Yildirim and Yana express the view that some of the machine-learning models used for predictive maintenance in aircrafts lack the ability to effectively harness temporal correlation characteristics that are present within time series data. This can hamper their competence at making accurate and precise predictions in some circumstances [39]. This appears to constitute a challenge associated with the usage of such models that needs to be overcome.

Further to this, Stanton et al. indicate that the datasets that are utilised for training the machine-learning models for predictive-maintenance systems used in aviation are frequently imbalanced. This is due to the fact that faults are not common in aircrafts. This results in the available data being skewed towards normal operation, leading models to struggle to adequately learn about system failures. According to Stanton et al., in some cases there is a complete lack of failure data as faulty components are replaced prior to reaching the failure stage [28].

Stanton et al. also note that the substantial numbers of sensors in aircrafts can result in high dimensionality within the data that is collected [28]. This can result in what they refer to as the "curse of dimensionality", which is when "the higher the dimension space, the denser the data samples are required" [28, p. 230]. This constitutes yet another potential challenge.

Korba et al. highlight the need for incorporating AI-driven predictive maintenance into existing workflows. This suggests that there are challenges associated with factoring it into preexisting routines and processes. They point out that the algorithms that are utilised need to be able to be used in conjunction with existing operational procedures, databases and databases and sources of data. Failure to do this can result in disruptions to existing workflows [27].

Cheng et al. highlight the fact that there is a plethora of different machine-learning-based predictive-maintenance methods available that are suitable for use in aviation and indicate that this in and of itself poses challenges. They point out that each technique has its own unique limitations and applicable conditions associated with it. According to Cheng et al., this means that it is often hard to determine which method is likely to perform best in which situation. They claim that this issue is compounded by the fact that prediction performance can be compromised by applying deep-learning-based methods to diverse scenarios, meaning that the results suffer as a consequence. This can make it harder to pinpoint which method is best for which purpose [40].

The literature also highlights some challenges that specifically relate to deep-belief-network-based predictive-maintenance systems. They underlying models struggle to model temporal relaxations. Furthermore, they suffer from a tendency to overfit [28].

Furthermore, ongoing innovation and the need for continuous adaptation constitutes another challenge, as rapid technological development means that predictive-maintenance technologies need to be sufficiently flexible to incorporate new analytical methods, sensors and other technologies as they emerge. This necessitates a commitment to ongoing investment and improvement. As predictive maintenance systems incorporate more and more data from different sources throughout the years to come, the need for complex algorithms that are able to efficaciously combine and analyse disparate data types will grow. Changes will also be required in order to accommodate this [38].

Kabashkin and Perekrestov also note that AI models for predictive maintenance need to be capable of evolving with changing operational practices and aircraft technologies. They need to be easily updated in order to maintain their relevance and accuracy as time goes on [14]. This could also potentially prove challenging.

Now that the insights provided within the literature have been synthesised and detailed, it is necessary to present a detailed discussion of them. This allows for them to be understood within the context of the subject of the review. It is consequently done throughout the course of the subsequent chapter.

5 DISCUSSION

As can be seen in Fig. 2, the challenges in aircraft maintenance associated with the shift from preventative to AI-powered predictive maintenance fit broadly into five main categories. The first of these categories is finance- and resource-based challenges. These challenges include the fact that additional resources need to be put towards implementing predictive-maintenance systems and ensuring compliance [23, 28]. This category is connected to the next category; legal, accountability-oriented and ethical challenges; as the additional legal and ethical complications brought about by the move to machine-learning-oriented predictive maintenance necessitates the additional costs and resources [2]. Challenges within this category include questions regarding accountability, more complex considerations concerning compliance and difficulties in gaining approval for new systems [2, 23]. Ways of mitigating challenges in this area include a number of different measures geared towards bolstering transparency [2, 25].

The two aforementioned categories are also linked to the next category, human factors, as staff can be classed as a resource and human beings are inherently bound up in the ethical and legal decision-making within a firm. Human factors that pose challenges include attitudinal issues such as concerns about AI replacing jobs and scepticism about it and also those related to staff being adequately skilled to deal

with the new predictive-maintenance systems. Solutions involve change management and the provision of training [2].

All three of these categories are connected to the category of privacy and cybersecurity, as this requires resources and human input and breaches have potential legal consequences. Means of mitigating these challenges include encrypting and anonymising sensitive information and putting comprehensive data-governance policies in place [2].

The final category is performance- and efficiency-based issues, which is linked to human factors, as the models themselves are implemented by human actors, and finance- and resource-based challenges, as some are related to resources such as computing power [2]. A wide variety of challenges arose in this area, some specific to certain types of AI and some having implications across the spectrum. They include but are not limited to issues related to computational intensity, handling complexity, dealing with noise and anomalies and coping with outliers [2, 29, 31, 32]. Solutions include the careful choice of AI models to ensure the most suitable is selected for each purpose and a variety of strategies tailored to various individual problems [2, 32, 39].

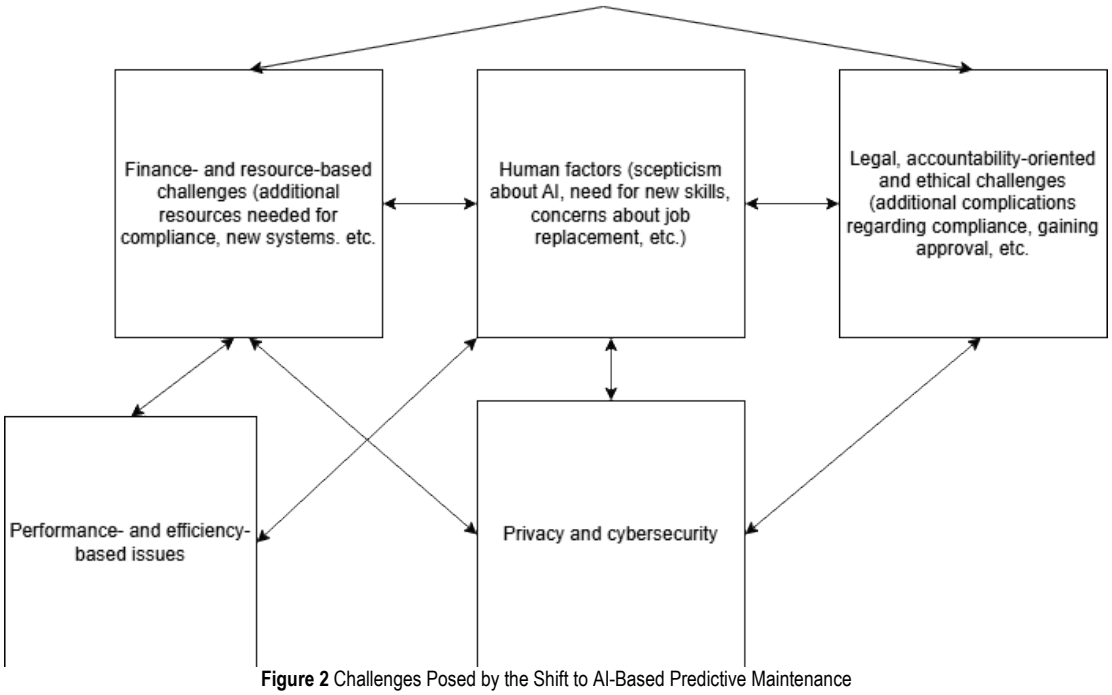


Figure 2 Challenges Posed by the Shift to AI-Based Predictive Maintenance

The findings also suggest are differences in the challenges posed by different AI models. They suggest that HetGPO lacks sufficient competency at modelling more complex multi-level and multi-dimensional relationships [29], whereas logistic regression is ineffective at dealing with complex, non-linear patterns [2]. They indicate that K-nearest-neighbour-based predictive-maintenance systems for aircrafts tend to be especially sensitive to noisy datasets, outliers and missing data points and that they are unable to effectively deal with large numbers of dimensions [28], whereas neural networks and random-forest-based and particle-filler-based predictive-maintenance systems have high computational costs associated with them [28, 29]. This suggests that rather than adhering to the perspective that there is a single optimum model, they should be chosen according to the degree to which their weaknesses compromise the specific purposes for which they are utilised.

Overall, the results highlight a multitude of different obstacles that need to be effectively navigated in a number of different areas. While some remedies emerged from the literature, other challenges do not appear to have a clear solution. This suggests a pressing need for additional research in this area.

6 CONCLUSION

This literature set out to identify the challenges associated with the challenges in aircraft maintenance associated with the shift from preventative to AI-powered predictive maintenance. It revealed that five main types of challenges exist: those related to the requisite financial and non-financial resources and the strain that implementing a new system entails with regards to them; human factors; legal, accountability-related and ethical challenges; novel privacy- and cybersecurity-related issues stemming from the new systems; and issues related to the performance and efficiency of the AI models. Pressing challenges that arose throughout the course of the review include the need for transparency to more easily demonstrate compliance; the limitations associated with the functionality of various different AI models; staff resistance to a new system and scepticism of AI; and economic constraints.

A number of different solutions arose from within the literature. Change management could be implemented to reduce scepticism and ill feeling towards AI. Training could be provided to staff in order to endow them with the relevant competencies. A concerted effort could be made to keep

sensitive operational data secure and federated learning could be deployed. Sensitive information could also be encrypted and anonymised. Privacy constraints should be identified from the outset, as should data-sharing requirements, and comprehensive data-governance policies can help to overcome challenges regarding upholding data privacy. Sufficient data processing can go some way towards remedying issues with noise and outliers, as can the careful selection of AI models that are appropriate for the specific purpose within aviation that they are being used for.

Further to this, efforts could also be made to strike a balance between human diligence and the use of AI in order to avoid errors stemming from overconfidence in the latter. Lightweight AI models can be deployed to mitigate issues with computational requirements. A commitment to ongoing investment in AI-based predictive maintenance and the improvement of systems can also help to safeguard against future issues.

It is notable that the rapid rate at which technology changes means that the findings of this review might not necessarily hold true within the immediate future. It is therefore recommended that future literature reviews on this topic are carried out at regular interviews. Furthermore, given the heterogenous nature of machine learning models, primary research focussing upon specific models should also be conducted at frequent intervals.

Another limitation of this study is the fact that it merely reviewed published sources, meaning that grey literature and sources yet to be published were excluded. It can therefore be considered to be a review of a large number of relevant sources as opposed to being exhaustive. Its findings could also potentially have been impacted upon by publication bias.

Furthermore, given that the search strategy was devised by the researchers, the possibility exists that this might have influenced the sources that were included in the review. However, the use of a rigorous search strategy means that efforts were made to reduce the extent to which this affected the results. Efforts were also made to ensure that the keywords were as tightly related to the subject of the review as possible.

In order to enhance aviation companies' ability to further mitigate the challenges that have been identified, it is recommended that additional research is carried out on predictive maintenance that takes the industry's strict safety requirements into consideration given that a dearth of empirical research in this area was mentioned within the literature. Furthermore, given that the literature identified a lack of studies that evaluate the feasibility of machine-learning-driven predictive-maintenance systems across critical operational and economic dimensions, it is also advised that this is an area in which additional research is required.

It is also recommended that future studies should focus upon who is accountable for faults on the part of the AI involved in predictive-maintenance systems, given that the literature also indicates a gap in this area. It is also advised that additional research into outlier removal algorithms, denoising and data-complexity reduction methods should take place. Future studies into interpretable models for machine learning and explanation frameworks intended for blackbox models could also help to mitigate issues with the transparency of the AI used in predictive maintenance in aviation.

This literature review has helped to provide a clear insight into the various different obstacles, problems and pitfalls associated with the ongoing shift towards the usage of AI and predictive maintenance in aviation. It has shed light on not only what challenges have manifested but also those that are likely to become prominent throughout the years to come. Furthermore, it has illuminated the precise form that some of these issues take.

Furthermore, the study has shown what the current base of literature says about how to overcome some of the challenges that were identified. Although solutions were not identified for some problems, it is nevertheless useful to gain an insight into their characteristics so that future remedies can be established. This paper has therefore contributed a thorough overview of the issues associated with this facet of Industry 4.0 within the context of aviation. Given the importance of safety within this field, its findings have clear utility and can act as a firm basis for future research in this domain.

7 REFERENCES

[1] Karim, R., Galgar, D., & Kumar, U. (2023). *AI factory: Theories, applications and case studies*. Boca Raton, FL: CRC Press, 2023. https://doi.org/10.1201/9781003208686

[2] Kabashin, I., & Susanin, V. (2024). Unified ecosystem for data sharing and AI-driven predictive maintenance in aviation. *Computer*s, *13*(12). https://doi.org/10.3390/computers13120318

[3] Doogakant, B., Fowdur, T., Milovanovic, D., & Bojkovic, Z. (2025). 5G/6G-based sustainable systems for Industry 4.0, in T. Fowdur, D. Milovanovic & Z. Bojkovic (eds.), *Intelligent and Sustainable Engineering Systems for Industry 4.0 and Beyond* (pp. 29-58). CRC Press. https://doi.org/10.1201/9781003511298-2

[4] Bhatia, V., Sidhartha, S., Khare, S., Ghorpade, S., Kumar, P., Kumar, A., & Agarwal, A. (2024). Intelligent manufacturing in aerospace: Integrating Industry 4.0 technologies for operational excellence and digital transformation, in A. Kumar, P. Kumar & Y. Liu (eds.), *Industry 4.0 Driven Manufacturing Technologies*. Springer Natuire Switzerland. https://doi.org/10.1007/978-3-031-68271-1_18

[5] Mazumdar, B., Mishra, M., Ghatak, A., & Verma, V. (2025). AI for industry 4.0 with real-world problems, in A. Tyagi, S. Tiwari and G. Soni (eds.), *Data Analytics and Artificial Intelligence for Predictive Maintenance in Smart Manufacturing* (pp. 326-352). CRC Press. https://doi.org/10.1201/9781003480860-15

[6] Sivakumar, M., Maranco, M., & Krishnaraj, N. (2025). Data analytics and artificial intelligence for predictive maintenance and manufacturing, in A. Tyagi, S. Tiwari and G. Soni (eds.), *Data Analytics and Artificial Intelligence for Predictive Maintenance in Smart Manufacturing* (pp. 29-55). CRC Press. https://doi.org/10.1201/9781003480860-3

[7] Kumar, D. & Batra, J. (2024). The pioneering convergence of data analysis and predictive maintenance. In Gatti R., & Singh, C. (eds.), *Shaping the Future of Automation with Cloud-enhanced Robotics* (pp. 38-64). IGI Global.

https://doi.org/10.4018/979-8-3693-1914-7.ch003
[8] Shukla, A., & Dubey, K. (2024). Deployment issues in industrial resolution. In Dubey, A., Yadav, V., & Trivedi, M. (eds.), *Computational Intelligence in the Industry 4.0* (pp. 161-188.). CRC Press. https://doi.org/10.1201/9781003479031-10
[9] Ucar, A., Karakose, M., & Kirimca, N. (2024). Artificial intelligence for pedictive maintenance applications: Key components, trustworthiness, and future trends. *Applied Sciences*, *4*(2). https://doi.org/10.3390/app14020898
[10] Padmavathi, S., Nevetha, B., Sakthi, R. & Varshini, K. (2025). Machine learning for reliable industrial operations: Predictive maintenance. In Murugan, T., & Singh, W. (eds.), *Cybersecurity and Data Science Innovations for Sustainable Development of HEICC* (pp. 316-330). CRC Press. https://doi.org/10.1201/9781032711300-22
[11] Kristoffersen, E., Aremu, O. O., Blomsma, F., Mikalef, P., & Li, J. (2019). Exploring the Relationship Between Data Science and Circular Economy: An Enhanced CRISP-DM Process Model. In Pappas, I. O., Mikalef, P., Dwivedi, Y. K., Jaccheri, L., Krogstie, J., & Mäntymäki, M. (eds) *Digital Transformation for a Sustainable Society in the 21st Century. I3E 2019*. Lecture Notes in Computer Science, vol 11701. Springer, Cham. https://doi.org/10.1007/978-3-030-29374-1_15
[12] Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management*, *60*. https://doi.org/10.1016/j.ijinfomgt.2021.102383
[13] Kavanagh, K. (2024). *Google machine learning and generative AI for solutions architects*. Packt Publishing.
[14] Kabashkin, I., & Perekrestov, I. (2024). Ecosystem of aviation maintenance: Transition from aircraft health monitoring to health management based on IoT and AI synergy. *Applied Sciences*, *14*(11). https:// doi.org/10.3390/app14114394
[15] Hicham, T., Khalifa, M., Kamal, E., & Fatiha, A. (2023). Machine and deep learning applications in Industry 4.0, in Institute of Electrical and Electronics Engineers (ed.), *Proceedings of the 2023 International Conference on Technology, Engineering, and Computing Applications (ICTECA).* Institute of Electrical and Electronics Engineers. Retrieved from https://ieeexplore.ieee.org/document/10490844
[16] Gera, I., & Singh, S. (2024). Integration revolution: Building trust with technology. In Leong, W., Chuah, J., & Tee, B. (eds.), *The Nine Pillars of Technologies for Industry 4.0*. Institution of Engineering & Technology. https://doi.org/10.1049/PBTE088E_ch10
[17] Suleiman, Z., Shaikholla, S., Dikhanbayeva, D., Shehab, E., & Turkyilmaz, A. (2022)., Industry 4.0: clustering of concepts and characteristics. *Cogent Engingeering*, *9*(1). https://doi.org/10.1080/23311916.2022.2034264
[18] Silvestri, L., Forcina, A., Introna, V., Santolamazza, A., & Cesarotti, V. (2020). Maintenance transformation through Industry 4.0 technologies: A systematic literature review. *Computers in Industry*, *123*. https://doi.org/10.1016/j.compind.2020.103335
[19] Jasiulewicz-Kaczmarek, M. (2024). Maintenance 4.0 technologies for sustainable manufacturing. *Applies Sciences*, 14(6). https://doi.org/10.3390/app14167360
[20] Vasquez, G., & Castrillon, H. (2022). Scientific literature review for the search of vulnerabilities and pentesting tools for SCADA system. In Mata-Rivera, M., & Zagal-Flores, R. (eds.), *Telematics and Computing* (pp. 90-103). Springer Nature Switzerland. https://doi.org/10.1007/978-3-030-89586-0_7
[21] Zaaba, M., Zainuddin, S., Abdullah, T., Zaib, S., Mohamed, A., Kadir, H., & Hashim, N. (2024). Blockchain technology and food supply chain: A bibliometric analysis. In El Khoury, R. (ed.), *Technology-Driven Business Innovation: Unleashing the Digital Advantage* (vol. 2) (pp. 531-540). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-62656-2_47
[22] Ma, D. (2023). A bibliometric viualisation of CNKI-based visual metaphor studies. In Chen, C., Singh, S., Saxena, S., & Wheeb, A. (eds.), *Proceedings of the 3rd International Conference on Digital Economy and Computer Application (DECA 2023)* (vol. 17) (pp. 644-659). Atlantis Press. https://doi.org/10.2991/978-94-6463-304-7_68
[23] Kabashkin, I., Federov, R. & Perekrestov, V. (2025). Decision-making framework for aviation safety in predictive maintenance strategies. *Applied Sciences*, *15*(3). https://doi.org/10.3390/app15031626
[24] Kabashin, I. (2024). The iceberg model for integrated aircraft health monitoring based on AI, blockchain, and data analytics. *Electronics*, *13*(19). https://doi.org/10.3390/electronics13193822
[25] Kabashkin, I. (2024). Integration of foundation models and federated learning in AIoT-based aircraft health monitoring systems. *Mathematics*, *12*(21). https://doi.org/10.3390/math12213428
[26] Apostolidis, A., Pelt, M., & Stamoulis, K. (2020). Aviation data analytics in MRO operations: Prospects and pitfalls, in Institute of Environmental Sciences and Technology (ed.), *Proceedings of the 2020 Annual Reliability and Maintainability Symposium (RAMS 2020)*. Institute of Environmental Sciences and Technology. https://doi.org/10.1109/rams48030.2020.9153694
[27] Korba, P., Svab, P., Veres, M., & Lukac, J. (2023). Optimizing aviation maintenance through algorithmic approach of real-life data. *Applied Sciences*, *13*(6). https://doi.org/10.3390/app13063824
[28] Stanton, I., Munir, K., Ikram, A., & El-Bakry, M. (2022). Predictive maintenance analytics and implementation for aircraft: Challenges and opportunities. *Systems Engineering*, 26(2), 216-237. https://doi.org/10.1002/sys.21651
[29] Kang, L., He, M., Zhou, J., Hou, Y., Xu, B., & Liu, H. (2025). A method for the predictive maintenance resource scheduling of aircraft based on heterogeneous hypergraphs. *Electronics*, 14(4). https://doi.org/10.3390/electronics14040782
[30] Lee, K., Hussain, K., & Oh, I., (2025). Validation for predictive maintenance of aircraft systems using AI models developed with rotor blade vibration data. *IEEE Access*, 13, 48173-48178. https://doi.org/10.1109/ACCESS.2025.3550822
[31] Zhou, M., Miu, K., Sun, J., Shen, Y., & Han, B. (2024). Data-driven modeling of aero-engine performance degradation models. *IEEE Access*, *12*, 150020-150031. https://doi.org/10.1109/ACCESS.2024.3460872
[32] Berghout, T., Bentrcia, T., Lim, W., & Benbouzid, M. (2023). A neural network weights initialization approach for diagnosing real aircraft engine inter-shaft bearing faults. *Machines*, *11*(19). https://doi.org/10.3390/machines11121089
[33] Basora, L., Olive, X., & Dubot, T. (2019). Recent advances in anomaly detection methods applied to aviation. *Aerospace*, *6*(11). https://doi.org/10.3390/aerospace6110117
[34] Farina, D., Machrafi, H., Queeckers, P., Dong, P., & Iorio, C. (2024). Innovative AI-enhanced ice detection system using graphene-based sensors for enhanced aviation safety and efficiency. *Nanomaterials*, *14*(13). https://doi.org/10.3390/nano14131135

[35] Apostolidis, A., Bouriquet, N., & Stamoulis, K. (2022). AI-based exhaust gas temperature prediction for trustworthy safety-critical applications. *Aerospace*, *9*(11), 1-16. https://doi.org/10.3390/aerospace9110722
[36] Klekowicki, M., Szymanski, G., Waligorski, M., & Misztal, W. (2025). Application of large language models in diagnostics and maintenance of aircraft propulsion systems. *Advances in Science and Technology Research Journal*, *19*(2), 304-320. https://doi.org/10.12913/22998624/196264
[37] Kumar, D., Sowmya, B., Kanavalli, A., Supreeth, S., Shruthi, G., & Rohith, S. (2024). Deep learning methodologies based on metaheuristics for predictive engine maintenance. Acta *IMEKO*, *13*(2), 1-15. https://doi.org/10.21014/actaimeko.v13i2.1667
[38] Khalife, A., & AlMeqdadi, S. (2024). Machine learning-based remaining useful life predictions and its application on predictive maintenance, in Institute of Electrical and Electronics Engineers (eds.), *Proceedings of the 22nd International Conference on Research and Education in Mechatronics (REM)*. Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/REM63063.2024.10735638
[39] Yildrim, S., & Rana, Z. (2024). Enhancing aircraft safety through advanced engine health monitoring with Long Short-Term Memory. *Sensors*, 24(2). https://doi.org 10.3390/s24020518
[40] Cheng, Y., Zeng, J., Wang, Z., & Song, D. (2023). A health state-related ensemble deep learning method for aircraft engine remaining useful life prediction. *Applied Soft Computing*, *135*. https://doi.org/10.1016/j.asoc.2023.110041

Authors' contacts:

Belma Atapek-Yağan, Dipl. Ing., Ph.D. student
(Corresponding author)
Department of Materials and Technology,
Faculty of Electrical Engineering, University of West Bohemia,
Univerzitní 2795/26, 301 00 Pilsen, Czech Republic
belma22@fel.zcu.cz

Jiří Tupa, Doc. Ing. PhD
Department of Materials and Technology,
Faculty of Electrical Engineering,
University of West Bohemia,
Univerzitní 2795/26, 301 00 Pilsen, Czech Republic
tupa@fel.zcu.cz

František Steiner, Doc. Ing. PhD
Department of Materials and Technology,
Faculty of Electrical Engineering,
University of West Bohemia,
Univerzitní 2795/26, 301 00 Pilsen, Czech Republic
steiner@fel.zcu.cz

CC BY

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20250922083008

Subject review
Received: 2025-09-22, Accepted: 2026-02-12

Potential, Challenges, and Opportunities in the Development of Geothermal Energy Projects

Pavao Mesarić, Lovorka Gotal Dmitrović*, Boris Jalušić

Abstract: Geothermal energy is an important yet insufficiently explored renewable source. Known for its supply stability and low environmental impact, it offers significant potential, however, broader adoption faces notable technical, economic, and regulatory challenges. This paper first examines the origin of geothermal energy and its historical use, followed by the technological evolution of systems for both direct thermal use and electricity generation. It also reviews global trends and statistical indicators, providing context for current utilization patterns. The analysis places special emphasis on the Republic of Croatia, highlighting existing projects, available geological data, and strategic development plans. Key barriers to wider deployment are identified, and potential solutions are discussed, including legislative adjustments, investment incentives, and stronger institutional support. The paper concludes by emphasizing the need for a systematic approach to integrate geothermal energy into Croatia's energy transition, thereby enhancing energy independence and reducing greenhouse gas emissions.

Keywords: Croatia; energy independence; energy transition; geothermal energy; geothermal power plants; renewable energy

1 INTRODUCTION

The increasing energy demands of modern society require innovative energy solutions. In this context, renewable energy sources will play a key role. In addition to the inherent limitations of fossil and nuclear fuels as global resources and their actual and potential harmful environmental impacts, energy supply chains are under growing pressure due to the current unstable geopolitical situation. Therefore, new energy solutions are being sought that, in addition to reducing greenhouse gas emissions and the exploitation of finite natural resources, must also ensure energy independence through local energy production. Geothermal energy naturally emerges as an attractive alternative. Compared to renewable sources such as solar or wind energy, this form of energy enables energy production both day and night, regardless of weather conditions and throughout the year, which makes its supply stable and predictable. Geothermal energy is a virtually inexhaustible, clean, and renewable energy source, but it faces numerous challenges, and its utilization is currently rather limited.

2 ORIGIN OF GEOTHERMAL ENERGY

Geothermal energy is the thermal energy present within the Earth's interior that is, in the Earth's crust, mantle, and core. There are two main sources of this heat: primordial or accretion heat, i.e., the heat remaining from the formation of the Earth, and the heat released due to the decay of radioactive elements in the Earth's interior [1]. In [2] another important mechanism of heat source within the Earth is proposed, namely, a gravitational heat source generated by viscous friction in the upward flows of material deep from the mantle towards the surface. These flows occur because mantle material gradually loses iron over time. The denser iron sinks towards the core, while the remaining, lighter material rises, bringing additional heat closer to the surface. The heat released by these mechanisms drives further processes that enhance local heat transfer within the Earth's interior all the way to the surface. Therefore, in addition to pure heat conduction, convective heat transfer occurs due to the movement of viscous mantle material, which consequently causes shearing, subduction, overthrusting, collision and uplifting of tectonic plates, processes that also release large amounts of energy. As a result, geothermal energy manifests at the Earth's surface in the form of volcanoes, geysers, hot springs, and similar phenomena.

3 HISTORICAL AND PRESENT-DAY USES OF GEOTHERMAL SOURCES

Humans have always been aware of the existence of geothermal energy because they witnessed the previously mentioned surface manifestations. There is evidence that people used hot springs for bathing and cooking more than ten thousand years ago. The oldest known bathhouses that used geothermal water were built in China during the Qin dynasty in the 3rd century BC. Many bathhouses from the time of the Roman Empire are also well known, and the Romans used geothermal springs for medicinal purposes also, particularly for treating skin and eye diseases, as they believed that the minerals in geothermal water had healing properties. The Romans also used geothermal sources for underfloor heating in buildings, such as in Pompeii. The oldest centralized heating system was built in France in the 14th century. In addition to bathing, healing, and building heating, geothermal water is also valuable as a resource. For example, in the early 20th century in Tuscany, Italy, it was used to produce boric acid [3]. On that same geothermal field, Larderello, the first geothermal power plant was commissioned in 1904. Today, the Larderello field is one of the largest geothermal fields in the world with a total installed capacity for electricity generation of 800 MW [4]. Today, geothermal sources are used for a wide range of purposes. They are used both as an energy source and as a source of raw materials. As an energy source, the most common use is the direct use of geothermal heat. Some of the applications of this heat include district heating of urban areas, heating of large facilities such as industrial plants or warehouses, heating of greenhouses for growing fruits, vegetables, and

flowers, heating of fish farms, heating of pools, spas, and wellness centers, de-icing of roads, drying of wood, fruits, vegetables, and similar products, food processing such as milk pasteurization, salt production, textile washing and dyeing, and so on. It is also important to mention the use of heat pumps for households and geothermal cooling [5]. In addition to the direct use of heat, power plants that utilize geothermal sources for electricity production are gaining increasing importance. Finally, geothermal sources are also used to produce raw materials such as lithium, boron, silicon, and others [6].

4 CRITERIA FOR SELECTING OPTIMAL LOCATION AND TYPE OF GEOTHERMAL SYSTEM

Not all geothermal systems are suitable for all locations. For example, systems for direct use of heat can be constructed in areas where the geothermal reservoir has a significantly lower temperature than what would be required for the construction of a geothermal power plant. In this context, the concept of the geothermal gradient is of primary importance. It is a measure that describes how much the subsurface temperature increases with depth. It is usually expressed in °C/km. In the continental crust, it typically amounts to about 25 °C/km down to depths of approximately 3 to 5 km [7]. In addition to the fact that the geothermal gradient is not constant at a given location, that is, it varies with different geological layers and with increasing depth, there are also large differences in geothermal gradients between different regions. Fig. 1 shows a map of Europe with temperatures at the depth of 4500 m. It is evident that the highest subsurface temperatures are found in countries such as Iceland and Turkey, while regions such as the Pannonian Basin, parts of Italy, Germany, and France also exhibit a higher geothermal gradient than the European average.

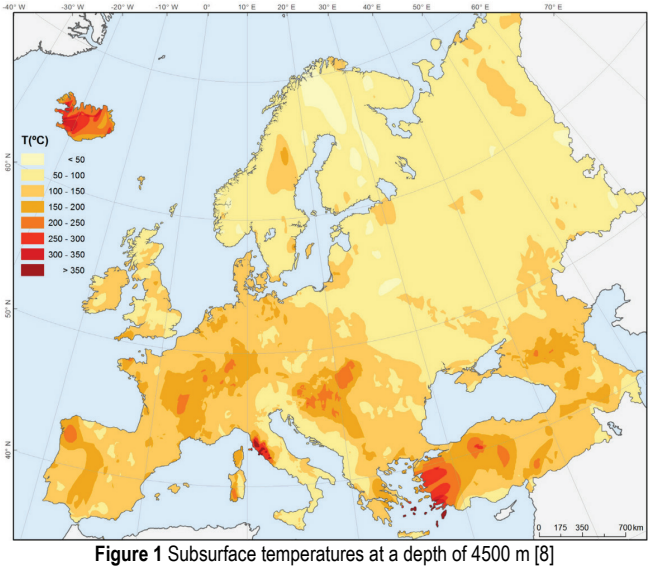


Figure 1 Subsurface temperatures at a depth of 4500 m [8]

Regardless of the type of geothermal system, the selection of a location requires the consideration of a wide range of different criteria. This is a multidisciplinary process that involves evaluating numerous factors. Primarily, these include locational criteria such as geographical and geological characteristics, but other aspects must also be considered, such as economic, regulatory, and social factors, in order to assess the need, feasibility, profitability, and acceptability of a geothermal project. For example, the construction of a geothermal system will be unfavorable or impossible on inaccessible terrain, in populated areas or natural protected areas, on remote mountains, or on bodies of water.

When evaluating geological criteria, it is first important to pay attention to surface manifestations. Clearly visible geothermal activity, such as the proximity of hot water springs, volcanic activity, and geysers, are clear indicators of favorable geological conditions. Additionally, gas and steam emissions from the ground may signal geothermal activity. Further indicators include craters, calderas, and other similar formations. These structures often occur alongside extrusive igneous rocks such as trachyte, rhyolite, and basalt, which further indicate favorable geothermal characteristics of the area. In addition to surface indicators, possible signs also include faults and fracture zones that allow geothermal fluids to flow through the rocks, acting as high-permeability channels [9].

In addition to surface indicators, valuable data is provided by surface geophysical measurements. These include multispectral and hyperspectral imaging, InSAR, LiDAR data, and similar sources. Subsurface data is obtained through methods such as seismographic, magnetotelluric, and gravimetric surveys. These exploration methods are useful for detecting large geological structures, rock composition, subsurface layer thickness, and general characteristics of a potential geothermal reservoir. Geophysical survey methods are typically used to analyze parameters such as density and to identify anomalies [10].

However, by far the highest quality and most reliable data comes from actual wells. The drilling of exploratory wells is the most expensive method of obtaining subsurface data, but it provides the most accurate information about the geothermal reservoir, including lithological data, permeability, porosity, and the overall quality of the reservoir. It also yields detailed data on the geothermal fluid (water) such as its composition, salinity, the quantity and type of dissolved minerals and gases, and similar characteristics. Naturally, it also provides information about reservoir temperature (of both water and rock) and geothermal gradients.

By using all this data, a precise 3D representation of the subsurface structure can be made. Such models can help assess the profitability of the project and allow engineers to plan drilling locations more efficiently, define well trajectories to maximize thermal output, and carry out economic analysis to select the most favorable design variant of the geothermal system [10]. A large amount of data is often already available, so it is always necessary to thoroughly examine existing archive data. In the Republic of Croatia, around 4500 wells have been drilled so far [11], which is, relative to the country's territory and the fact that Croatia is not a particularly significant producer of oil and

gas, a very large number. Therefore, there is a great deal of data on the Croatian subsurface, and it is considered well explored. Such a situation represents a major advantage, as it enables a much better assessment of geothermal potential and reduces the costs of data collection and exploration. In addition, old wells that were considered negative in terms of hydrocarbon production can be repurposed into geothermal wells, which significantly reduces costs.

Based on data on earthquakes, heat flow, sediment thickness, groundwater, tectonic data, elevation, and other factors, Coro and Trumpy [12] created a map shown in Fig. 2. The map shows the most promising areas on Earth for the development of geothermal systems, primarily located along the Pacific Ring of Fire, in China, the Middle East, parts of Africa at the boundary between the Nubian and Somali plates, Hawaii, New Zealand, Iceland, the Azores, and parts of Europe from Turkey towards southeastern and central Europe and Italy. Although geothermal systems can be constructed almost anywhere, these areas are significantly more suitable for their construction.

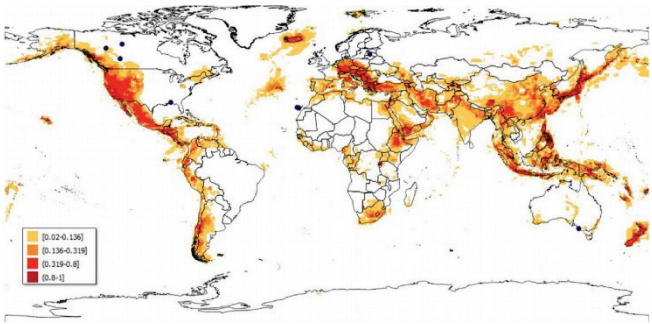


Figure 2 The most promising areas for the development of geothermal systems on Earth [12]

5 TYPES OF GEOTHERMAL SYSTEMS

Today, there are various types of systems for direct heat use. Direct heat systems range from simple, shallow pipe setups used for household heating and cooling, such as heat pumps, to large-scale systems with deep wells designed for heating urban districts or agricultural facilities like greenhouses.

Geothermal power plants can be divided into conventional and unconventional types. Conventional geothermal power plants are further classified into three main types: dry steam systems, flash steam systems, and binary cycle systems. In dry steam systems, steam from underground reservoirs is used directly to drive a turbine and generate electricity. These systems require high reservoir temperatures (150 °C or above) where dry steam is present. This system is also the first and simplest type of geothermal power plant ever developed. The flash steam system uses geothermal water in liquid form at a temperature of around 180 °C under very high pressure. The water is brought to the surface through a well into high-pressure tanks (separators). When the pressure drops in the tank (in one or more stages), the water flashes into steam, which is then directed to the turbine. The binary cycle differs from the previous two types in that geothermal water is not used directly to generate electricity. Instead, in a heat exchanger, the geothermal water transfers its heat to another working fluid. This working fluid, vaporized by the heat, drives the turbine. The working fluid is typically a fluid with a lower boiling point than water. This approach enables the system to operate with geothermal water at temperatures as low as 100 °C or even lower, making it suitable for lower-temperature geothermal reservoirs [13].

The second type includes unconventional geothermal systems. These are divided into two main groups: enhanced geothermal systems (EGS) and closed-loop geothermal systems (CLGS). In enhanced geothermal systems, a geothermal reservoir is not required, although it can still be used if present. What is needed are hot rocks, either dry (HDR) or wet (HWR). These systems artificially improve or create rock permeability between production and injection wells by hydraulic fracturing, allowing water to circulate and transfer heat [14]. In closed-loop geothermal systems, there is no geothermal water present in the system. Instead, closed pipe circuits are used to circulate a working fluid through underground formations, where it absorbs heat from hot rocks [15]. There are different designs of closed systems, such as U-tubes, multilateral wells, coaxial systems, and others. The types of conventional geothermal power plants are shown in Fig. 3, while the types of unconventional geothermal power plants are shown in Fig. 4.

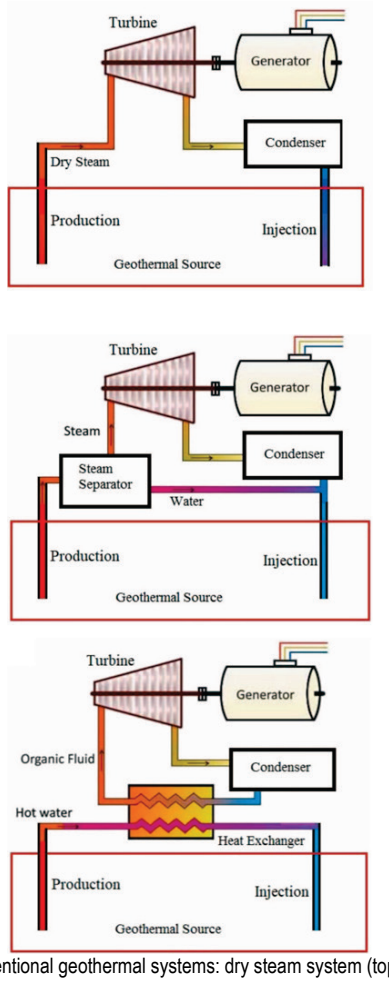


Figure 3 Conventional geothermal systems: dry steam system (top), flash steam system (center), and binary cycle (bottom) [13]

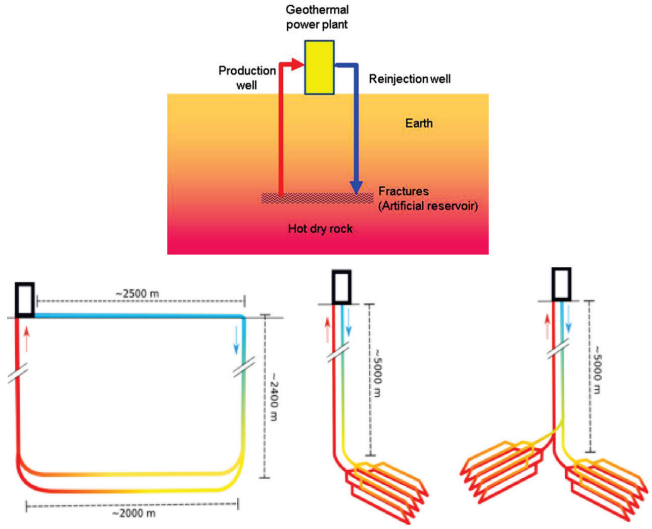


Figure 4 Unconventional geothermal systems: EGS (top) [14] and CLGSs (bottom) [15]

There are also hybrid systems that combine features of two or more systems, as well as so-called cascade systems, in which geothermal water is used at multiple levels (e.g., first for electricity production, and then the remaining warm water, before reinjection into the reservoir, is used for heating greenhouses).

6 USE OF GEOTHERMAL ENERGY WORLDWIDE

The International Geothermal Association (IGA) regularly updates its database on the use of geothermal energy around the world. Data on installed geothermal energy capacity is shown in the charts in Fig. 5.

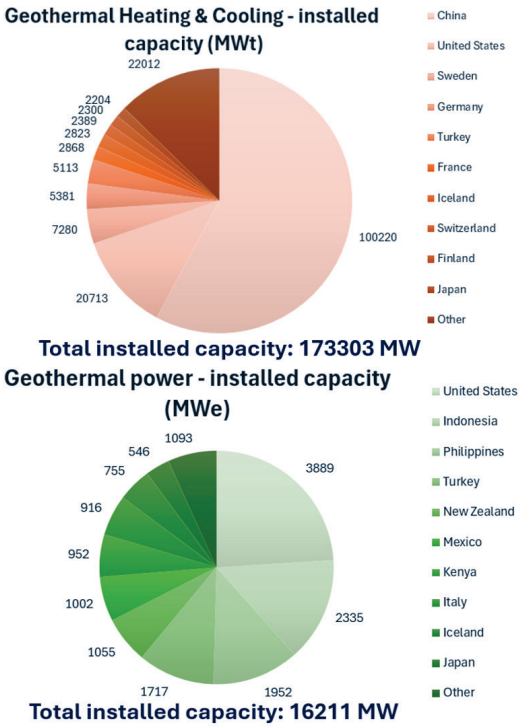


Figure 5 Installed geothermal capacities worldwide [16]

From the chart in Fig. 5, it is evident that, when it comes to direct use of heat, China accounts for more than half of the world's capacity, and the top five countries alone account for over 80%. In the case of electricity production, the capacities are distributed somewhat more evenly among the top ten countries. It is interesting to note that most of the electricity generation capacity is located in countries along the Pacific Ring of Fire (USA, Indonesia, the Philippines, New Zealand, Mexico, Japan). A large share of the remaining installed capacity is found in Turkey, Italy, and Iceland, as well as in Kenya, a significant representative of geothermal energy in Africa. Also, although the capacities for direct heat use and for electricity generation are not easily directly comparable, it is still evident that geothermal energy today is far more commonly used as a direct source of heat than for electricity production. The reason for this lies in the already mentioned fact that much more prominent geothermal reservoir characteristics are required for electricity production. When this data is compared with the map in Fig. 2, a significant correlation can be observed. This fact confirms the great importance of geographical location, i.e., geological conditions, for the potential of geothermal energy use.

Taking into account the estimate that just within ten kilometers below the Earth's surface lies a geothermal energy potential 50 thousand times greater than the energy potential of all oil and gas reserves in the world [17], and that geothermal energy accounts for only 0.34% of global electricity production [18], it can be concluded that geothermal energy is still significantly underutilized as an energy source and that there is an enormous room for growth.

7 GEOTHERMAL ENERGY IN CROATIA

The favorable position of Croatia in regard to geothermal energy is already visible from Fig. 1 and Fig. 2. The Pannonian region of Croatia exhibits a geothermal gradient up to 60% higher than the European average [19]. Using temperature data from 154 wells and by preforming interpolation, Macenić et al. [20] produced a detailed geothermal gradient map. Fig. 6 presents this gradient map, while Fig. 7 shows the geothermal fields in Croatia, including granted permits, projects in the exploration phase, and projects in the production phase.

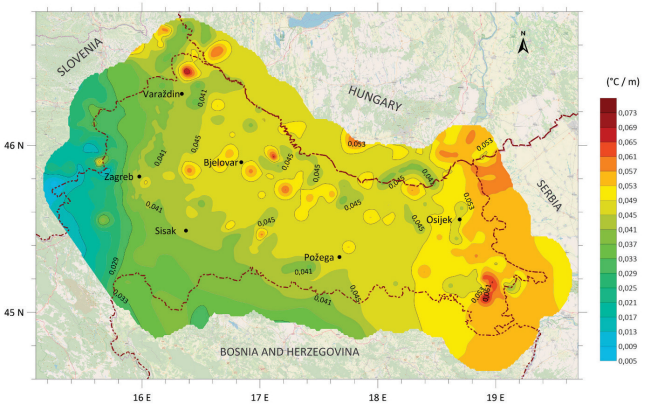


Figure 6 Geothermal gradient map of Croatia [20]

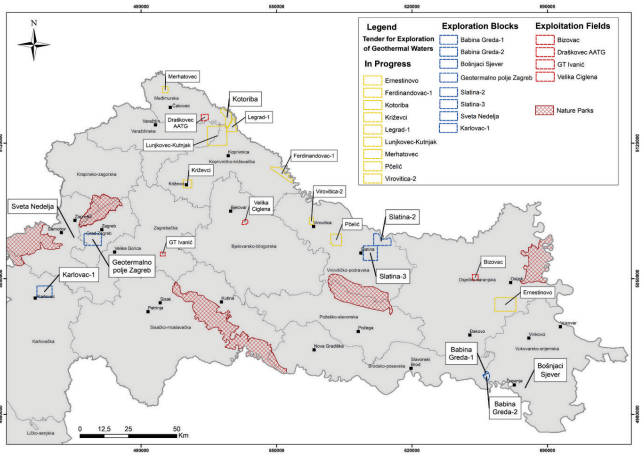


Figure 7 Geothermal fields in Croatia [20]

Croatia currently has 75 MW of installed capacity for thermal energy and 16.5 MW of installed capacity for electricity production [16]. It is estimated that the total geothermal potential in Croatia ranges between 750 and 1300 MW for thermal energy and about 500 MW for electricity production [21], with more recent estimates suggesting that the potential for electricity production could reach up to 1000 MW [22]. The 16.5 MW actually come from just one location, Velika 1 geothermal power plant, the only geothermal power plant in Croatia, operational since 2019, and the largest binary cycle geothermal power plant in mainland EU [23]. According to data from 2023, only around 0.21% of gross electricity production in Croatia comes from geothermal sources [24]. Based on the above, it can be concluded that Croatia utilizes only about 6 to 10% of its total potential for direct use of geothermal heat and only around 1.7 to 3.3% of its total potential for geothermal electricity production. Therefore, it is expected that installed capacities in Croatia will increase. This claim is supported by the fact that there are currently 24 active exploration areas in Croatia, 12 production fields are already in the production phase, and six preliminary exploration areas are focused on the development of district heating systems [25].

8 CHALLENGES IN THE DEVELOPMENT OF GEOTHERMAL PROJECTS AND POSSIBLE SOLUTIONS

When examining the challenges in developing geothermal projects, it is useful to categorize them broadly into technical, economic, regulatory, social, and other challenges.

8.1 Technical Challenges

Drilling technology is a proven technology whose modern development began in the 19th century when large-scale drilling started in search of oil. Although the drilling process for hydrocarbon exploration and for geothermal project development is very similar, there are certain differences. These primarily relate to the temperatures encountered during drilling and the type and characteristics of the rocks. High temperatures limit the performance and lifespan of drilling tools and reduce the effectiveness of drilling fluids. Drilling fluids easily degrade, lose functionality, and the accuracy of measurement instruments during drilling decreases. Rubber parts have difficulty achieving proper sealing, and controlling the drilling trajectory is more challenging. Additionally, high temperatures affect the stability of cement slurry and the quality of cementing, which complicates later reservoir development. Since the goal is often to find fractured rocks, circulation losses of drilling fluids frequently occur as fluids enter the fractures. Favorable geothermal reservoirs are also often located in hard rocks, which further complicates drilling [26]. Moreover, geothermal wells are usually of somewhat larger diameter than oil wells because, especially for lower temperature reservoirs, a higher water flow rate is required for electricity production [27]. Also, during the construction of geothermal wells, it is often necessary to install casing along the entire length of the well, which is not always required in oil production.

The problems mentioned, although solvable, make geothermal projects much more expensive. Solutions include the development of temperature resistant fluids, instruments and tools, advanced cementing methods, and innovative drilling technologies (gas or foam instead of drilling mud, PDC bits, etc.), but further research and development of drilling technology for even higher temperatures and depths are needed [26].

However, the biggest technical challenges actually relate to the unreliability of surface measurements. No matter how good the data obtained from these measurements are, they can never be completely reliable. Before a reservoir is reached by drilling, the quality of the reservoir cannot be determined with certainty. Data such as permeability, porosity, reservoir fluid pressure, temperature, composition of reservoir fluids, etc., can only be reliably obtained by sampling rocks and fluids directly from the reservoir. It is a great advantage when well data from old wells exist in an area. Furthermore, if old wells have not been abandoned (liquidated), they can often be re-entered with measuring equipment and sampling tools, allowing reliable data about the subsurface to be obtained. This is a great opportunity for Croatian geothermal projects as many old wells can be revitalized and reused.

8.2 Economic Challenges

It is estimated that, in Croatia, the costs for the surface part of a geothermal power plant amount to around 1.5 million euros per MW. However, due to the high drilling costs and the aforementioned technical challenges involved in drilling such wells, costs can multiply, in some cases reaching up to 6 million euros per MW of installed capacity. For this reason, innovation centers, companies, and universities are developing new technologies aimed at accelerating and reducing the cost of the drilling process [28].

However, the biggest problem is precisely the unreliability of surface geological and geophysical data. This fact causes significant scepticism toward investing in the early exploratory phase of geothermal projects. As for the costs of above-ground (surface) equipment and infrastructure, these are usually comparable to those of similar projects [29], so they do not cause as much uncertainty. Therefore, the extremely high initial, i.e., capital

investment costs and the long payback period represent a challenge for the development of such projects. Banks are generally conservative investors and are usually willing to participate only in the production phase of a project, so such projects are often, at least at the early stage, financed from other sources, often own funds [28].

The profitability of investment in geothermal power plants could largely depend on the level of the premium price set by the Croatian Energy Market Operator (HROTE). With appropriate support, the investment payback period could be reduced to eight to ten years [28], which could motivate investors to invest in such projects. Alongside considering the price premium, it would be beneficial to review clear quotas for promoting production from geothermal sources, which could contribute to creating a more favorable investment environment and further strengthen the sense of regulatory predictability and security among investors [28]. Likewise, the development of risk-sharing schemes in the exploration phase (e.g., government guarantees or funds to cover unsuccessful drilling) could stimulate investments in the risky phases of projects [30]. Also worth mentioning are the 75 million euros that have been secured for the development of geothermal projects in Croatia through the European Regional Development Fund and the Recovery and Resilience Plan [25].

8.3 Regulatory, Social, and Other Challenges

Geothermal projects, unlike other renewable energy sources, are subject not only to the usual legislation (such as the Renewable Energy Sources Act, the Electricity Market Act, and the Construction Act), but also to the Hydrocarbons Exploration and Exploitation Act. This additional law regulates the exploration, evaluation, and exploitation of geothermal reservoirs, as well as the design and construction of the plant. As a result, geothermal energy investors must navigate two parallel administrative processes: one for the exploitation of geothermal water and another for the construction of the power plant. Such a system is rather complex and time-consuming, so unifying the regulations under a dedicated geothermal energy law could significantly facilitate project development. Additionally, cooperation with investors in drafting new laws and regulations would be desirable at an early stage, not only during public consultations [28].

The development of a geothermal project is a complex and lengthy process involving public tenders for exploration permits, the preparation of technical studies and economic assessments, approvals from expert commissions, environmental impact assessments, and obtaining numerous permits and energy approvals. All these phases prolong project timelines and increase risks for investors. Therefore, it is necessary to further simplify and accelerate procedures for geothermal projects, for example, by limiting the entire process to a maximum of two years and to more clearly define national policy toward this type of energy, especially regarding electricity generation capacities and grid access. It would also be beneficial to ease administrative obstacles, particularly in spatial planning procedures before the public tender [30].

The National Plan for the Development of Geothermal Energy by 2030 foresees increasing the share of renewable energy sources to at least 36.6%, reducing emissions, and developing geothermal energy for electricity generation as well as for heating and cooling. By 2030, the plan aims to install 68 MW of geothermal power plants [19]. The plan includes strengthening exploration activities, encouraging investment, building infrastructure, improving the regulatory framework, and involving local governments, with an emphasis on sustainable resource use and environmental protection. With its above-average geothermal potential, Croatia aims to accelerate its energy transition and strengthen supply security and energy independence by 2030 [31].

However, other challenges also arise, such as public mistrust due to environmental concerns (e.g., contamination of groundwater or gas emissions) [32] and issues related to property rights (land ownership, leases, compensation, etc.). Solutions to these problems include transparency with the public and all stakeholders, educational campaigns, and presentations of successful projects.

9 NEW GEOTHERMAL TECHNOLOGIES

As mentioned before, several companies have emerged that, through their innovations, aim to significantly change the approach to geothermal energy. Two such interesting examples are Eavor and Quaise. The Canadian company Eavor Technologies Inc. is developing advanced closed-loop systems, commercially known as Eavor-Loop™, which utilize a network of multilateral wells with complex geometries to efficiently transfer heat from the subsurface to a working fluid, with the goal of making geothermal energy scalable and applicable in nearly any location [33]. The American company Quaise Energy Inc. is developing a technology that, instead of conventional drilling, uses microwaves to melt and vaporize rock, with the purpose of enabling drilling into deeper and harder layers and accessing virtually unlimited geothermal resources. This approach could overcome the current limitations in the availability and scalability of geothermal sources [34].

10 CONCLUSION

By its nature, geothermal energy represents practically inexhaustible, clean, and renewable source of energy. In just the first 10 kilometers of the Earth's crust, there is 50 thousand times more energy potential than in all hydrocarbon reserves combined. This type of energy has been used by humans since ancient times, and at the beginning of the 20th century, it was used for the first time to generate electricity. However, despite this enormous potential, geothermal energy contributes very little to electricity production, both globally and in the Republic of Croatia. The reasons for this lie in numerous technical, economic, regulatory, and other challenges faced by geothermal projects. Technical problems could be solved through the development of innovative technologies and new, more modern drilling methods, which are already being researched by certain companies and institutions. Economic challenges will have to be addressed through incentivized feed-in tariffs, risk-sharing, and subsidies. Regulatory challenges should be mitigated through

amendments to regulations aimed at simplifying and accelerating administrative processes, which would also help to further reduce the risks of investing in such projects. In addition to all of this, there is a need to strengthen institutional support and apply a systematic approach in promoting geothermal energy, which could play a very important role in the green transition and in achieving energy independence for Croatia and other countries. Lastly, it is important not to overlook the potential role of oil and gas companies, whose technical expertise, extensive experience, and substantial capital could prove invaluable for the future research, development, and deployment of geothermal energy.

11 REFERENCES

[1] Igwe, C. I. (2021). Geothermal energy: A review. *International Journal of Engineering Research & Technology (IJERT)*, *10*(3), 420–431. https://www.ijert.org/research/geothermal-energy-a-review-IJERTV10IS030164.pdf

[2] Vacquier, V. (1998). A theory of the origin of the Earth's internal heat. *Tectonophysics*, *291*, 1–7. https://doi.org/10.1016/S0040-1951(98)00026-2

[3] Carney All Seasons. (n.d.). *The history of geothermal energy*. https://www.carneyallseasons.com/blog/geothermal/the-history-of-geothermal-energy/ (Accessed: June 4, 2025)

[4] GlobalData. (n.d.). *The oldest geothermal power plant: Larderello, Italy*. https://www.power-technology.com/features/oldest-geothermal-plant-larderello/ (Accessed: June 4, 2025)

[5] ThinkGeoEnergy. (n.d.). *Geothermal energy production & utilisation*. https://www.thinkgeoenergy.com/geothermal/geothermal-energy-production-utilisation/ (Accessed: June 4, 2025)

[6] Goldberg, V., Winter, D., Nitschke, F., Rath, M., Held, S., Spitzmüller, L., Budach, I., Pavez, M., Morata, D., Koschikowski, J., & Kohl, T. (2021). The potential of raw material extraction from thermal brines – Successful milestones of the BrineMine project. *Oil Gas European Magazine*, *47*, 26–33. https://doi.org/10.19225/210306

[7] DiPietro, J. A. (2013). Keys to the interpretation of geological history. In *Landscape evolution in the United States: An introduction to the geography, geology, and natural history* (pp. 327–344). Elsevier. https://doi.org/10.1016/B978-0-12-397799-1.00020-8

[8] Chamorro, C. R., García-Cuesta, J. L., Mondéjar, M. E., & Pérez-Madrazo, A. (2014). Enhanced geothermal systems in Europe: An estimation and comparison of the technical and sustainable potentials. *Energy*, *65*, 250–263. https://doi.org/10.1016/j.energy.2013.11.078

[9] Omwenga, B. M. (2020). Geothermal well site suitability selection using geographic information systems (GIS) and remote sensing: Case study of the Eburru geothermal field. In *Proceedings of the 45th Workshop on Geothermal Reservoir Engineering* (SGP-TR-216), Stanford University.

[10] Dynamic Graphics, Inc. (n.d.). *Visualizing 3D geothermal exploration data*. https://www.dgi.com/blog/visualizing-3d-geothermal-exploration-data/ (Accessed: June 4, 2025)

[11] Macenić, M. (n.d.). *Eksploatacija geotermalne energije revitalizacijom napuštenih naftnih i plinskih bušotina*. https://www.rgn.unizg.hr/hr/component/content/article/2643-eksploatacija-geotermalne-energije-revitalizacijom-napustenih-naftnih-i-plinskih-busotina (Accessed: June 4, 2025) (in Croatian)

[12] Coro, G., & Trumpy, E. (2020). Predicting geographical suitability of geothermal power plants. *Journal of Cleaner Production*, *267*, 121874. https://doi.org/10.1016/j.jclepro.2020.121874

[13] Yousef, B. A. A., Hachicha, A. A., Rodriguez, I., Abdelkareem, M. A., & Inyaat, A. (2021). Perspective on integration of concentrated solar power plants. *International Journal of Low-Carbon Technologies*, *16*, 1098–1125. https://doi.org/10.1093/ijlct/ctab034

[14] Sakai, Y. (2017). Advanced geothermal steam turbines. In *Advances in steam turbines for modern power plants* (pp. 455–486). Woodhead Publishing. https://doi.org/10.1016/B978-0-08-100314-5.00019-1

[15] Alagoz, E., AlNasser, F., Ozkan, Y., Dundar, E. C., & Shiriyev, J. (2023). Overview of closed-loop enhanced geothermal systems. *International Journal of Earth Sciences Knowledge and Applications*, *5*(1), 158–164.

[16] International Geothermal Association. (n.d.). *Geothermal energy database*. https://worldgeothermal.org/geothermal-data/geothermal-energy-database (Accessed: June 4, 2025)

[17] International Renewable Energy Agency (IRENA). (2017). *Geothermal power: Technology brief*. https://www.irena.org/-/media/Files/IRENA/Agency/Publication/2017/Nov/IRENA_Geothermal_Power_2017.pdf

[18] Gutiérrez-Negrín, L. C. (2024). Evolution of worldwide geothermal power 2020–2023. *Geothermal Energy*, *12*, 14. https://doi.org/10.1186/s40517-024-00290-w

[19] Croatian Chamber of Economy. (n.d.). *Geotermalna energija važan je faktor energetskog miksa, Hrvatska ima potencijal jedinstven u Europi i mora to iskoristiti na vrijeme*. https://www.hgk.hr/geotermalna-energija-vazan-je-faktor-energetskog-miksa-hrvatska-ima-potencijal-jedinstven-u-europi-i-mora-to-iskoristiti-na-vrijeme (Accessed: June 4, 2025) (in Croatian)

[20] Macenić, M., Kurevija, T., & Medved, I. (2020). Novel geothermal gradient map of the Croatian part of the Pannonian Basin System based on data interpretation from 154 deep exploration wells. *Renewable and Sustainable Energy Reviews*, *132*, 110069. https://doi.org/10.1016/j.rser.2020.110069

[21] Tumara, D., & Pavlović, D. (2019). Geotermalna energija i njezin potencijal u vremenu energetske diversifikacije i tranzicije Republike Hrvatske. *Nafta i Plin*, *39*(161–162), 53–62. (in Croatian)

[22] Lider Media. (n.d.). *Domitrović: Za realizaciju geotermalnih elektrana potrebna je premijska cijena*. https://lidermedia.hr/zeleno-i-digitalno/domitrovic-bez-financijskih-fondova-lokalne-zajednice-nece-moci-realizirati-toplinarske-projekte-161720/ (Accessed: June 4, 2025) (in Croatian)

[23] OIEH. (n.d.). *Croatia's first geothermal power plant Velika 1 has been officially opened*. https://oie.hr/en/croatias-first-geothermal-power-plant-velika-1-has-been-officially-opened/ (Accessed: June 4, 2025)

[24] Ministry of Economy. (2024). *Energy in Croatia 2023: Annual energy report (Energija u Hrvatskoj 2023: Godišnji energetski pregled)*. https://www.eihp.hr/wp-content/uploads/2024/12/Energija-u-HR-2023_WEB_novo.pdf (Accessed: June 4, 2025) (in Croatian)

[25] Ministry of Economy. (n.d.). *Ministar Šušnjar: Geotermalna energija ima potencijal da postane jedan od glavnih stupova energetske sigurnosti*. https://mingo.gov.hr/vijesti/ministar-susnjar-geotermalna-energija-ima-potencijal-da-postane-jedan-od-glavnih-stupova-energetske-sigurnosti/9973 (Accessed: June 4, 2025) (in Croatian)

[26] Liu, W., & Gao, K. (2017). Drilling technical difficulties and solutions in development of hot dry rock geothermal energy.

Advances in Petroleum Exploration and Development, *13*(1), 63–69. https://doi.org/10.3968/9456

[27] Hickson, C. J., Kumataka, M., Cotterill, D. K., & Huang, K. (2020). Geothermal drilling and completions in comparison to oil and gas practices in low temperature (<170°C) sedimentary basin reservoirs. *GeoConvention 2020*, Calgary, Alberta, Canada, May 11–13, 2020.

[28] Nacional. (n.d.). *Dragutin Domitrović: U Hrvatskoj je u zadnjih šest godina izdano 30 dozvola za istraživanje geotermalnih voda*. https://www.nacional.hr/dragutin-domitrovic-u-hrvatskoj-je-u-zadnjih-sest-godina-izdano-30-dozvola-za-istrazivanje-geotermalnih-voda/ (Accessed: June 4, 2025) (in Croatian)

[29] Erdogmus, B., Toksoy, M., Ozerdem, B., & Aksoy, N. (2006). Economic assessment of geothermal district heating systems: A case study of Balcova-Narlidere, Turkey. *Energy and Buildings*, *38*, 1053–1059. https://doi.org/10.1016/j.enbuild.2006.01.001

[30] Investor.hr. (2024). *Hrvatska ima veliki potencijal geotermalne energije*. https://investor.hr/2024/02/27/hrvatska-ima-veliki-potencijal-geotermalne-energije/ (Accessed: June 4, 2025) (in Croatian)

[31] Ministry of Economy and Sustainable Development. (2023). *Plan razvoja geotermalnog potencijala Republike Hrvatske do 2030. godine*. https://mingo.gov.hr/UserDocsImages/UPRAVA%20ZA%20ENERGETIKU/Plan%20razvoja%20geotermalnog%20potencijala%20Republike%20Hrvatske%20do%202030..pdf (Accessed: June 4, 2025) (in Croatian)

[32] European Parliament. (2023). *Nacrt izvješća o geotermalnoj energiji (2023/2111(INI))*. Odbor za industriju, istraživanje i energetiku, reporter: Zdzisław Krasnodębski. https://www.europarl.europa.eu/meetdocs/2014_2019/plmrep/COMMITTEES/ITRE/PR/2023/10-12/1285295HR.pdf (Accessed: June 4, 2025) (in Croatian)

[33] Eavor Technologies Inc. (n.d.). *Technology – The Eavor-Loop™*. https://eavor.com/technology/ (Accessed: June 4, 2025)

[34] Quaise Energy Inc. (n.d.). *Quaise technology*. https://www.quaise.energy/ (Accessed: June 4, 2025)

Authors' contacts:

Pavao Mesarić
Sveučilište Sjever,
Ulica 104. brigade 3 42000 Varazdin, Croatia
pamesaric@unin.hr

Lovorka Gotal Dmitrović, Associate Professor Dr.
(Corresponding author)
Sveučilište Sjever,
Ulica 104. brigade 3 42000 Varazdin, Croatia
lgotaldmitrovic@unin.hr

Boris Jalušić, Assistant Professor Dr.
Sveučilište Sjever,
Ulica 104. brigade 3 42000 Varazdin, Croatia
bjalusic@unin.hr

ISSN 1846-6168 (Print), ISSN 1848-5588 (Online)
https://doi.org/10.31803/tg-20.....

Categorization of paper
Received: y-m-d, Accepted: y1-m1-d1

14pt
14pt

Article Title Only in English (Style: Arial Narrow, Bold, 14pt)

14pt

Ivan Horvat, Thomas Johnson, Marko Marić (Style: Arial Narrow, Normal, 10pt)

14pt

Abstract: Article abstract contains maximum of 150 words and is written in the language of the article. The abstract should reflect the content of the article as precisely as possible. TECHNICAL JOURNAL is a trade journal that publishes scientific and professional papers from the domain(s) of mechanical engineering, electrical engineering, civil engineering, multimedia, logistics, etc., and their boundary areas. This document must be used as the template for writing articles so that all the articles have the same layout. (Style: Arial Narraw, 8pt)

8pt

Keywords: keywords in alphabetical order (5-6 key words). Keywords are generally taken from the article title and/or from the abstract. (Style: Arial Narraw, 8pt)

10pt
10pt

1 INTRODUCTION (Article Design)

(Style: Arial Narrow, Bold, 10pt)

10pt

(Tab 6 mm) The article is written in Latin script and Greek symbols can be used for labelling. The length of the article is limited to eight pages of international paper size of Letter (in accordance with the template with all the tables and figures included). When formatting the text the syllabification option is not to be used.

10pt

1.1 Subtitle 1 (Writing Instructions)

(Style: Arial Narrow, 10pt, Bold, Align Left)

10pt

The document format is Letter with margins in accordance with the template. A two column layout is used with the column spacing of 10 mm. The running text is written in Times New Roman with single line spacing, font size 10 pt, alignment justified.

Article title must clearly reflect the issues covered by the article (it should not contain more than 15 words).

Body of the text is divided into chapters and the chapters are divided into subchapters, if needed. Chapters are numbered with Arabic numerals (followed by a period). Subchapters, as a part of a chapter, are marked with two Arabic numerals i.e. 1.1, 1.2, 1.3, etc. Subchapters can be divided into even smaller units that are marked with three Arabic numerals i.e. 1.1.1, 1.1.2, etc. Further divisions are not to be made.

Titles of chapters are written in capital letters (uppercase) and are aligned in the centre. The titles of subchapters (and smaller units) are written in small letters (lowercase) and are aligned left. If the text in the title of the subchapter is longer than one line, no hanging indents.

10pt

Typographical symbols (bullets), which are being used for marking an item in a list or for enumeration, are placed at a beginning of a line. There is a spacing of 10pt following the last item:

- Item 1
- Item 2
- Item 3

10pt

The same rule is valid when items are numbered in a list:

1) Item 1
2) Item 2
3) Item 3

10pt

1.2 Formatting of Pictures, Tables and Equations

(Style: Arial Narrow, 10pt, Bold, Align Left)

10pt

Figures (drawings, diagrams, photographs) that are part of the content are embedded into the article and aligned in the centre. In order for the figure to always be in the same position in relation to the text, the following settings should be defined when importing it: text wrapping / in line with text.

Pictures must be formatted for graphic reproduction with minimal resolution of 300 dpi. Pictures downloaded from the internet in ratio 1:1 are not suitable for print reproduction because of unsatisfying quality.

10pt

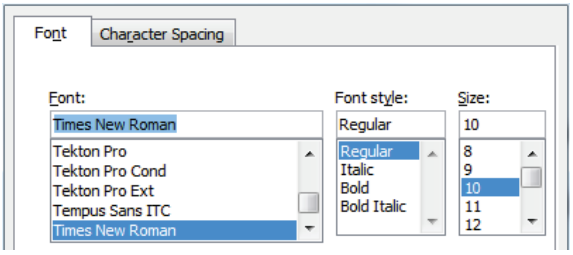


Figure 1 Text under the figure [1]
(Style: Arial Narraw, 8pt, Align Centre)

10pt

The journal is printed in black ink and the figures have to be prepared accordingly so that bright tones are printed in a satisfactory manner and are readable. Figures are to be in colour for the purpose of digital format publishing. Figures in the article are numbered with Arabic numerals (followed by a period).

Text and other data in tables are formatted - Times New Roman, 8pt, Normal, Align Center.

When describing figures and tables, physical units and their factors are written in italics with Latin or Greek letters, while the measuring values and numbers are written upright.

10pt

Table 1 Table title aligned centre
(Style: Arial Narraw, 8pt, Align Centre)

	1	2	3	4	5	6
ABC	ab	ab	ab	ab	ab	ab
DEF	cd	cd	cd	cd	cd	cd
GHI	ef	ef	ef	ef	ef	ef

10 pt

Equations in the text are numbered with Arabic numerals inside the round brackets on the right side of the text. Inside the text they are referred to with equation number inside the round brackets i.e. ".... from Eq. (5) follows ...." (Create equations with MathType Equation Editor - some examples are given below).

10pt

$$F_{\text{avg}}(t, t_0) = \frac{1}{t}\int_{t_0}^{t_0+t} F\left[q(\tau), p(\tau)\right] \mathrm{d}\tau, \tag{1}$$

$$\cos\alpha + \cos\beta = 2\cos\frac{\alpha+\beta}{2}\cdot\cos\frac{\alpha-\beta}{2}, \tag{2}$$

$$(\boldsymbol{AB})^{\mathrm{T}} = \boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{\mathrm{T}}, \tag{3}$$

$$AAMC = \frac{1}{n}\sum_{i=1}^{n} PVMC_i. \tag{4}$$

10pt

Variables that are used in equations and also in the text or tables of the article are formatted as *italics* in the same font size as the text.

10pt

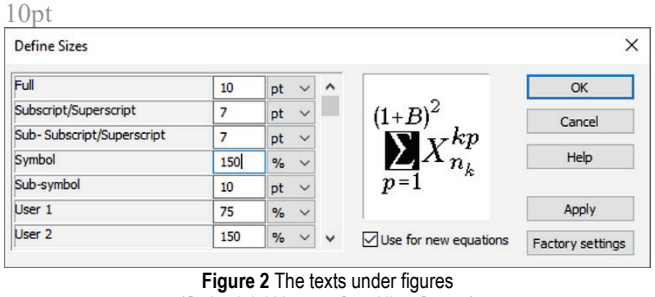


Figure 2 The texts under figures
(Style: Arial Narraw, 8pt, Align Centre)

10pt

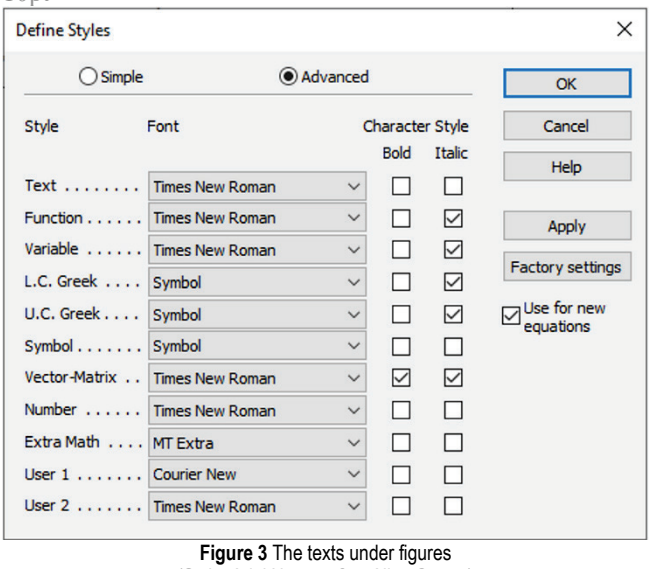


Figure 3 The texts under figures
(Style: Arial Narraw, 8pt, Align Centre)

Figures and tables that are a part of the article have to be mentioned inside the text and thus connected to the content i.e. " … as shown in Fig. 1…" or "data from Tab. 1…" and similar.

10pt

2 PRELIMINARY ANNOTATION

10pt

Article that is offered for publication cannot be published beforehand, be it in the same or similar form, and it cannot be offered at the same time to a different journal. Author or authors are solely responsible for the content of the article and the authenticity of information and statements written in the article.

Articles that are accepted for publishing are classified into four categories: original scientific papers, preliminary communications, subject reviews and professional papers.

Original scientific papers are articles that according to the reviewer and the editorial board contain original theoretical or practical results of research. These articles need to be written in such a way that based on the information given, the experiment can be repeated and the results described can be achieved together with the author's observations, theoretical statements or measurements.

Preliminary communication contains one or more pieces of new scientific information, but without details that allow recollection as in original scientific papers. Preliminary communication can give results of an experimental research, results of a shorter research or research in progress that is deemed useful for publishing.

Subject review contains a complete depiction of conditions and tendencies of a specific domain of theory, technology or application. Articles in this category have an overview character with a critical review and evaluation. Cited literature must be complete enough to allow a good insight and comprehension of the depicted domain.

Professional paper can contain a description of an original solution to a device, assembly or instrument, depiction of important practical solutions, and similar. The article need not be related to the original research, but it should contains a contribution to an application of known scientific results and their adaptation to practical needs, so it presents a contribution to spreading knowledge, etc.

Outside the mentioned categorization, the Editorial board of the journal will publish articles of interesting content in a special column. These articles provide descriptions of practical implementation and solutions from the area of production, experiences from device application, and similar.

10pt

3 WRITING AN ARTICLE

10pt

Article is written in the English language and the terminology and the measurement system should be adjusted to legal regulations, standards and the International System of Units (SI) (Quantities and Units: ISO 80 000 - from Part 1 to Part 14). The article should be written in third person.

Introduction contains the depiction of the problem and an account of important results that come from the articles that are listed in the cited literature.

Main section of the article can be divided into several parts or chapters. Mathematical statements that obstruct the reading of the article should be avoided. Mathematical statements that cannot be avoided can be written as one or more addendums, when needed. It is recommended to use an example when an experiment procedure, the use of the work in a concrete situation or an algorithm of the suggested method must be illustrated. In general, an analysis should be experimentally confirmed.

Conclusion is a part of the article where the results are being given and efficiency of the procedure used is emphasized. Possible procedure and domain constraints where the obtained results can be applied should be emphasized.

AI and AI-assisted tools do not qualify for authorship under TG/TJ's authorship policy. Authors who use AI or AI-assisted tools during the manuscript writing process are asked to disclose their use in a separate section of the manuscript. The publishing agreement process works as usual, with the authors keeping the copyright to their own work.

10pt

4 RECAPITULATION ANNOTATION

10pt

In order for the articles to be formatted in the same manner as in this template, this document is recommended for use when writing the article. Finished articles written in MS Word for Windows and formatted according to this template must be submitted using our The Paper Submission Tool (PST) (https://tehnickiglasnik.unin.hr/authors.php) or eventually sent to the Editorial board of the Technical Journal to the following e-mail address: tehnickiglasnik@unin.hr

The editorial board reserves the right to minor redaction corrections of the article within the framework of prepress procedures. Articles that in any way do not follow these authors' instructions will be returned to the author by the editorial board. Should any questions arise, the editorial board contacts only the first author and accepts only the reflections given by the first author.

10pt

5 REFERENCES (According to APA)

10pt

The literature is cited in the order it is used in the article. No more than 35 references are recommended. Individual references from the listed literature inside the text are addressed with the corresponding number inside square brackets i.e. "… in [7] is shown …". If the literature references are web links, the hyperlink is to be removed as shown with the reference number 8. Also, the hyperlinks from the e-mail addresses of the authors are to be removed. In the literature list, each unit is marked with a number and listed according to the following examples (omit the subtitles over the references – they are here only to show possible types of references):

9pt

[1] See http://www.bibme.org/citation-guide/apa/

[2] See http://sites.umuc.edu/library/libhow/apa_examples.cfm

[3] (Style: Times New Roman, 9pt, according to APA)

[4] Amidzic, O., Riehle, H. J. & Elbert, T. (2006). Toward a psychophysiology of expertise: Focal magnetic gamma bursts as a signature of memory chunks and the aptitude of chess players. *Journal of Psychophysiology, 20*(4), 253-258. https://doi.org/10.1027/0269-8803.20.4.253

[5] Reitzes, D. C. & Mutran, E. J. (2004). The transition to retirement: Stages and factors that influence retirement adjustment. *International Journal of Aging and Human Development, 59*(1), 63-84. Retrieved from http://www.baywood.com/journals/PreviewJournals.asp?Id=0091-4150

[6] Jans, N. (1993). *The last light breaking: Life among Alaska's Inupiat Eskimos*. Anchorage, AK: Alaska Northwest Books.

[7] Miller, J. & Smith, T. (Eds.). (1996). *Cape Cod stories: Tales from Cape Cod, Nantucket, and Martha's Vineyard*. San Francisco, CA: Chronicle Books.

[8] Chaffe-Stengel, P. & Stengel, D. (2012). *Working with sample data: Exploration and inference*. https://doi.org/10.4128/9781606492147

[9] Freitas, N. (2015, January 6). People around the world are voluntarily submitting to China's Great Firewall. Why? Retrieved from http://www.slate.com/blogs/future_tense/2015/01/06/tencent_s_wechat_worldwide_internet_users_are_voluntarily_submitting_to.html (Style: Times New Roman, 9pt, according to APA)

10pt

10pt

Authors' contacts:

8pt

Full Name, title
Institution, company
Address
Tel./Fax, e-mail

8pt

Full Name, title
Institution, company
Address
Tel./Fax, e-mail

Note: Gray text should be removed in the final version of the article because it is for guidance only.

Oberwolfach Meetings 2027

Mathematisches Forschungsinstitut Oberwolfach

Member of Leibniz Association

Workshops

Participants of the Workshops at Oberwolfach are invited personally by the Director of the Institute. Participation is subject to such an invitation. Interested researchers, in particular early career scientists, can contact the administration of the Institute. Since the number of participants is restricted, not all inquiries can be considered.

03.01. – 08.01.2027 ID: 2701

Moduli Theory in Differential Geometry and Analysis
B. Hanke (Augsburg), I. Alonso Lorenzo (Piscataway), C. Mantoulidis (Houston), M. Upmeier (Aberdeen)

10.01. – 15.01.2027 ID: 2702

Geometry of Motives
S. Gille (Edmonton), C. Ravi (Mumbai), P. Sechin (Regensburg), A. Vishik (Nottingham)

17.01. – 22.01.2027 ID: 2703

Wave-Kinetic Interfaces: Dispersive Dynamics Meets Kinetic Theory
M. Iacobelli (Zürich), N. Pavlovic (Austin), G. Staffilani (Cambridge MA), N. Visciglia (Pisa)

24.01. – 29.01.2027 ID: 2704

Geometric Methods in Data Science
M. Bauer (Tallahassee), X. Pennec (Sophia Antipolis), S. Sommer (København), G. Steidl (Berlin)

31.01. – 05.02.2027 ID: 2705

Non-commutative Function Theory: A Link between Operator Theory and Random Matrices
M. Banna (Abu Dhabi), G. Cébron (Toulouse), M. Hartz (Saarbrücken), J. E. McCarthy (St Louis)

07.02. – 12.02.2027 ID: 2706a

Mahler Measures and L-Functions
M. Lalín (Montréal), D. van Straten (Mainz), W. Zudilin (Nijmegen)

07.02. – 12.02.2027 ID: 2706b

Lagrangian Torus Fibrations
M. Abouzaid (Stanford), E. Mazzon (Paris), H. Ruddat (Stavanger)

14.02. – 19.02.2027 ID: 2707

Scientific Machine Learning: Mathematical Foundations and Algorithmic Advances
M. Han Veiga (Columbus), A. Heinlein (Delft), C. B. Schönlieb (Cambridge UK), J. Zech (Heidelberg)

21.02. – 26.02.2027 ID: 2708

Hypoelliptic Operators in Geometry
W. Bauer (Hannover), X. Ma (Tianjin/Paris), N. Savale (Dublin), Y. Wang (Baltimore)

28.02. – 05.03.2027 ID: 2709

Statistical Analysis of Black Box Models
P. Bartlett (Berkeley), S. Langer (Bochum), J. Rousseau (Paris), S. Wang (Berlin)

Proposals invited

Mini-Workshops

01.08. – 06.08.2027 (2731a/b)
deadline for appl.: 31.01.2027

08.08. – 13.08.2027 (2732a/b)
deadline for appl.: 31.01.2027

15.08. – 20.08.2027 (2733a/b)
deadline for appl.: 31.01.2027

22.08. – 27.08.2027 (2734a/b)
deadline for appl.: 31.01.2027

Time slots for four further Mini-Workshops in winter 2027 will be announced on our website in 2026.

Tandem- or Mini-Workshops

05.09. – 10.09.2027 (2736a/b)
deadline for appl.: 01.09.26/31.01.27

Oberwolfach Seminars

16.05. – 21.05.2027 (2720a/b)
to be announced in winter 2026

24.10. – 29.10.2027 (2743a/b)
to be announced in winter 2026

21.11. – 26.11.2027 (2747a/b)
to be announced in winter 2026

Arbeitsgemeinschaft (study group)

14.03. – 19.03.2027 (2711)
to be announced in spring 2026

10.10. – 15.10.2027 (2741)
to be announced in fall 2026

17.10. – 22.10.2027 (2742),
to be announced in fall 2026

07.03. – 12.03.2027 ID: 2710

Representations of Finite Groups
O. Dudas (Marseilles), R. Kessar (Manchester), B. Späth (Wuppertal), P. H. Tiep (Rutgers)

21.03. – 26.03.2027 ID: 2712

Homotopy Theory
J. Grodal (København), M. Hill (Minneapolis), B. Richter (Hamburg)

28.03. – 02.04.2027 ID: 2713

Generalized Computability Theory
J. P. Aguilera (Wien), N. Greenberg (Wellington), B. Löwe (Hamburg/Cambridge UK), M. I. Soskova (Madison)

04.04. – 09.04.2027 ID: 2714

Aperiodic Order: Spectral Theory, Harmonic Analysis and Dynamical Systems in Interaction
D. Damanik (Houston), D. Lenz (Jena), N. P. Frank (Poughkeepsie), R. Yassawi (London)

Check online!

11.04. – 16.04.2027 ID: 2715

Geometric, Algebraic and Topological Combinatorics
M. Juhnke (Osnabrück), E. Nevo (Jerusalem/Valladolid), F. Santos (Santander), V. Welker (Marburg)

18.04. – 23.04.2027 ID: 2716

Structure Preserving Methods for Computational Geometric Mechanics
E. Celledoni (Trondheim), S. Ober-Blöbaum (Paderborn), A. Palha (Delft), A. Stern (St Louis)

25.04. – 30.04.2027 ID: 2717

Mathematical Analysis of Fluid Dynamics Models
I. Kukavica (Los Angeles), A. Mazzucato (University Park), E. S. Titi (Cambridge UK), E. Wiedemann (Erlangen)

02.05. – 07.05.2027 ID: 2718

Statistical Physics and Random Surfaces
C. Garban (Lyon), J. Miller (Cambridge UK), S. Sheffield (Cambridge MA)

09.05. – 14.05.2027 ID: 2719

Complexity Theory
M. Bläser (Saarbrücken), I. Dinur (Princeton), R. O'Donnell (Pittsburgh), S. Vadhan (Harvard)

23.05. – 28.05.2027 ID: 2721

Interactions between Algebraic Geometry and Non-commutative Algebra
W. Crawley-Boevey (Bielefeld), O. Schiffmann (Orsay), S. Spenko (Bruxelles), M. Wemyss (Glasgow)

30.05. – 04.06.2027 ID: 2722a

Nonlinear Optics: Physics, Analysis, and Numerics
K. Busch (Berlin), R. El-Ganainy (St Louis), M. Hochbruck (Karlsruhe), C. Lasser (München)

30.05. – 04.06.2027 ID: 2722b

Methods of Mathematization around 1900
H. Junker (Bonn), D. Kent (St Andrews), J. Lorenat (Claremont)

06.06. – 11.06.2027 ID: 2723

Arithmetic of Shimura Varieties
K. Madapusi (Chestnut Hill), T. Richarz (Darmstadt), E. Viehmann (Münster), M. Zhang (Princeton)

13.06. – 18.06.2027 ID: 2724

Homological Commutative Algebra and Applications to Geometry
H. Brenner (Osnabrück), D. Eisenbud (Berkeley), E. Faber (Graz), C. Miller (Syracuse)

20.06. – 25.06.2027 ID: 2725

Recent Trends in Algebraic Geometry
G. Farkas (Berlin), M. Mustata (Ann Arbor), R. Pandharipande (Zürich), C. Voisin (Paris)

27.06. – 02.07.2027 ID: 2726

Mathematical Horizons in Artificial Intelligence
C. Boyer (Paris), M. Hein (Tübingen), R. Nowak (Madison), F. Yang (Zürich)

04.07. – 09.07.2027 ID: 2727

Differentialgeometrie im Grossen
O. Chodosh (Stanford), J. Lotay (Oxford), A. Lytchak (Karlsruhe), L. Wang (New Haven)

11.07. – 16.07.2027 ID: 2728

Dynamische Systeme
S. Crovisier (Gif-sur-Yvette), U. L. Hryniewicz (Aachen), T. M-Seara (Barcelona)

18.07. – 23.07.2027 ID: 2729

Discrete Nonlinear Optimization: Mathematics and Algorithms
J. Lee (Ann Arbor), R. Sotirov (Tilburg), A. Wiegele (Klagenfurt)

25.07. – 30.07.2027 ID: 2730

Partial Differential Equations
A. Fraser (Vancouver), T. Lamm (Karlsruhe), C. Mantoulidis (Houston), X. Ros-Oton (Barcelona)

29.08. – 03.09.2027 ID: 2735

Mathematical Aspects of General Relativity
C. Cederbaum (Tübingen), G. Holzegel (Münster), J. Luk (Stanford), T. Oliynyk (Clayton)

12.09. – 17.09.2027 ID: 2737

Homotopical Algebra and Higher Structures
T. Dyckerhoff (Hamburg), J. Holstein (Hamburg), M. Livernet (Paris), M. Markl (Prague)

19.09. – 24.09.2027 ID: 2738

Proof Complexity and Beyond
P. Beame (Seattle), M. Mahajan (Chennai/Mumbai), J. Nordström (Copenhagen/Lund), P. Pudlák (Praha)

26.09. – 01.10.2027 ID: 2739

Quantum Groups – Algebra, Analysis and Category Theory
S. Neshveyev (Oslo), D. Nikshych (Durham NH), J. Plavnik (Bloomington), A. Skalski (Warsaw)

03.10. – 08.10.2027 ID: 2740

Polytopal Finite Element Methods for PDEs: Theory, Algorithms, and Machine Learning Enhancement
P. Antonietti (Milano), L. Beirão da Veiga (Milano), J. Droniou (Montpellier), P. Wriggers (Hannover)

31.10. – 05.11.2027 ID: 2744

New Directions in Tomography: Theory and Applications
G. Ambartsoumian (Arlington), S. Arridge (London), B. Hahn-Rigaud (Stuttgart), A. Wald (Göttingen)

07.11. – 12.11.2027 ID: 2745

Applications of Optimal Transportation
G. Carlier (Paris), E. Radici (L'Aquila), F. Santambrogio (Lyon), B. Schmitzer (Göttingen)

14.11. – 19.11.2027 ID: 2746

Combinatorial Optimization
J. de Loera (Davis), V. Traub (Zürich), L. Végh (Bonn)

28.11. – 03.12.2027 ID: 2748

Applied Harmonic Analysis and Data Science
R. Alaifari (Aachen), A. Bandeira (Zürich), F. Krahmer (Garching), R. Ward (Austin)

05.12. – 10.12.2027 ID: 2749

Machine Learning in Infinite Dimensions
M. de Hoop (Houston), Y. Korolev (Bath), N. Mücke (Braunschweig), M. Thorpe (Coventry)

12.12. – 17.12.2027 ID: 2750

Convex Geometry and its Applications
S. Artstein-Avidan (Tel Aviv), D. Hug (Karlsruhe), E. Werner (Cleveland)

www.mfo.de/scientific-program

TEHNIČKI GLASNIK / TECHNICAL JOURNAL – GODIŠTE / VOLUME 20 – BROJ / ISSUE 4

PROSINAC 2026 / DECEMBER 2026 – STRANICA / PAGES 511-684

Sveučilište Sjever

SVEUČILIŠTE SJEVER / UNIVERSITY NORTH – CROATIA – EUROPE

ISSN 1846-6168 (PRINT) / ISSN 1848-5588 (ONLINE)

TEHNICKIGLASNIK@UNIN.HR – HTTP://TEHNICKIGLASNIK.UNIN.HR